

LES DONNÉES, MUSES ET FRONTIÈRES DE LA CRÉATION

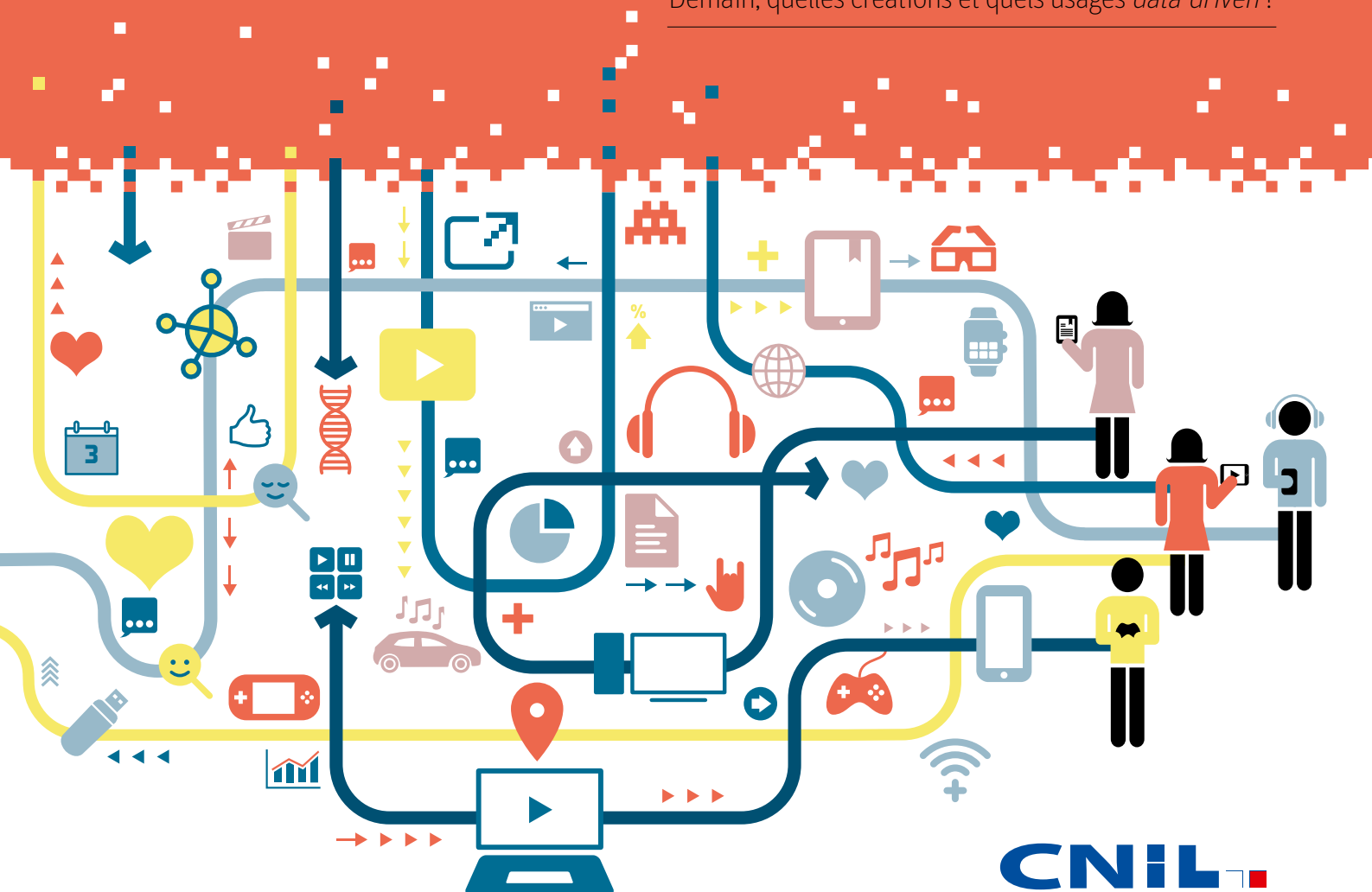
LIRE, ÉCOUTER, REGARDER ET JOUER
À L'HEURE DE LA PERSONNALISATION

Industries créatives, contenus numériques et données

Les contenus culturels vus au travers du prisme des données

Le graal de la recommandation et de la personnalisation

Demain, quelles créations et quels usages *data-driven* ?



Au sein de la Direction des technologies et de l'innovation de la CNIL, l'équipe Innovation & Prospective pilote des projets d'études et d'explorations de sujets émergents touchant aux enjeux liés aux données personnelles et à la vie privée, se situant dans ses travaux à l'interface entre innovation, technologies, usages, société, régulation et éthique.

La collection des cahiers IP, pour Innovation & Prospective, a vocation à présenter et à partager les travaux et études prospectives conduits par la CNIL. Il s'agit ainsi de contribuer à une réflexion pluridisciplinaire et ouverte dans le champ Informatique & Libertés et de nourrir les débats sur les sujets éthique dans le numérique.

L'équipe Innovation, Études et Prospective : Geoffrey Delcroix, Olivier Desbiey et Lucie Le Moine.

Commission Nationale
de l'Informatique et des Libertés
8 rue Vivienne – CS 30223
75083 Paris Cedex 02
Tél. : 01 53 73 22 22
ip@cnil.fr – www.cnil.fr



Retrouvez-nous sur notre site [www.cnil.fr/ip] en flashant le code ou sur :

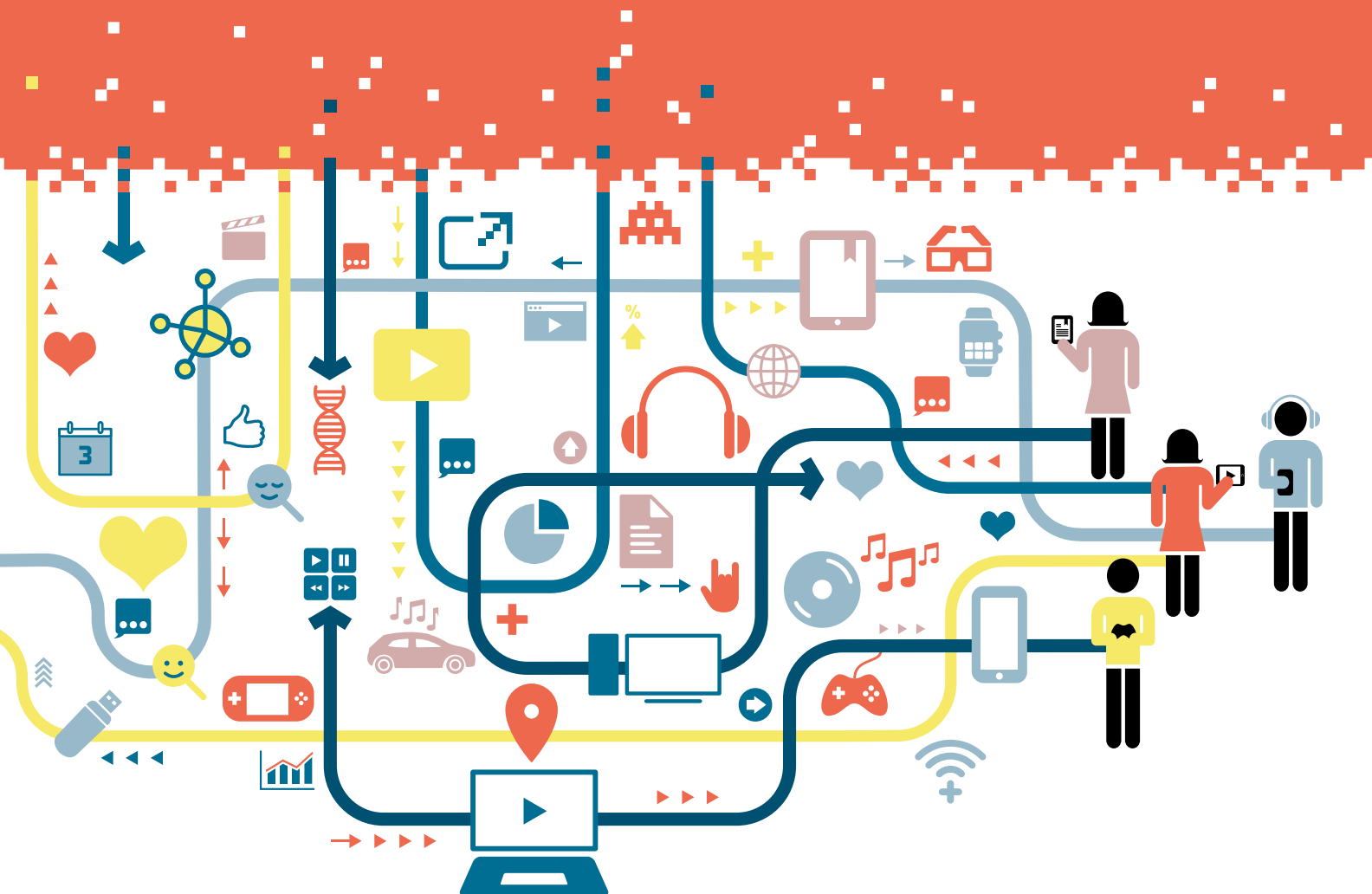


Édition annuelle / Directeur de la publication : Édouard Geffray / Rédacteurs de ce cahier : Geoffrey Delcroix, Olivier Desbiey, Lucie Le Moine avec l'aide de Marie Leroux, Stéphanie Chapelle, Mylène Podvin et Arthur Bouyx / Rédacteur en chef : Gwendal Le Grand / Conception graphique : EFIL 02 47 47 03 20 / www.efil.fr / Impression : Imprimerie Champagnac - 04 71 48 51 05 / Crédit Photos : IStock Photos, sauf p.13 Crédit Romuald Meigneux et p.55 Crédit André Gunthert / ISSN : 2263-8881 / Dépôt légal : à publication / Cette œuvre excepté les illustrations et sauf mention contraire est mise à disposition sous licence Attribution 3.0 France. Pour voir une copie de cette licence, visitez <http://creativecommons.org/licenses/by/3.0/fr/>



Les points de vue exprimés dans cette publication ne reflètent pas nécessairement la position de la CNIL.

La CNIL remercie vivement l'ensemble des experts interviewés ou ayant participé aux ateliers pour leur contribution.



ÉDITO

Les sondages nous rappellent régulièrement la place importante qu'occupent les pratiques culturelles dans la vie quotidienne des Français. Ces enquêtes ratent pourtant quelque chose d'essentiel. Elles sont impuissantes à nous révéler le rôle capital que joue la création culturelle dans notre existence aussi bien individuelle que collective. La lecture, la musique, les films et les séries mais aussi sans doute les jeux vidéo, bien au-delà du divertissement qu'ils nous procurent, ne sont-ils pas les lieux par excellence où ne cesse de s'élaborer et de se réinventer notre identité ? L'espace imaginaire qu'ils nous ouvrent n'est-il pas d'abord un laboratoire, un atelier où il nous est permis de confronter ce que nous pensons être à d'autres vies que les nôtres ; de nous projeter dans d'autres façons de penser, de sentir, d'aimer, en un mot, de conduire nos existences ? Pour le dire encore autrement, c'est au contact des œuvres que nous ne cessons d'inventer et de comprendre les parents, les amis, les citoyens, les héros, les amants que nous sommes ou aspirons à être.

Reproduites et diffusées par ailleurs massivement selon des logiques industrielles et économiques, supports de l'élaboration d'imaginaires sensibles et politiques partagés, les œuvres se situent donc au carrefour du plus collectif et du plus intime, au cœur de nos destins publics autant que personnels. Placée aux avant-postes de la révolution numérique qui bouleverse un à un tous les pans de nos sociétés, la Cnil ne pouvait manquer d'interroger les impacts de cette mise en données du monde sur ce point focal de la construction de nos identités que constitue l'univers de la création et de la distribution de contenus culturels.

Les évolutions sur ce front sont en effet spectaculaires et nos nouvelles pratiques et consommations deviennent, ici comme ailleurs, productrices de données personnelles. Les données alimentent des algorithmes de personnalisation et de recommandations. Compagnons de plus en plus indispensables pour naviguer dans l'immensité des catalogues de contenus, ces algorithmes peuvent tout autant favoriser la découverte et la diversité qu'enfermer les individus dans des goûts stéréotypés ou des horizons très limités. N'entrevoit-on pas ici le risque d'une création et d'une consommation affadies qui ne seraient plus l'occasion de la rencontre d'une altérité et d'un imaginaire singuliers mais la scène d'un pâle monologue indéfiniment rejoué ?

La discussion ne doit pas restreindre la question de la valeur apportée à la personne par l'utilisation des données personnelles à la seule gratuité proposée à l'utilisateur. Elle doit au contraire envisager ce qui fait la valeur irremplaçable de la consommation culturelle : l'élargissement du champ des possibles et des pensables.

Face à ces tendances, les citoyens semblent mutiques. Pas ou peu de débat : la collecte de données personnelles serait devenue le « mal nécessaire » de l'accès à certains services numériques innovants. Mais peut-on se résoudre à cette perspective ? Ne faut-il pas alors développer une culture de la donnée ►►

Partie 0.1

INDUSTRIES CRÉATIVES, CONTENUS NUMÉRIQUES ET DONNÉES

DÉMATÉRIALISATION DES CONTENUS ET
NOUVELLES PRATIQUES CULTURELLES **08**

FACE À LA NUMÉRISATION DES CONTENUS,
DE NOUVEAUX MODÈLES ÉCONOMIQUES
POUR LES INDUSTRIES CULTURELLES **12**

LES DONNÉES PERSONNELLES, AU CŒUR
DES INDUSTRIES CRÉATIVES ET CULTURELLES **16**

Partie 0.2

LES CONTENUS CULTURELS VUS AU TRAVERS DU PRISME DES DONNÉES

SPÉCIFICITÉS DES DONNÉES PERSONNELLES
CULTURELLES : UN ESSAI DE TYPOLOGIE **22**

DU LIVRE NUMÉRISÉ AUX ÉCRITURES
NUMÉRIQUES ET NOUVELLES EXPÉRIENCES
DE LECTURE **28**

DE LA PLAYLIST À LA CONTEXTUALISATION
ET À LA MUSICALISATION DU QUOTIDIEN **32**

DE LA VOD AUX CONTENUS ADAPTÉS
ET ADDICTIFS **36**

DES JEUX EN LIGNE AUX INTERACTIONS
HOMME-MACHINE AUGMENTÉES DANS
DES ENVIRONNEMENTS IMMERSIFS **42**

Partie 0.3

LE GRAAL DE LA RECOMMANDATION ET DE LA PERSONNALISATION

POURQUOI RECOMMANDER EST-IL
SI IMPORTANT ? **50**

COMMENT RECOMMANDE-T-ON ? **51**

MYTHES & RÉALITÉS DE LA RECOMMANDATION **52**

DE LA RECOMMANDATION À LA
PERSONNALISATION : QUELLE EXPÉRIENCE
POUR L'UTILISATEUR ? **56**

Partie 0.4

DEMAIN, QUELLES CRÉATIONS ET QUELS USAGES DATA-DRIVEN ?

QUELS SCÉNARIOS D'USAGE ET QUELS
MODÈLES ÉCONOMIQUES À 5 ANS ? **64**

HYPERMOVIE GENERATOR **66**

LIFE TRACKS **68**

TRAUMDATA **70**

THE COLLECTIVE MOOD MASTER **72**

DE QUOI CES SERVICES SONT-ILS LE NOM ? **74**

Annexes

76

01



Partie 0.1

INDUSTRIES CRÉATIVES, CONTENUS NUMÉRIQUES ET DONNÉES

DÉMATÉRIALISATION DES CONTENUS ET NOUVELLES PRATIQUES CULTURELLES	08
FACE À LA NUMÉRISATION DES CONTENUS, DE NOUVEAUX MODÈLES ÉCONOMIQUES POUR LES INDUSTRIES CULTURELLES	12
LES DONNÉES PERSONNELLES, AU CŒUR DES INDUSTRIES CRÉATIVES ET CULTURELLES	16

DÉMATÉRIALISATION DES CONTENUS ET NOUVELLES PRATIQUES CULTURELLES

► Ces industries culturelles et créatives ont connu la transformation numérique très tôt, et les usages comme les modèles économiques ont une certaine maturité, ce qui en fait un « laboratoire de la mise en données » de l'économie et des usages.

► Les écritures littéraires (romans, bande-dessinées...), les vidéos, la musique enregistrée et les jeux vidéo sont des secteurs aux caractéristiques proches et subissent des transformations numériques similaires dans la place que prennent les données dans la création, la distribution et la consommation des œuvres.

► Ces secteurs, sont tous marqués par des pratiques intenses et une dématérialisation croissante des contenus (qui représente déjà environ 40 % des revenus).

LES INDUSTRIES CRÉATIVES ET CULTURELLES

Comment définir un périmètre aux industries créatives et culturelles ? Le [Panorama des Industries Créatives et Culturelles en France](#), publié par EY (*Ernst & Young*) en novembre 2013, distingue 9 marchés culturels et créatifs (Arts graphiques et plastiques, Musique, Spectacle vivant, Cinéma/Télévision, Radio, Jeux vidéo, Livres, Presse/journaux et magazines) et illustre leur importance en termes de chiffres d'affaires et d'emplois pour l'économie française. Ces contenus peuvent être distribués sur une base unitaire (un film, un album, un livre, un jeu vidéo) et exploités dans la durée : un film sera par exemple exploité en salle, mais aussi par la diffusion télévisuelle, par la vente ou la location de DVDs, par la mise à disposition sur un catalogue de vidéos à la demande¹... La chaîne de valeur est donc durable et multiforme.

Ce cahier se concentre sur 4 secteurs particulièrement consommateurs et producteurs de données à caractère personnel :

- le livre ainsi que les écritures littéraires et de fiction (à l'exclusion donc des médias d'information) ;
- les œuvres vidéo et « d'images animées », particulièrement les marchés de la vidéo à la demande et de la vidéo délinéarisée² ;
- la musique, comprenant la création, la production, la distribution, l'écoute de musique enregistrée et, notamment l'émergence du marché de l'écoute en streaming³ ;
- le jeu vidéo qui recouvre aussi bien la production pour ordinateurs ou consoles de salon que le jeu sur smartphones et tablettes ou encore le jeu sur les réseaux sociaux.

La France conserve une position relativement privilégiée dans ces secteurs avec quelques géants internationaux dans le domaine des catalogues de contenus (*Universal Music Group, Hachette, Ubisoft...*) et quelques pépites sur les nouveaux marchés numériques (par exemple *Deezer* ou *DailyMotion*).

¹ Cette définition du périmètre est également utilisée par l'IDATE (IDATE, [Content Economics](#), septembre 2014).

² C'est-à-dire à la consommation au moment souhaité par le consommateur et non au moment choisi par le diffuseur.

³ Le streaming, ou diffusion en continu, désigne des techniques d'accès à des contenus (images, vidéos, sons...) à la demande permettant leur lecture « en direct ».

UNE PART IMPORTANTE DU « TEMPS DE CERVEAU DISPONIBLE » DES FRANÇAIS CONSACRÉE AUX CONTENUS CRÉATIFS ET CULTURO-LUDIQUES.

Les pratiques culturelles sont un champ d'investigation ancien⁴ du fait de leur importance dans le quotidien des Français : en moyenne, ils passent 9 heures par jour sur les différents supports et contenus⁵. Certes, la télévision est toujours l'écran devant lequel les Français passent le plus de temps, avec environ 3 heures par jour en moyenne, mais l'intensité des autres pratiques est loin d'être négligeable puisqu'en moyenne 1h45 par jour sont consacrées à jouer à des jeux vidéo, presque 1 heure à écouter de la musique, 47 minutes à lire des livres et 35 minutes à regarder des vidéos.

Bien évidemment, les composantes numériques dématérialisées sont chaque année en augmentation. Mais elles sont d'ores et déjà importantes : cette part va en France de 4% pour la vente de livres à 40% pour le jeu vidéo, en passant par environ 25% du chiffre d'affaires de la musique enregistrée et environ 20% de la consommation de vidéos⁶. La dématérialisation est en croissance dans tous les secteurs et elle constitue la part la plus dynamique de ces marchés. Selon l'IDATE, les contenus dématérialisés représentent aujourd'hui environ 40% des

revenus de ces secteurs - contre seulement 17% en 2010 ; ils devraient atteindre 63% en 2018⁷.

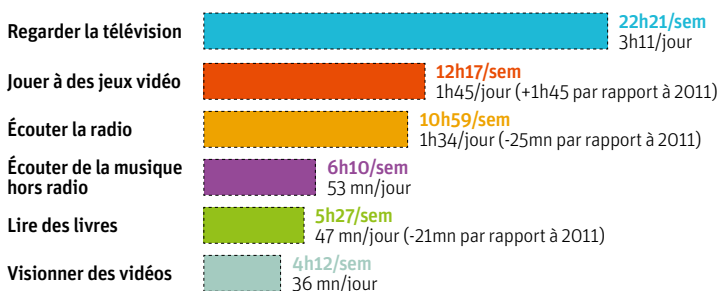
LE NUMÉRIQUE DÉVORE-T-IL LA CHAÎNE DE VALEUR DES INDUSTRIES CULTURELLES ?

Les modèles économiques du marché des contenus culturels ou ludiques ont déjà connu plusieurs bouleversements majeurs liés aux évolutions du financement de la création, à la préservation de la diversité culturelle, au partage de la valeur ajoutée entre les acteurs, à la propriété intellectuelle et au piratage. Est-ce à dire que les « barbares » numériques⁸ ont déjà mis à terre les modèles anciens des entreprises culturelles ? (voir chapitre I-3). C'est la position de Nicolas Colin dans la conférence « les barbares attaquent », concernant l'industrie du livre⁹ : « historiquement, depuis le début de la révolution numérique au milieu des années 90, les industries culturelles sont parmi les premières à voir leur modèle d'affaires, leurs façons de travailler, l'organisation de leur filière profondément transformés par le numérique ».

Qu'il s'agisse de la musique en *streaming*, de la vidéo à la demande, des jeux vidéo sur des plateformes en ligne ou encore du livre numérique, des « nouveaux entrants » ont en effet bouleversé les équilibres de marché dans chaque domaine, faisant tomber les barrières à l'entrée qui protégeaient les acteurs traditionnels. >>>

TEMPS HEBDOMADAIRE ALLOUÉ À LA CONSOMMATION D'ACTIVITÉS CULTURELLES ET CRÉATIVES PAR LES FRANÇAIS DE 10 ANS ET PLUS

(2012 par rapport à 2011)



Source : GfK - enquête 2012 auprès de 2305 répondants/EY, Panorama des Industries Créatives et Culturelles en France, novembre 2013.

⁴ MINISTÈRE DE LA CULTURE. [Les pratiques culturelles des Français](#), 2008. Ce dispositif d'enquête existe depuis les années 70 et sa dernière itération a été publiée en 2008.

⁵ EY. [Panorama des Industries Créatives et Culturelles en France](#), novembre 2013.

Ces chiffres sont évidemment basés sur des usages mesurés non "exclusifs" (d'autres activités ont lieu en même temps...).

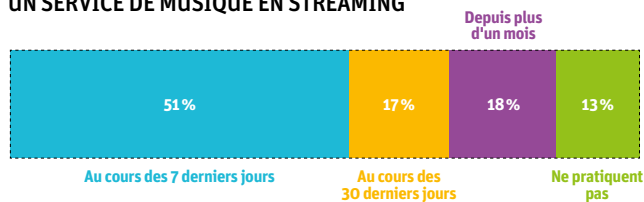
⁶ Essentiellement : EY (déjà cité), GfK, le Syndicat National de l'Édition, le Syndicat National de l'Édition Phonographique et le Syndicat National du Jeu Vidéo.

⁷ IDATE. [Content Economics](#), septembre 2014.

⁸ DUPONT-CALBO, Julien. « [Survivre aux "barbares" numériques](#) », *Les Echos*, 20 mai 2015.

⁹ Voir : <http://barbares.thefamily.co/le-livre> (Nicolas Colin, co-fondateur de The Family, Inspecteur des finances, et également Commissaire à la CNIL).

POURCENTAGE DES FRANÇAIS CONSOMMATEURS DE MUSIQUE NUMÉRIQUE AYANT UTILISÉ UN SERVICE DE MUSIQUE EN STREAMING



Source : Syndicat National de l'Édition Phonographique, édition 2015, p. 43 et Hyperworld.

MUSIQUE NUMÉRIQUE : LE *STREAMING*, MODE DE CONSOMMATION MAJORITAIRE

Alors qu'en 2011 le *streaming* rapportait quasiment deux fois moins de revenus que le téléchargement de musique selon le [Syndicat National de l'Édition Phonographique \(SNEP\)](#)¹⁰, ces deux modes de consommation représentent, aujourd'hui en France, un revenu équivalent. Selon [GfK](#)¹¹, le temps d'écoute de musique numérique dépasse celui de supports physiques et 60 à 70% des Français ont déjà utilisé plus ou moins régulièrement un service d'écoute de musique en *streaming*. D'ailleurs, les Français en sont parmi les plus avides utilisateurs : selon le SNEP, la part du chiffre d'affaires du *streaming* au sein des ventes numériques est deux fois plus importante en France qu'aux États-Unis, au Royaume-Uni ou en Allemagne (19% en moyenne). L'existence d'un leader français de ce marché (*Deezer*), et son offre préférentielle à destination des abonnés de l'opérateur *Orange*, expliquent certainement cette spécificité. De même, le fait que *Spotify* soit une création suédoise contribue au fait que le *streaming* compte pour 94% des revenus du secteur dans ce pays. Mais quoiqu'il en soit, la croissance du *streaming* s'observe au niveau global : en 2013, il a progressé de 39% pour représenter désormais plus du quart des revenus numériques mondiaux.

VIDÉO À LA DEMANDE ET TÉLÉVISION DE RATTRAPAGE, INCARNATIONS DE LA DÉLINÉARISATION DE LA CONSOMMATION DES FILMS ET SÉRIES

Si la consommation télévisuelle reste encore la reine des pratiques de consommation de contenus culturels et créatifs, la transformation numérique s'opère également dans ce secteur via l'apparition des téléviseurs

connectés, des offres de télévision de rattrapage et de l'importance des « seconds écrans » (smartphones et tablettes).

Deux tendances complémentaires sont observées. La première est la stratégie dite de l'« event TV »¹², qui se matérialise notamment dans la télévision sociale ayant recours aux services de réseaux sociaux pour impliquer le public et renforcer l'audience du programme en appelant aux réactions du public en direct (concours, votes dans les émissions de télé-réalité...). La seconde est la tendance croissante à la « délinéarisation » télévisuelle, c'est-à-dire à la consommation du média au moment souhaité par le consommateur et non au moment choisi par le diffuseur, à laquelle viennent s'ajouter les transformations des pratiques de consommation d'œuvres cinématographiques « à domicile ».

Selon un sondage réalisé par Médiamétrie fin 2014, environ un tiers des internautes français avait déjà regardé une vidéo à la demande¹³. C'est ainsi que venant largement succéder à la location de DVD, la consommation de vidéo à la demande s'installe rapidement, soit par l'intermédiaire d'acteurs de la télévision payante tels que le groupe *Canal+* ou la chaîne *HBO*, soit par l'intermédiaire de kiosques mis à disposition des téléspectateurs par des opérateurs de télécommunications sur leurs *boxes* ou par des nouveaux acteurs comme *Netflix*¹⁴, dont le service est devenu pour beaucoup synonyme de catalogue de vidéos à la demande.

LE JEU VIDÉO, UN MARCHÉ *BORN DIGITAL*¹⁵ EN BONNE SANTÉ, MAIS EN PLEINE MUTATION

Affichant un taux de croissance annuelle de 6 à 10% depuis 40 ans, l'industrie du jeu vidéo se porte bien et semble à contre-courant du contexte économique actuel des autres secteurs pour lesquels la dématérialisation intervient dans une situation générale de contraction des ventes. Cette industrie n'en est pas moins en pleine mutation ! Il y a 15 ans, les usages pouvaient se résumer à une utilisation domestique, par un public relativement jeune, de contenus sur support physique (CD-ROMs, DVD...). Aujourd'hui, le jeu vidéo est davantage intégré au sein du foyer avec plus de 28 millions de joueurs en France en 2011 (contre 11 millions

¹⁰ SYNDICAT NATIONAL DE L'ÉDITION PHONOGRAPHIQUE. [L'Économie de la production musicale](#), édition 2014.

¹¹ GfK. [Étude REC+](#), juillet 2013 (communiqué de presse).

¹² Voir : http://en.wikipedia.org/wiki/Event_television.

¹³ MEDIAMETRIE. [VOD 360](#), vague de décembre 2014. (Communiqué de presse).

¹⁴ L'offre de Netflix a été lancée en France en septembre 2014.

¹⁵ Document ou information créé dans une forme électronique et non numérisé *a posteriori*.

¹⁶ EY. [Panorama des Industries Créatives et Culturelles en France](#), novembre 2013, page 57.

en 2000) [selon le CNC cité par EY](#)¹⁶. Surtout, il concerne dorénavant toutes les tranches d'âges puisqu'en 10 ans, le nombre de joueurs de plus de 50 ans a doublé. Cette montée en popularité a manifestement été aidée par l'avènement du jeu mobile, car ce qui était autrefois dévolu à un usage spécialiste et sédentaire s'est popularisé au travers d'un jeu plus grand public et nomade. Le jeu correspond ainsi à 49% du temps passé par les utilisateurs sur des équipements mobiles¹⁷.

Désormais, deux modèles d'affaires coexistent : d'un côté, les grands studios développent des projets coûteux, nécessitant un investissement considérable en termes de marketing et les jeux sont de ce fait proposés à un prix élevé ; d'un autre côté, une multitude de studios (*Gameloft, Zynga, King...*) proposent des jeux *freemium*¹⁸ ou entièrement gratuits, avec des prétentions plus modestes, un potentiel viral¹⁹ plus élevé et une logique de monétisation très différente. Les enjeux émergents semblent quant à eux concerner principalement le *cloud gaming*²⁰, tandis que des possibilités encore mal définies de la réalité augmentée et des terminaux immersifs (comme le casque de réalité virtuelle *Oculus Rift*, acquis par Facebook pour près de 2 milliards de dollars en 2014²¹) font leur apparition.

DU LIVRE NUMÉRIQUE AUX CONTENUS ÉCRITS INNOVANTS

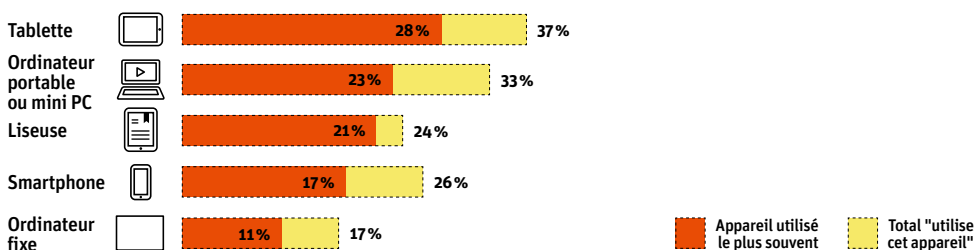
Alors que le monde de l'édition française a connu un recul de 1,3% de son activité en 2014²², le marché du livre numérique poursuit son développement avec un chiffre d'affaires en hausse

de 45% entre 2013 et 2014 en France, pour atteindre 64 millions d'euros. Selon [la Hadopi](#)²³, 11% des Français de plus de 15 ans lisent des livres numériques.

Si la liseuse a souvent été perçue comme le symbole du développement de la lecture numérique, elle n'est pourtant utilisée que par 24% des lecteurs, et principalement par les consommateurs de romans. Les smartphones et ordinateurs portables sont utilisés respectivement par 26% et 33% des lecteurs numériques. La tablette, privilégiée pour la lecture de bandes dessinées et mangas, arrive en tête, avec 37% d'utilisation totale.

Les acteurs de ce marché doivent donc penser des usages très concurrentiels de lecture sur terminaux²⁴ car contrairement à la liseuse, qui ne fait finalement que proposer la transposition simple du livre papier sur écran, les environnements mobiles, tactiles et foisonnants, proposent d'imaginer du contenu et des services innovants capables de maintenir le lecteur à l'intérieur de l'histoire. ■

APPAREILS UTILISÉS POUR LA LECTURE DE LIVRES NUMÉRIQUES



Source : HADOPI, (note d'analyse du Département Recherche, Études et Veille). [Livre numérique et terminaux de lecture, quels usages ?](#), 19 mars 2015.

¹⁷ SYNDICAT NATIONAL DU JEU VIDEO et IDATE. [Baromètre annuel du jeu vidéo en France](#), édition 2014.

¹⁸ Offre associant un accès gratuit en libre accès et un accès payant (*Premium*), plus haut de gamme. Il s'agit d'offrir « gratuitement » les fonctionnalités de base et de faire payer pour les fonctionnalités avancées du service.

¹⁹ Capacité d'un produit à se faire connaître plus ou moins rapidement sans publicité au sens traditionnel mais par le bouche à oreilles, par le biais des réseaux sociaux par exemple.

²⁰ C'est-à-dire le jeu à distance sans téléchargement.

²¹ Voir le [texte posté](#) par Mark Zuckerberg sur Facebook le 25 mars 2014.

²² GFK. [2014, une année en demi-teinte pour le marché du livre](#), 18 mars 2015.

²³ HADOPI, (note d'analyse du Département Recherche, Études et Veille). [Livre numérique et terminaux de lecture, quels usages ?](#), 19 mars 2015.

²⁴ Tablettes, ordinateurs et smartphones.

FACE À LA NUMÉRISATION DES CONTENUS, DE NOUVEAUX MODÈLES ÉCONOMIQUES POUR LES INDUSTRIES CULTURELLES

► Désormais en grande partie dématérialisée, l'offre de contenus a explosé, tout comme les moyens d'y accéder. Fini alors le modèle de la vente d'un produit physique et unique ? De nouveaux modèles économiques émergent, mobilisant les données personnelles des utilisateurs comme monnaie d'échange et moyen de personnaliser leur expérience du service.

► À terme, irons-nous jusqu'à la détermination dynamique des prix pour chaque utilisateur ?

« Les individus ne sont plus prêts à payer pour posséder des contenus, désormais seul l'accès compte. » C'est probablement l'adage que l'on retrouve le plus souvent lorsque l'on s'intéresse aux mutations induites par la numérisation des contenus dans les industries culturo-ludiques. Il renvoie implicitement à l'idée qu'avec le numérique, l'offre de contenus s'est considérablement accrue aussi bien en nombre qu'en variété, tout comme les moyens d'y accéder, légaux ou non, rendant obsolète le modèle traditionnel de distribution.

¹ On parle ici des biens culturels dans leur dimension divisible comme les livres, les disques... ce qui n'inclue pas le spectacle vivant qui est impacté différemment par le numérique et constitue souvent une forme de valeur refuge (ex. la musique).

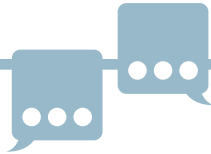
La remise en cause de ce modèle, basé sur la vente d'un produit physique ou unique, distribué par des intermédiaires et financé directement par le public, impacte naturellement le maillon le plus en amont de la chaîne, c'est à dire la création. Depuis la fin des années 90, on assiste ainsi à l'émergence de nouveaux modèles économiques hybridant le modèle traditionnel et la logique majoritaire du web, créant ainsi un marché publicitaire biface où les données constituent une monnaie d'échange permettant de créer des services toujours plus personnalisés.

Les difficultés rencontrées par les industries culturelles sont en partie le reflet de ces dilemmes et les modèles économiques des nouveaux acteurs font figure de parades. Deux tendances majeures se dégagent avec d'une part, la volonté de proposer une expérience et des services qui vont au-delà du simple contenu (qui peut souvent être accessible gratuitement), et d'autre part la personnalisation de cette expérience.

Pour l'utilisateur, la numérisation apporte aussi des effets ambigus combinant à la fois un accroissement de l'offre qui complique davantage le choix et une évolution des services numériques qui permet de consommer plus rapidement un grand nombre de contenus. Mais avoir accès à un catalogue de plusieurs millions de titres ne signifie pas pour autant qu'il est aisé d'avoir accès à du contenu pertinent. Les plateformes se différencient donc à travers le développement d'outils d'assistance à l'utilisateur, afin de l'aider à éclairer ses choix. Le recours à des données sur les caractéristiques de consommation de l'utilisateur (temps d'écoute, de lecture, goûts, intensité de consommation, etc. à découvrir dans la partie II) devient alors souvent nécessaire pour affiner et personnaliser le processus de découverte. La prescription en tant que terme générique pour désigner la recommandation (personnalisée ou non), qu'elle transite par des tiers (pairs, critiques, etc) ou par des algorithmes, découle de la nature même des biens culturels (voir partie III).

QUEL PRIX POUR LES PRODUITS CULTURELS

Le numérique a révélé ou modifié les propriétés économiques des biens culturels¹. Dès lors que les contenus se retrouvent disponibles et consommables dans un format numérique, ils deviennent indéfiniment reproductibles à un coût proche de zéro et sans perte de qualité. Ces propriétés font écho à la définition de bien public dans l'analyse économique. On parle de



3 QUESTIONS À... NICOLAS CURIEN, CONSEILLER AU CONSEIL SUPÉRIEUR DE L'AUDIOVISUEL



© Romuald Meigneux

Lors du MIP TV 2014, le président du CSA, Olivier Schrameck, évoquait la transformation du secteur audiovisuel en parlant notamment du « passage d'une logique d'offre (top down) centrée sur les choix de programmation de l'éditeur à une logique de demande (bottom up) fondée sur les préférences du téléspectateur ». Comment le CSA analyse-t-il ces changements dans les pratiques actuelles et futures des consommateurs ?

Audiovisuel et internet sont désormais intimement liés et leur entrelacement se traduit par de nouveaux usages. Ainsi, la croissance de la consommation des services non linéaires, notamment distribués directement sur internet, s'observe en particulier chez les plus jeunes ; cette tendance devrait vraisemblablement s'accroître dans les années à venir, avec l'essor soutenu des terminaux connectés : téléviseurs, tablettes et smartphones. Toutefois, la place de la télévision linéaire demeure encore prépondérante dans les habitudes et les usages des français, constituant à ce stade un support majeur d'exposition des contenus.

En tout état de cause, la transition numérique s'accompagne d'un basculement d'une logique d'offre vers une logique de demande. Ce mouvement porte en lui des promesses, notamment une meilleure adéquation aux attentes du public, mais aussi des risques,

notamment en termes de maintien de la diversité culturelle. À ce titre, la promotion de la diversité et la protection de la propriété intellectuelle doivent figurer parmi les politiques prioritaires des pouvoirs publics. Ces politiques devront toutefois être menées d'une manière qui soit adaptée à l'environnement numérique, c'est à dire qui ne bride pas la dynamique des technologies de l'information et de la communication, en elle-même porteuse d'un accroissement du bien-être social.

De quelle manière les offres et les services des acteurs du secteur évoluent-ils ?

La consommation de contenus vidéo compte désormais parmi les tout premiers usages sur internet. Les diffuseurs de contenus audiovisuels prolongent en conséquence leur pratique professionnelle et adaptent leur modèle économique, leur activité en ligne permettant d'accumuler un potentiel de captation d'audience. Le Big Data audiovisuel – c'est-à-dire l'analyse et l'exploitation des données personnelles issues de la consommation de contenus audiovisuels – constitue aujourd'hui l'un des principaux leviers de l'innovation en matière de diffusion des contenus culturels. Cette tendance de fond s'étendra à court terme à l'ensemble des acteurs, historiques et nouveaux, qui sont actifs sur la chaîne de valeur de l'audiovisuel à l'ère numérique.

Dans un tel contexte, il est indispensable que l'innovation s'inscrive dans une logique de promotion de la diversité culturelle, afin d'opérer pleinement au bénéfice des consommateurs. On doit à cet égard souligner que les nouveaux algorithmes de recherche, de recommandation, de prescription et de personnalisation ne sont pas nécessairement synonymes d'un appauvrissement du champ des propositions faites au consommateur. Au contraire, ces outils peuvent devenir de précieux

vecteurs d'orientation de la demande, facilitant certes la recherche de ce qui plaît déjà, mais suscitant aussi la découverte de ce qui est susceptible de plaire. À cet effet, les algorithmes ne devront pas uniquement procéder à l'analyse des choix passés du consommateur, mais également permettre à ce dernier d'accéder à des contenus inattendus, selon la logique dite de la « sérendipité ». Un tel mécanisme hybride correspond d'ailleurs aux us du téléspectateur-internaute, devenu « consommateur » de contenus et déjà coutumier d'une navigation bimodale, faisant appel, tantôt à des liens forts pour l'exploitation du connu, tantôt à des liens faibles pour l'exploration de l'inconnu.

Selon vous, quelles seront les tendances émergentes les plus intéressantes à observer au cours des prochaines années ?

Tout d'abord, le lancement de nouveaux services non linéaires, et plus particulièrement d'offres de vidéo à la demande par abonnement, accessibles directement sur internet via des plateformes de distribution OTT, devrait stimuler le développement de nouvelles offres sur le marché.

On peut également tabler sur l'enclenchement d'un cercle vertueux en matière d'innovation de programmes, s'agissant des contenus et des formats, avec, par exemple, l'écriture participative ou encore la combinaison de différents médias : télévision, radio, internet (cross-média). Ce mouvement de « modernisation » de la production et de la diffusion audiovisuelles ne sera évidemment pas cantonné au seul périmètre des nouveaux acteurs et il transformera aussi les chaînes « classiques » de télévision. Le consommateur bénéficiera ainsi de cette variété accrue de l'offre, quels que soient ses modes de consommation de contenus audiovisuels.

bien public lorsqu'il n'est pas possible d'exclure des individus de sa consommation et qu'il n'existe pas de « rivalité » dans l'accès². Le fichier numérique devient ainsi support et véhicule des contenus.

En économie, rattacher un bien à ces propriétés permet de mettre en évidence les problématiques d'incitations et de financement : pourquoi les individus paieraient-ils pour un bien qu'ils pourraient obtenir « gratuitement » ? Comment définir un prix de marché ? Qui va financer la production de biens que personne ne souhaite acheter ?

Les propriétés des contenus culturels ont conduit à l'origine à des fixations de prix « expérimentales »³ : les prix des versions dématérialisées étaient déduits du prix des biens physiques similaires (un album acheté en MP3 coûtant entre la moitié et les deux-tiers du prix d'un CD, un film à la demande coûtant à peu près le même prix qu'un DVD loué et à peine moins cher qu'un DVD acheté, etc). Progressivement, des modèles alternatifs sont apparus. La fixation du prix unitaire étant délicate et les utilisateurs achetant l'usage plus que la propriété, les modèles les plus explorés sont désormais les abonnements ou la consultation gratuite financée par la publicité⁴.

DES MODÈLES VARIÉS ET HYBRIDES : CONTENU GRATUIT FINANCÉ PAR LA PUBLICITÉ, ABONNEMENT ILLIMITÉ, PAIEMENT À L'UNITÉ...

Le schéma le plus courant pour accompagner l'évolution des pratiques de consommation consiste à proposer des plateformes qui reposent sur des modèles dits *freemium*. Ce schéma est particulièrement courant dans la musique en ligne où des services comme *Deezer* ou *Spotify* permettent l'écoute gratuite de titres mais avec des restrictions auxquelles les abonnés payants ne sont pas soumis. Par exemple, sur *Spotify*, l'utilisateur gratuit ne peut pas stocker de morceaux pour bénéficier d'une écoute hors connexion et il est par ailleurs exposé à des publicités qui s'intercalent dans sa liste d'écoute. En quelque sorte, il s'agit de fournir un service dégradé aux utilisateurs gratuits pour les inciter à opter pour une formule payante.

Le modèle payant par abonnement ou de paiement à l'unité est quant à lui particulièrement répandu pour les services de livres numériques et de vidéo à la demande. Plus

marginale, l'acceptation de publicités permet parfois d'accéder à une liseuse à un prix préférentiel comme dans le cas du *Kindle* d'*Amazon*, qui peut s'acheter à un prix moins élevé lorsque l'acheteur choisit la version sur laquelle des publicités personnalisées s'affichent sur l'écran de veille.

Les plateformes de distribution de contenus basent leur modèle soit sur des ventes à l'unité (modèle surtout présent dans la VOD) soit sur un accès payant et illimité à un catalogue de contenus (modèle dominant dans le *streaming musical*) assorti d'un ensemble de services ayant vocation à personnaliser l'expérience de l'utilisateur et sur lesquels la concurrence fait rage. Il peut s'agir de vente de tickets de concerts associée aux artistes écoutés ou plus généralement de réductions sur des produits pertinents dans un même univers de consommation. La valeur d'un contenu culturel n'étant connu qu'après sa consommation, offrir aux utilisateurs la possibilité d'avoir un aperçu (bande annonce, extraits) auquel ils pourraient accéder d'un côté, et conserver leur historique de consommation de l'autre, sont des mécanismes très répandus pour optimiser la rencontre d'un contenu avec une audience.

ILLUSTRATION SECTORIELLE : LE JEU VIDÉO

Selon le [Baromètre annuel](#) du jeu vidéo en France 2014⁵, le marché de la consommation de jeux vidéo a crû de presque 300 % en 10 ans.

Ce marché est organisé verticalement avec des développeurs, des studios de jeu, des éditeurs, des distributeurs, et enfin des fabricants de consoles qui contrôlent la chaîne de valeur. Mais récemment, l'apparition des jeux sur smartphones (et donc l'arrivée d'*Apple* et de *Google*) a réduit cette chaîne, et la dématérialisation *quasi* complète du marché a bouleversé les équilibres, permettant l'essor des jeux en ligne (par exemple les jeux sociaux⁶).

De par la multiplicité des supports et des utilisateurs (comparativement aux autres industries des biens culturels), le secteur des jeux vidéo est désormais celui qui offre le plus de variété en matière de modèles économiques. À une extrémité du spectre se trouve le jeu « gratuit » principalement présent sur mobile et tablette, et de l'autre côté se trouve le jeu payant à l'unité qui est plutôt la norme sur console⁷. Entre les deux, il existe un ensemble [de déclinaisons dont le freemium constitue la médiane](#)⁸ :

² Par exemple, le fait qu'un individu soit protégé n'empêche pas les autres de l'être avec la même qualité.

³ IDATE. [Content Economics](#), septembre 2014.

⁴ Après une phase d'exploration des techniques de « *bundling* » où le support numérique était vendu avec le support physique.

⁵ Syndicat national du jeu vidéo et IDATE. [Baromètre annuel du jeu vidéo en France](#), édition 2014.

⁶ Communications & stratégies n°94 " [Video game business models and monetization](#) ", juin 2014.

⁷ Les jeux pour console semblent être l'un des derniers bastions où les standards propriétaires permettent la persistance d'un marché de vente physique à l'unité.

⁸ HADOPI/IFOP. [Le jeu vidéo dématérialisé](#), décembre 2014.

- Abonnement Payant : mensuel pour jouer en ligne ;
- AYCE : « All You Can Eat » abonnement illimité sur un catalogue de jeux en ligne ;
- Freemium (cf. supra) ;
- F2P : « Free 2 Play » jeu gratuit avec micro-paiement pendant le jeu ;
- P&PM : « Pay and Play More » achat unitaire en ligne avec micropaiement.

La variété de modèles est rendue possible par l'engagement⁹ des joueurs qui offre aux éditeurs l'accès à de nombreuses données comportementales et la possibilité de s'adapter à la typologie des jeux et des joueurs (voir encadré dans chapitre II-5). En d'autres termes, il s'agit d'exploiter au mieux leurs dispositions à payer. Par ailleurs, la gratuité n'est souvent qu'apparente car à la manière du modèle des applications mobiles (voir les résultats du projet CNIL-Inria "[Mobilitics](#)"), les données des utilisateurs, notamment leur géolocalisation, sont souvent utilisées pour affiner le ciblage publicitaire qui n'est rien d'autre que la véritable contrepartie du jeu.

DEMAIN, UN PRIX DIFFÉRENT POUR CHAQUE CONSOMMATEUR GRÂCE AU BIG DATA ?

Le point commun à tous ces marchés est sans doute « l'économie des hits » (*winner takes all*) : que ce soit les *best sellers* littéraires, les *blockbusters* hollywoodiens, les hits musicaux ou les jeux vidéo AAA¹⁰ (nom donné aux plus grosses productions de l'industrie du jeu vidéo), il y a beaucoup d'appelés et peu d'élus, mais ces élus sont extraordinairement rentables. Les clés de ces succès tiennent essentiellement au professionnalisme, à l'instinct ou au talent des participants à ces productions. Il tient aussi de plus en plus à la capacité à utiliser les données pour détecter ce qui marche, voire imaginer une sorte d'algorithme du succès. À cet égard, Netflix n'a jamais raté une occasion de souligner que le choix de produire la série à succès *House of cards* était un succès du *big data* (voir chapitre II-4).

D'autres modèles économiques peuvent-ils apparaître demain ? Cela est très probable du fait que le marché dispose, grâce aux données, de toutes les clés pour explorer d'autres voies de détermination des prix des contenus culturels et créatifs dématérialisés. IDATE¹¹ imagine ainsi une hypothèse de détermination dynamique du prix.

Aujourd'hui, les champions de la gestion dynamique des prix¹² sont les professionnels des transports (compagnies aériennes ou même... Uber, qui a fait de la variation en temps réel des prix une des clés de voûte de sa proposition de service). Amazon est également très largement en mesure de faire varier rapidement les prix des produits vendus.

Pour autant le prix d'un livre, d'un album musical ou d'un jeu vidéo change de manière assez rigide, avec au maximum deux ou trois prix différents lors de son cycle de vie de vente. Mais la dématérialisation permet beaucoup plus de flexibilité. Il est aujourd'hui imaginable de faire varier en temps réel le prix d'un produit en fonction de la demande et donc de la popularité d'un contenu. Le jeu vidéo est déjà relativement avancé dans cette voie, par exemple en proposant des accès exclusifs en avance, des éditions « de luxe »... qui sont en fait autant de moyens de segmenter la demande selon le « consentement à payer »¹³, que les vendeurs seront capables de deviner par leur connaissance de l'acheteur qui se présente sur leur plateforme. Le ciblage en ligne ne sera alors plus seulement publicitaire mais permettra bien de définir algorithmiquement un prix pour chaque consommateur. Si cette hypothèse paraît très prospective, la Maison Blanche l'a jugé suffisamment crédible pour rendre public un rapport s'intéressant à la capacité du *big data* à influencer sur les rapports de force entre acheteurs et vendeurs en ligne¹⁴. Dans ce rapport, la Maison Blanche cite incidemment une étude de Benjamin Shiller qui estimait que si Netflix utilisait des données comportementales pour personnaliser les prix de son abonnement, les profits supplémentaires pourraient atteindre jusqu'à +12%¹⁵. L'équipe de prospective de France Télévision pose d'ailleurs la question dans [son cahier de tendances Méta-média numéro 9](#) : « et si demain, les prix de la culture devenaient variables ? Comme ceux des billets d'avion ou des chambres d'hôtels. Le *yield management* et la gestion des prix en temps réel pourraient ainsi faire varier les prix en fonction de l'utilisateur, de ses goûts, du moment de découverte et de consommation, de l'écran utilisé. Et pourquoi pas de la qualité ? »¹⁶ ■

⁹ Cette expression est utilisée ici, et dans le reste du présent cahier, en se référant à son utilisation dans les domaines du marketing, du commerce et de la publicité dans lesquelles cette notion est devenue cardinale. L'engagement se définira comme une interaction significative entre un consommateur et un contenu, un service, un produit, un média. Ce concept est la conséquence directe de l'appropriation par le monde du marketing et du commerce de la prédiction du « [Manifeste des évidences](#) », texte fondateur datant de 1999 : « les marchés sont des conversations ».

¹⁰ En référence aux notes scolaires dans le système nord-américain.

¹¹ IDATE. [Content Economics](#), septembre 2014.

¹² Également appelée « [yield management](#) », c'est-à-dire la tarification différenciée à des fins d'optimisation des ventes.

¹³ Ce que les économistes appellent de la discrimination par les prix « de troisième type » (https://fr.wikipedia.org/wiki/Discrimination_par_les_prix).

¹⁴ WHITE HOUSE. [Big Data and differential pricing](#), février 2015.

¹⁵ Netflix a aujourd'hui une stratégie de prix unique, et rien n'indique que l'entreprise ne compte pas changer cette stratégie à brève échéance, ne serait-ce que parce que sa dernière modification de grille tarifaire lui avait valu une levée de boucliers des abonnés aux États-Unis. Ces stratégies sont donc très dangereuses en termes d'image... lorsqu'elles sont visibles !

¹⁶ SCHERER, Eric. « [La TV demain : 10 enjeux de transformation](#) », *métablog* #9, juin 2015.

LES DONNÉES PERSONNELLES, AU CŒUR DES INDUSTRIES CRÉATIVES ET CULTURELLES

- ▶ Les modèles d'affaires du secteur des contenus culturels et créatifs sont de plus en plus enrichis en données personnelles.
- ▶ De nouveaux acteurs économiques internationaux sont devenus les symboles de ces évolutions : Amazon pour les livres, Spotify (et Deezer) pour la musique, Netflix pour la vidéo à la demande et Steam pour les jeux vidéo.
- ▶ Le recours croissant aux données est donc une tendance transversale des industries culturelles et créatives mais pour rendre quel service, et surtout, avec quel alignement d'intérêt entre les utilisateurs et les firmes ?
- ▶ En quoi cet enrichissement en données modifie-t-il la nature de la relation contractuelle et de la relation de confiance entre les utilisateurs et les diffuseurs ?

Historiquement, les questions de protection des données personnelles et de la vie privée ne pesaient pas lourd dans cette industrie. Mais les choses changent ! Le présent Cahier IP se propose donc d'explorer dans quelle mesure la numérisation des contenus culturels produits, distribués, consommés impacte la capacité à « enrichir en données personnelles » les modèles d'affaires.

La véritable plus-value des services émergents de distribution de contenus numériques culturels et ludiques se concentre aujourd'hui dans leur capacité à traiter les données de consommation et d'habitude des utilisateurs, afin de les accompagner.

Les services se trouvent dans une position privilégiée pour cibler les individus et leur proposer des recommandations pertinentes en cumulant plusieurs types d'informations. À travers ces possibilités, la collecte de données devient un enjeu majeur de différenciation concurrentielle et la recommandation personnalisée sera peut-être le juge de paix qui distinguera les acteurs qui prennent le contrôle de la chaîne de valeur et ceux dont le rôle sera moins important. Comme dans beaucoup de secteurs, des acteurs « disruptifs » sont devenus, à juste titre ou à grand renfort de *storytelling*, les symboles de cette transformation de l'équilibre des pouvoirs de marché. Pour une fois, il ne faut pas nécessairement parler des GAFAM (Google, Amazon, Facebook, Apple, Microsoft). Certes, tous sont déjà des géants dans le domaine : Google avec YouTube et Google Music, Facebook avec les jeux en ligne, Apple bien évidemment pour la musique avec iTunes et le rachat de Beats, Amazon sur l'ensemble des marchés et Microsoft sur les jeux vidéo mais aussi la musique et les films avec Xbox et Xbox store. Mais à l'exception (très notable) d'Amazon, ils n'ont pas lancé cette transformation-là.

LES SYMBOLES DE LA TRANSFORMATION DATA-CENTERED DES CONTENUS CULTURELS : AU-DELÀ DES GAFAM, LES ASNS (AMAZON, SPOTIFY, NETFLIX, STEAM).

Amazon a lancé cette transformation en en faisant la pierre angulaire de son offre de livres numériques dans sa gamme de liseuses Kindle. Amazon collecte en effet un grand nombre d'informations lors de l'utilisation de l'appareil et de la lecture des ouvrages électroniques. L'ensemble des fabricants de liseuses a d'ailleurs adopté des pratiques similaires à des degrés différents¹. Par exemple, certains éditeurs peuvent mettre en avant les passages les plus souvent surlignés ou sauvegardés sur des encarts publicitaires. Le *Wall Street Journal* avait décrit ce paysage dès 2012² : « un lecteur moyen met juste 7 heures pour lire le livre final de la trilogie *Hunger Games* de Suzanne Collins sur la liseuse Kobo, soit à peu près 57 pages par heure.

¹ ELECTRONIC FRONTIER FOUNDATION. *E-Reader Privacy Chart*. 2012 Edition.

² ALTER, Alexandra. "Your E-Book Is Reading You", *Wall Street Journal*, 19 juillet 2012.



Environ 18 000 lecteurs sur *Kindle* ont surligné la même ligne du deuxième livre de la série (...) et sur le *Nook* de *Barnes & Noble*, la première chose que font la plupart des lecteurs dès qu'ils ont achevé le premier tome de *Hunger Games* est de télécharger le suivant.»

Côté vidéo à la demande, l'exemple *Netflix* est tout aussi parlant dans la mesure où l'entreprise affirme elle-même que 75% des contenus visionnés par ses clients vient d'une recommandation personnalisée de son moteur³. Que cette affirmation soit ou non surestimée, *Netflix* a tout intérêt la mettre en avant puisque le premier objectif est de permettre au spectateur de trouver facilement un programme qui lui convient dans l'immense catalogue. La donnée permet donc au diffuseur de tenir la main du consommateur dans ses pérégrinations au cœur d'une offre foisonnante, afin d'assurer un taux d'utilisation important du service et donc de réduire l'attrition des abonnements (le « *Churn*⁴ »). Mais ces informations peuvent également être valorisées pour faire des choix éditoriaux, ou informer sur les programmes que l'entreprise produit.

Concernant la musique, des services comme *Deezer* ou *Spotify* sont en mesure de faire des recommandations de plus en plus pertinentes à partir d'un ensemble de données sur les morceaux, croisées avec les données de consommation et les données de réseaux sociaux. Ces acteurs souhaitent maintenant tirer parti de toutes ces informations pour les partager avec des tiers de manière agrégée et aider les producteurs à faire de la musique « qui plaît »⁵. Demain, ces services utiliseront probablement aussi des

données de géolocalisation ou d'autres capteurs du smartphone pour deviner si l'utilisateur est au travail, en voiture, à une soirée... et lui fournir des recommandations améliorées en fonction du contexte (voir partie IV).

Enfin, le domaine du jeu vidéo n'échappe pas à cette tendance. Sans même parler du jeu sur smartphone, par nature connecté et très gourmand en données⁶, le fait que les joueurs (aussi bien sur console que sur ordinateur) soient dorénavant connectés sur des plateformes comme *Steam*, *Xbox live* ou *PlayStation network*, permet la collecte d'informations pendant que le client joue. Là aussi, l'usage de ces données semble pour le moment limité à de l'analyse d'usage (par exemple pour mesurer si la difficulté est bien équilibrée ou non sur certains passages du jeu) et à de la recommandation ou de la personnalisation de certaines offres commerciales (par exemple, achats de contenus additionnels personnalisés). Mais demain, de véritables tests de comportement à l'intérieur même de jeux pourraient voir le jour, afin d'analyser et mesurer les réactions de la personne.

DATA-BASED SERVICES VS. BUNDLE

Le recours aux données est donc une tendance transversale des industries culturelles (voir partie II).

Si les données permettent sans aucun doute de rendre un service à valeur ajoutée, elles peuvent avoir vocation à remplir des objectifs bien distincts :

- s'assurer qu'il s'agit du bon utilisateur, qu'il dispose des droit d'accès au contenu, ►►

³ SMITH, Chris. « [How Spotify, Netflix and Amazon control your online habits](#) », *Techradar*, 18 janvier 2014.
⁴ Le *churn* est un anglicisme qui renvoie tout simplement au taux d'abandon du service, au taux de désabonnement.
⁵ SHUBBER, Khadim. « [Music analytics is helping the music industry see into the future](#) », *The Guardian*, 9 avril 2014.
⁶ Comme le montrent les [résultats du projet CNIL-Inria « Mobilities »](#)



que le contenu proposé est de bonne qualité... mais aussi en allant plus loin, d'améliorer le service rendu par exemple par la personnalisation de celui-ci en fonction des goûts, de l'historique de consommation, etc.

- construire des connaissances à valeur ajoutée, par exemple pour déterminer des modèles d'habitudes de consommation qui peuvent permettre de mieux catégoriser les utilisateurs (en particulier pour des annonceurs potentiels). Ces informations peuvent concerner un utilisateur identifié ou identifiable ou au contraire concerner des ensembles de données anonymes (par exemple sur les statistiques de comportement).

Autrement dit, certaines données utilisées sont nécessaires pour rendre le service quand d'autres vont permettre de construire une connaissance qui pourra bénéficier au fournisseur de services, à des tiers ou à l'utilisateur final⁷. Ces formes de participation des utilisateurs à la création de richesse des acteurs économiques s'apparentent à ce que certains chercheurs appellent le *digital labor*⁸. Cette situation est déjà troublante lorsque l'utilisateur gratuit d'un service « paye » ainsi la plus-value d'utilité « offerte » par le service, mais elle l'est encore plus du fait que les véritables « clients » du service sont souvent des annonceurs et publicitaires.

Pour autant, la collecte de données est-elle plus limitée lorsque l'on paye ? Probablement pas vraiment, comme le laisse penser certaines annonces sur des services et produits nouveaux (par exemple les annonces de Spotify en août 2015 concernant sa nouvelle politique de confidentialité et les éventuelles transmissions

de données à des tiers à des fins marketing ou commerciales). À l'adage souvent répété « si c'est gratuit, c'est vous le produit », il faut peut-être comme le propose Antonio Casilli ajouter « si vous ne payez pas, c'est que vous êtes le travailleur, l'ouvrier du service » mais aussi peut-être « et même si c'est payant, vous serez quand même le produit et/ou le travailleur ».

En outre, l'information fournie aux utilisateurs est-elle plus transparente, plus complète et plus explicite dans la mesure où celui-ci paye directement pour le service ? Puisque le service propose des recommandations, on pourrait penser que l'utilisateur a peut-être davantage conscience du fait que ses données en sont à l'origine. Pour autant, les mécanismes permettant la recommandation sont rarement détaillés par les fournisseurs du service.

Enfin, une autre problématique à mi-chemin entre des considérations « informatique et libertés » et de droit de la concurrence consiste à s'interroger pour déterminer si les grandes plateformes telles qu'*Amazon*, *Netflix*, *Deezer* ou *Spotify* fournissent un service d'accès à du contenu personnalisé – qui induirait l'exploitation de données pour la personnalisation – ou un ensemble de services packagés dans une offre groupée (bundle⁹) potentiellement séparables les uns des autres¹⁰. En droit de la concurrence, il s'agirait de s'intéresser au(x) périmètre(s) du ou des marché(s) pertinent(s) considéré(s). Dans une optique « protection des données », il pourrait s'avérer nécessaire que les services proposent notamment des mécanismes permettant de paramétrer la personnalisation (avec possibilité de s'opposer à l'analyse ou la combinaison des données de consommation par exemple, voire d'envisager un recueil préalable du consentement). ■

⁷ Les exemples d'*Amazon* utilisant les analyses des lectures sur son *Kindle* pour définir les stratégies de promotion des livres démontrent déjà qu'elles ne sont pas utilisées que dans le cadre de l'amélioration du service à l'utilisateur.

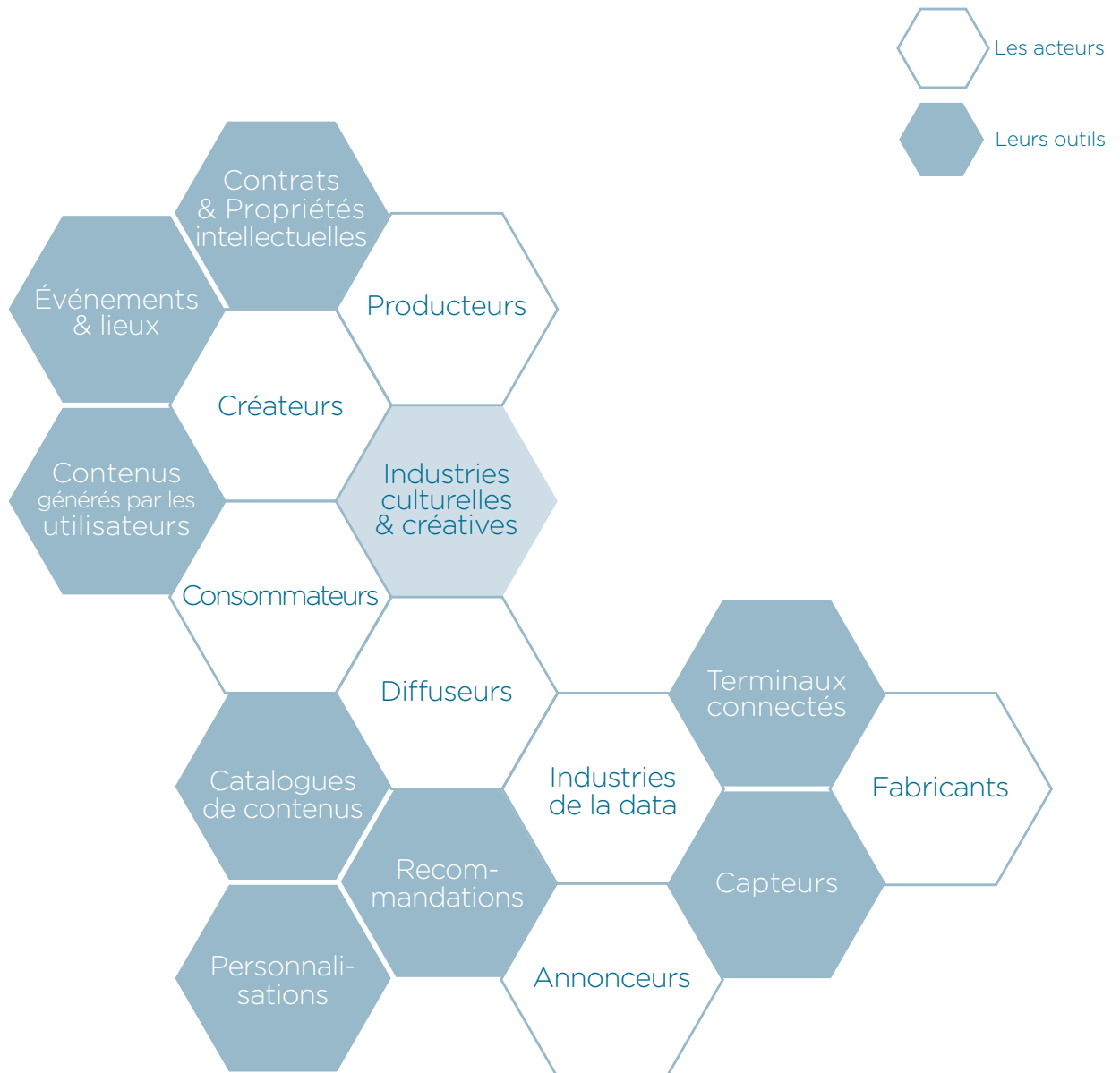
⁸ Expression utilisée par exemple par Antonio Casilli citée dans : GUILLAUD, Hubert. « *Digital Labor : comment répondre à l'exploitation croissante du moindre de nos comportements* », *Internet Actus*, 20 décembre 2014 (voir encadré Partie III).

⁹ Une *offre groupée* (*bundle* en anglais) est un ensemble de produits vendus ensemble dans une seule offre complète (par exemple, les abonnements internet + télévision + téléphone).

¹⁰ On pourrait tout à fait imaginer un service de *SVOD* fonctionnel sans moteur de recommandation et donc sans analyse des contenus consommés par exemple.

CARTOGRAPHIE DES INDUSTRIES CULTURELLES & CRÉATIVES

Quelle place pour les acteurs émergents dans l'écosystème des industries culturelles & créatives ?



02



Partie 0.2

LES CONTENUS CULTURELS VUS AU TRAVERS DU PRISME DES DONNÉES

SPÉCIFICITÉS DES DONNÉES PERSONNELLES CULTURELLES : UN ESSAI DE TYPOLOGIE	22
DU LIVRE NUMÉRISÉ AUX ÉCRITURES NUMÉRIQUES ET AUX NOUVELLES EXPERIENCES DE LECTURE	28
DE LA PLAYLIST À LA CONTEXTUALISATION ET À LA MUSICALISATION DU QUOTIDIEN	32
DE LA VOD AUX CONTENUS ADAPTÉS ET ADDICTIFS	36
DES JEUX EN LIGNE AUX INTERACTIONS HOMME-MACHINE AUGMENTÉES DANS DES ENVIRONNEMENTS IMMERSIFS	42

SPÉCIFICITÉS DES DONNÉES PERSONNELLES CULTURELLES : UN ESSAI DE TYPOLOGIE

► **Les données personnelles collectées et traitées dans le contexte des contenus culturels et créatifs ont des particularités liées au caractère très intime de la relation entre une personne et les œuvres : les pratiques culturelles contribuent à la construction de la personnalité et de l'identité.**

► **Dans ce domaine, les data-scientists sont en mesure d'utiliser des données d'apparence très anodine, mais très particulières : profils, descriptions des contenus, popularité, enrichissement, goûts et contexte.**

► **Ces données sont « au cœur de toutes les attentions des acteurs du big data » (EY, 2013).**

► **Les gisements de valeur pour les data sciences diffèrent d'un secteur à l'autre : les informations que l'on peut inférer de l'écoute musicale, des lectures, du visionnage de films ou séries ou de l'utilisation de jeux vidéo sont très différentes.**

Quelles sont les données personnelles dans les consommations et pratiques culturelles ? Fondamentalement, pourquoi ce domaine mérite-t-il une attention particulière, alors que des données sont collectées en permanence, par exemple lors de notre navigation sur les sites de e-commerce ?

Les internautes, s'ils ne comprennent pas toujours les mécanismes intimes de la traçabilité en ligne sont conscients qu'elle existe. La pédagogie sur les cookies n'y est pas étrangère, ni l'impression confuse de ciblage publicitaire que ressent tout internaute exposé à des bannières plus ou moins bien personnalisées. Personne ne semble donc étonné que le même type de *tracking* existe dans le domaine des données personnelles en lien avec les consommations de contenus créatifs et culturels.

EY a même cherché à promouvoir la notion de « données personnelles culturelles », dans son étude « [Comportements culturels et données personnelles au cœur du Big data](#) », publiée en 2013 pour le Forum d'Avignon, pour mettre en avant la spécificité des données du domaine. EY y définit les « données personnelles culturelles » comme « l'ensemble des données permettant de connaître les préférences culturelles des consommateurs, par exemple : discussions dans des forums, réseaux sociaux, avis et commentaires sur des œuvres/artistes... » et les distingue des données publiques culturelles¹ ainsi que des données personnelles plus traditionnelles (transactionnelles et marketing).

¹ Le document d'EY reprend la définition donnée par le « [guide Data culture](#) » du Ministère de la Culture : donnée produite ou détenue par un établissement, organisme ou service culturel ayant une activité culturelle réelle et effective. Ces données sont celles qui ont a priori vocation à être incluses dans une démarche d'ouverture des données publiques (*open data*).

² Définition de l'article 2 de la loi 78-17 du 6 janvier 1978 modifiée, dite « Informatique et Libertés » : « toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement ».

³ Hugo Bon, Co-fondateur et CEO de [Soundytics](#), a proposé d'analyser les différents types de données et métadonnées présentes dans les fichiers musicaux. Nous nous inspirons de cette typologie. BON, Hugo. [Encadrement des pratiques et des métadonnées : les nouveaux enjeux de la musique en ligne](#). Mémoire Master 2 - Droit du Numérique, Administration & Entreprises Université Paris 1 – Panthéon Sorbonne, mai 2012 (travail lui-même inspiré, entre autres de BOUTON, Rémi. [L'enjeu des métadonnées musicales. La musique n'est pas qu'un fichier son](#), 4 avril 2011 pour l'IRMA).

BENCHMARK DE DONNÉES : LA DONNÉE PERSONNELLE CULTURELLE SE DISTINGUE

	Données publiques culturelles	Données transactionnelles	Données marketing	Données personnelles culturelles
	★ ★ ★	★ ★ ★	★ ★ ★	★ ★ ★
Degré de digitalisation	Faible	Moyen	Fort	Moyen
Multiplicité des sources	Fort	Moyen	Moyen	Fort
Variété de format	Fort	Faible	Moyen	Fort
Volume de données	Faible	Fort	Moyen	Fort
Intimité	Faible	Moyen	Moyen	Fort

Source : EY©

TPOLOGIE DES DONNÉES

Sans vouloir créer une typologie définitive, quelles sont finalement les grands types de données liées à nos pratiques de consommation (ou création) de contenus créatifs et culturels ?

Pour autant, distinguer ainsi une catégorie générale « données personnelles culturelles » apporte-t-il une réponse suffisante ? En réalité, données transactionnelles (liées aux actes d'achat) et données marketing (de contact de qualification du client) sont toutes des données personnelles². Mais dans un « monde Big data », les traces et métadonnées prennent une importance croissante. Cet agrégat de données personnelles est comparable à la partie immergée d'un iceberg : peu visible et beaucoup plus importante parce que ces données sont abondantes mais aussi parce qu'elles sont souvent plus faciles à analyser et plus riches en interprétations possibles. La distinction créée par EY est donc utile, bien qu'elle crée une séparation relativement artificielle.

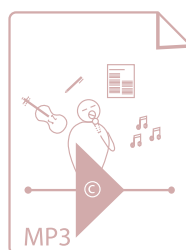
Sans vouloir créer une typologie définitive, quelles sont finalement les grands types de données liées à nos pratiques de consommation (ou création) de contenus créatifs et culturels³ ? Des données de profil, des données décrivant les contenus du catalogue, des données enrichissant ces contenus, des informations concernant la popularité, d'autres concernant les goûts de chaque utilisateur, des données de contexte (voir ci-contre).

Ces données peuvent provenir de sources diverses puisque certaines seront créées très en amont, par les créateurs, éditeurs, producteurs ou les sociétés de gestion (telles que la SACEM), d'autres sont enrichies par les diffuseurs. Elles peuvent en outre être produites algorithmiquement par des entreprises tierces ou créées par les utilisateurs eux-mêmes volontairement (*tags*, *playlists*, notes, commentaires).

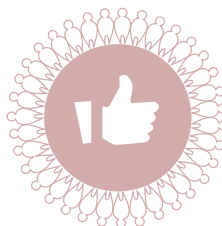
Les pratiques culturelles contribuent à construire la personnalité de l'individu. Elles sont donc communément vues comme très intimes. En 1988, suite à la divulgation dans la presse de la liste des locations de vidéos d'un candidat à la Cour Suprême, le Congrès des États-Unis avait ainsi édicté une des lois les plus protectrices de la vie privée (avec celles concernant les mineurs et concernant la santé de l'arsenal juridique américain). Dans le texte accompagnant ce *Video Privacy Protection Act*, Al McCandless, signataire de la première proposition, soulignait qu'il « y a un sentiment profond que les gens devraient être en mesure de lire des livres et regarder des films sans que le monde entier soit au courant. »



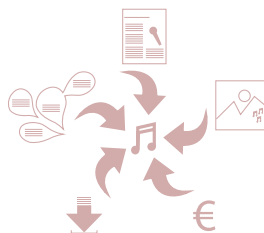
Des données personnelles au sens le plus classique : données d'identité, de contact et de **PROFIL** sociodémographique, qui sont finalement les données d'un fichier client et de gestion de la relation client traditionnel.



Des données **DESCRIPTIVES DES CONTENUS** : données de catalogage (artiste, auteur, interprète), de caractérisation (durée, genre, sous-genre) mais aussi données techniques (format, compression, échantillonnage) et données juridiques (concernant les droits de diffusion par exemple).



Des informations générales de consommation, c'est-à-dire renseignant sur les goûts collectifs comme la **POPULARITÉ** d'une chanson, le nombre total d'écoutes, les commentaires et citations sur les services de réseaux sociaux, etc.



Des données d'**ENRICHISSEMENT** : photos ou biographie des artistes, critiques, notes et évaluations, liens de téléchargement, prix, paroles de chansons. Ces données peuvent être fournies par des professionnels (par exemple les critiques) ou générées par les utilisateurs.



Des données concernant l'usage, les comportements et les **GOÛTS** de chaque utilisateur : ces données peuvent être très générales (type, quantité, durée, rythme d'achat/de consultation, playlists créées...) mais aussi très fines (les passages surlignés d'un livre, la vitesse de lecture...).



Des données de **CONTEXTE**, comme l'horodatage, la localisation ou toutes les données issues de capteurs (les mouvements, les émotions, l'état physiologique, l'humeur, etc).

Livres et films sont les vitamines intellectuelles qui alimentent la croissance de la pensée individuelle. L'ensemble de ce processus de maturité intellectuelle est foncièrement privé – un processus de calme et de réflexion. Ce processus intime devrait être protégé contre l'intrusion perturbante d'un regard inquisiteur⁴.

EY soulignait aussi en 2013 dans l'étude déjà citée que la consommation culturelle est un acte très intime et très engageant : « bien plus que la consommation de biens et services matériels, qui satisfont les besoins fondamentaux et primaires, celle de biens et services culturels en dit beaucoup sur nos préférences, nos centres d'intérêt et nos aspirations. Elle touche ainsi à notre identité, non seulement dans ce qu'elle a de plus intime, mais aussi dans ce qu'elle a de plus social et communautaire : les biens et services culturels que nous consommons, ou nos comportements culturels, participent en effet largement de l'image que nous souhaitons renvoyer dans notre environnement social. C'est pourquoi nous partageons volontiers notre expérience et notre avis sur un livre, un film, un jeu, un concert ou une exposition avec nos amis ou communautés en ligne, tout comme avec les éditeurs de contenus en ligne. La donnée personnelle culturelle renferme des informations contextuelles fortes et permet de qualifier de façon assez précise le pouvoir d'achat de l'être numérique mais aussi de prévoir son comportement. C'est pourquoi, en raison de sa valeur intrinsèquement élevée, elle figure aujourd'hui au centre de toutes les attentions des acteurs du big data. »

Les métadonnées, en apparence très anodines, elles sont en réalité très riches en informations sur le profil de l'individu. En outre, elles permettent d'interpréter ou de prédire le comportement, l'état d'esprit voire la psychologie réelle ou supposée des utilisateurs. Ceci est d'autant plus vrai qu'en réalité, il n'existe pas une grande catégorie de données personnelles culturelles mais tout un archipel de données très variées : chaque type de contenu créatif a des « propriétés » très fortes et distinctives, qui permettent des usages et analyses eux aussi très variés. Le UX designer⁵ Brett Goldstein, a ainsi [imaginé à quoi](#) pourrait ressembler des usages riches des données disponibles pour des services musicaux en *streaming* comme *Spotify*⁶. Dans un entretien avec les auteurs de ce cahier, Brett Goldstein expliquait que les données musicales sont assez

spécifiques, car « les chansons sont relativement courtes (2 à 4 minutes), et ont une très forte variabilité. Les livres sont par exemple beaucoup plus longs, prennent beaucoup plus de temps ».

À cet égard, ce sujet se rapproche de celui traité dans le cahier IP numéro 2 de la CNIL « [Le corps, nouvel objet connecté ?](#) »⁷, dans lequel nous explorions l'usage croissant de capteurs connectés capables de mesurer des constantes d'activité, de bien-être ou de santé, pouvant parfois sembler anodines de prime abord (nombre de pas, poids, cycles du sommeil). Certes, pris isolément, le nombre de pas d'une personne paraît insignifiant ; mais le fait de croiser de nombreuses données avec cette information sur une période de temps de plusieurs mois ou années devient autrement intrusif. Cela permet entre autres d'apprécier l'état de santé présent, voire futur de l'utilisateur.

Nous avons voulu étendre cette analyse pour explorer les spécificités des données des 4 grands secteurs de la création culturelle.

Dans le domaine musical les contenus sont donc très nombreux et variés, non seulement dans l'offre mais aussi dans la consommation. [Spotify indique ainsi](#) qu'en moyenne ses utilisateurs écoutent 1 200 chansons par an⁸. La variabilité des contenus est, elle aussi, relativement grande d'une unité de consommation (la chanson par exemple) à l'autre (styles, etc). En revanche, l'engagement pourra parfois être faible dans la mesure où il est facile de tester une nouvelle musique : un utilisateur de service en *streaming* acceptera relativement facilement d'écouter une chanson qui ne correspond pas a priori à ses goûts et préférences principaux... Autre particularité, la valeur sociale de la musique est très élevée, car l'adhésion à des types de musiques est souvent un marqueur d'appartenance à des communautés ou tribus, et de projection vers les autres de ses valeurs. La musique est un media d'adhésion et d'appartenance. À cet égard, dans le cadre de sa consommation musicale, l'individu peut adopter des stratégies de partage public sophistiquées et actives à travers les réseaux sociaux, à l'instar de ce qui existe concernant les photos (voir à ce sujet l'étude CNIL « [La place des photos dans la vie numérique](#) »)⁹. Ainsi, le fan de musique « exigeante » partagera facilement sa dernière découverte avant-gardiste, qui vient renforcer sa stature de découvreur de talent, en revanche,

⁴ *The Video Privacy Protection Act of 1988*, [sur le site de la Library of Congress](#) (traduction de l'équipe de rédaction).

⁵ *User Experience Designer*, c'est-à-dire concepteur d'expérience utilisateur.

⁶ GOLDSTEIN, Brett. [Spotify, we need to talk about your data](#), août 2014.

⁷ CNIL. [Le Corps, nouvel objet connecté ?](#), Cahier IP n°2, mai 2014.

⁸ LAMERE, Paul. « [How we listen to music](#) », festival SXSW, 2015

⁹ CNIL. [La Place des photos dans la vie numérique](#), 2012.

QUELLE EST LA RICHESSE EN DONNÉES DES UNITÉS DE CONTENUS*?

*déf.: Une unité de contenu est un média cohérent formant l'unité de base indivisible de ce type de contenus :



une chanson



un livre



un épisode de série/un film



un jeu vidéo

#volume

Combien d'unités de contenus différentes sont en moyenne consommées par un individu sur une période donnée



#densité

Quelle quantité de données est analysable lors de la consommation d'une unité ?



#agrégation

Quelle est la valeur de prédiction de l'ensemble des unités consommées ?



#valeur marginale

Quelle connaissance nouvelle apporte chaque unité supplémentaire consommée ?



#engagement

Quel est le degré d'investissement personnel pour l'utilisateur sur chaque unité ?



#réputation

L'utilisateur est-il sensible à l'image véhiculée en ligne par ses consultations et les partage-t-il ?



#contexte

À quel point la consommation d'une unité varie-t-elle en fonction du contexte ?



il se gardera de partager son écoute de musique plus commerciale. Enfin, l'écoute de musique peut-être extrêmement contextuelle. Ainsi, les styles de musique varient en fonction de son contexte social (seul, famille, travail, amis), du type d'activité du moment¹⁰ (sport, détente, concentration), et bien évidemment le style n'est pas le même lors d'une activité festive.

Dans le domaine du livre et des écritures numériques, la diversité des « points de données » est plus réduite. Si l'on se concentre sur le livre numérique, un lecteur donné lira un nombre relativement réduit de livres par an. La dernière enquête sur les pratiques culturelles des Français (2008) révèle que les Français ayant lu au moins un livre dans l'année disent lire en moyenne

16 livres par an. Il n'y a que 16% des lecteurs qui affirment lire plus de 20 livres par an.

En revanche la lecture de livres numériques permet la collecte d'énormément de données pendant la consultation du contenu : on peut mesurer le rythme de passage d'une page à l'autre, le temps passé sur une page, les durées de chaque séquence de lecture, le nombre d'arrêts et de reprises (permettant par exemple de mesurer en combien de fois le livre est lu), les signets et phrases surlignées et/ou sauvegardées, etc. (voir chapitre II-2).

Enfin, le taux « d'engagement » est fort, car la lecture est un effort, un investissement, souvent conséquent en temps. Si l'on peut tester un nouveau style, flâner d'un contenu musical à

¹⁰ Les *playlists* de motivation à l'activité sportive sont ainsi extrêmement répandues, de même que celles de musique au travail.

un autre, il est plus rare pour un lecteur de « lire au hasard ». Les données sont plus souvent cohérentes avec les goûts, et il sera donc difficile d'obfusquer volontairement ou involontairement ses goûts, ses choix, ses préférences.

Dans le secteur de la vidéo et de l'image animée, chaque « unité de contenus » a une taille intermédiaire : regarder un film ou un épisode de série est en moyenne moins long que lire un livre, mais évidemment plus long qu'écouter une chanson.

Selon *Netflix*, ses abonnés (en Amérique du Nord) regardent 1 milliard d'heures de vidéos par mois. *Netflix* indique par ailleurs avoir autour de 40 millions d'abonnés dans cette région du monde, cela veut dire que chaque abonné regarde en moyenne 25 heures de vidéos par mois, soit entre 10 et 20 contenus différents par mois. Sur un an, un abonné moyen regarderait donc de quelques dizaines à une ou deux centaines de contenus différents¹¹. Ces contenus sont très divers et très bien documentés (styles, acteurs, réalisateur, notes, critiques, recommandations d'autres utilisateurs, etc). En revanche, ils sont souvent consommés par un même foyer, ce qui implique de connaître celui-ci et de le décomposer autant que possible, même si cela ne s'avère pas toujours très difficile : un foyer d'adultes avec des enfants en bas âge aura un type de consommation relativement identifiable par rapport aux styles et aux horaires. Par ailleurs, ces contenus sont très sociaux : ils alimentent nombre d'échanges entre individus - la fameuse « discussion de la machine à café » sur le lieu de travail, qui est à rapprocher des pratiques croissantes de « binge watching », c'est-à-dire la pratique qui consiste à regarder à la suite en une seule fois de nombreux épisodes d'une même série.

Evidemment, ces chiffres ne sont pas tout à fait aussi justifiables si l'on élargit la réflexion à l'ensemble des contenus de télévision connectée (télévision de rattrapage ou contenus spéciaux) ou à des services plus larges de vidéo en *streaming* (*YouTube* ou *DailyMotion*) ; mais sur la partie de vidéo à la demande, ils donnent une image assez complète des pratiques de consommation.

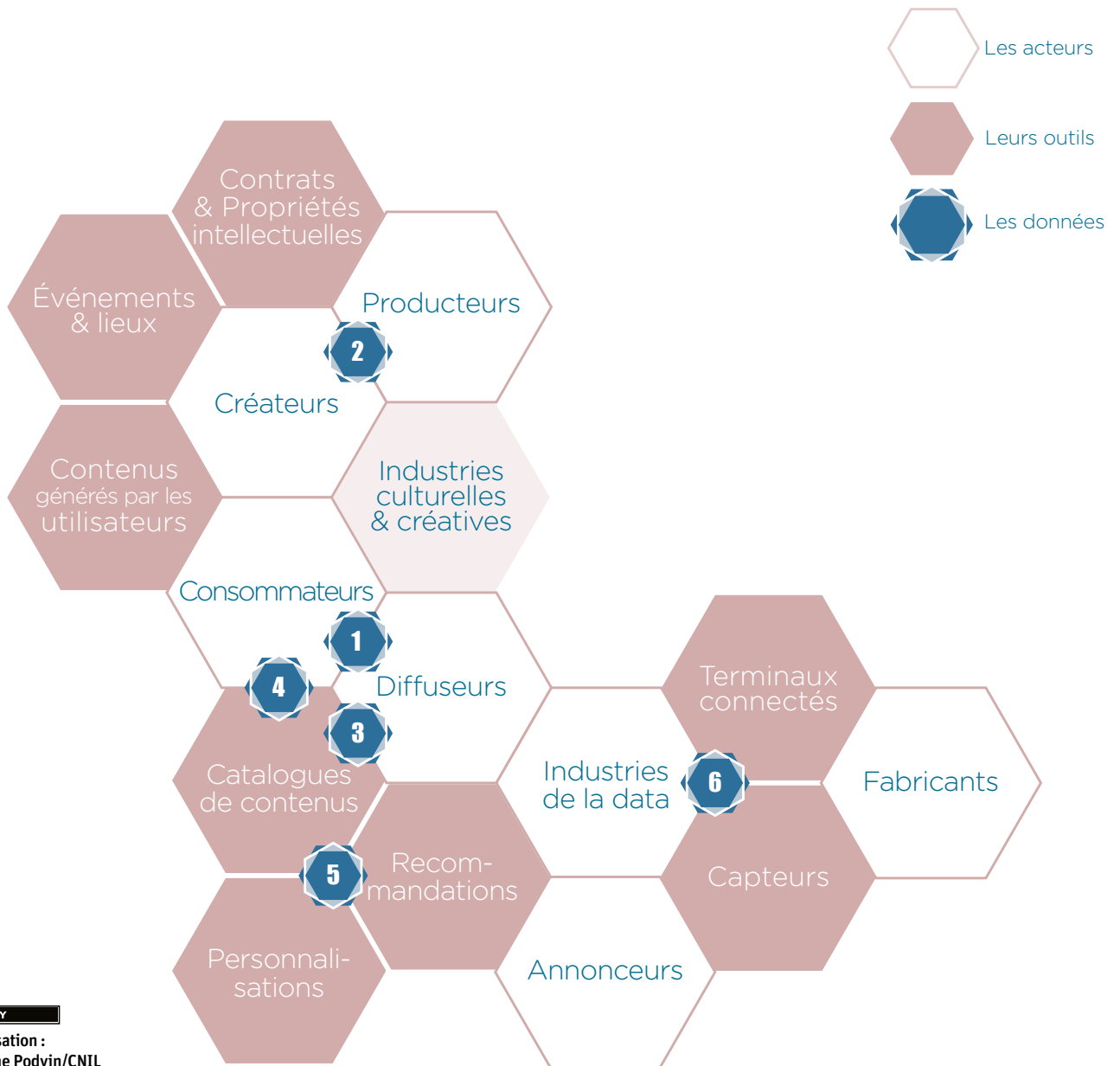
Enfin, le jeu vidéo est lui aussi très particulier. En ce qui concerne le nombre d'unités de contenus consommées, le secteur du jeu « traditionnel » se rapproche finalement du livre numérique : un jeu vidéo pour console ou ordinateur

coûtera généralement plusieurs dizaines d'euros à l'unité et sa durée de vie est couramment de plusieurs dizaines d'heures. Même si l'on prend en compte des formes plus « légères » de pratiques vidéo-ludiques, comme par exemple les jeux sur smartphones ou sur les services de réseaux sociaux, c'est bien évidemment cette caractéristique qui qualifie le mieux la particularité des données du secteur : les joueurs passent un temps important au cœur du contenu et ils y sont actifs en permanence. Une quantité importante de données est donc issue de l'usage lui-même. En réalité, les jeux permettent d'analyser les actions des utilisateurs et donc d'évaluer, presque en temps réel, leurs préférences, leurs choix ou leur comportement. ■

¹¹ Source des données : <http://blog.netflix.com/2014/01/new-isp-performance-data-for-december.html>

CARTOGRAPHIE DES INDUSTRIES CULTURELLES & CRÉATIVES

Quelles sont les données et où créent-elles de la valeur dans l'écosystème des industries culturelles et créatives ?



Réalisation :
Mylène Podvin/CNIL

- 1** données de profil
- 4** données de consommation & popularité
- 2** données descriptives de contenus
- 5** données de goûts et d'usages
- 3** données d'enrichissement
- 6** données de contexte

DU LIVRE NUMÉRISÉ AUX ÉCRITURES NUMÉRIQUES ET AUX NOUVELLES EXPÉRIENCES DE LECTURE

“ Que ce soit dans une bibliothèque bondée ou dans une sombre chambre à coucher, peu d'expériences sont aussi intimes que de lire un livre. Les livres concernent l'œil et la page, un cerveau humain en conversation avec un autre. Pourtant, le commerce des livres n'est jamais à propos de ces vertus élevées. Et comme toute industrie, l'édition doit toujours se concentrer sur le grand public, la masse : que veulent la plupart des gens ? Qu'est ce qui sera acheté par le plus grand nombre ? À quoi les gens réagiront-ils ? Entre ces deux facettes, il y a une étrange relation. Les entreprises collectent et analysent les données, mais les lecteurs ont rarement l'occasion de les voir. » ”

Robinson Meyer

« [The Most Popular Passages in Books, According to Kindle Data](#) », *The Atlantic*, novembre 2014

- ▶ « Dis-moi comment tu lis et j'adapterai des contenus à ton style de lecture. »
- ▶ La première transformation numérique, celle des liseuses, a certes été précoce et marquante, mais par certains côtés elle est aussi une impasse : il s'agissait simplement de « scanner » le livre papier, sans vraiment profiter du numérique pour changer l'œuvre.
- ▶ Un nouveau type de lecture est en train d'apparaître avec des contenus et des formats plus différenciés, ainsi que des systèmes de recommandation.
- ▶ L'enjeu industriel est double : garder le lecteur à l'intérieur de l'ouvrage et ouvrir le livre vers la communauté du lecteur.

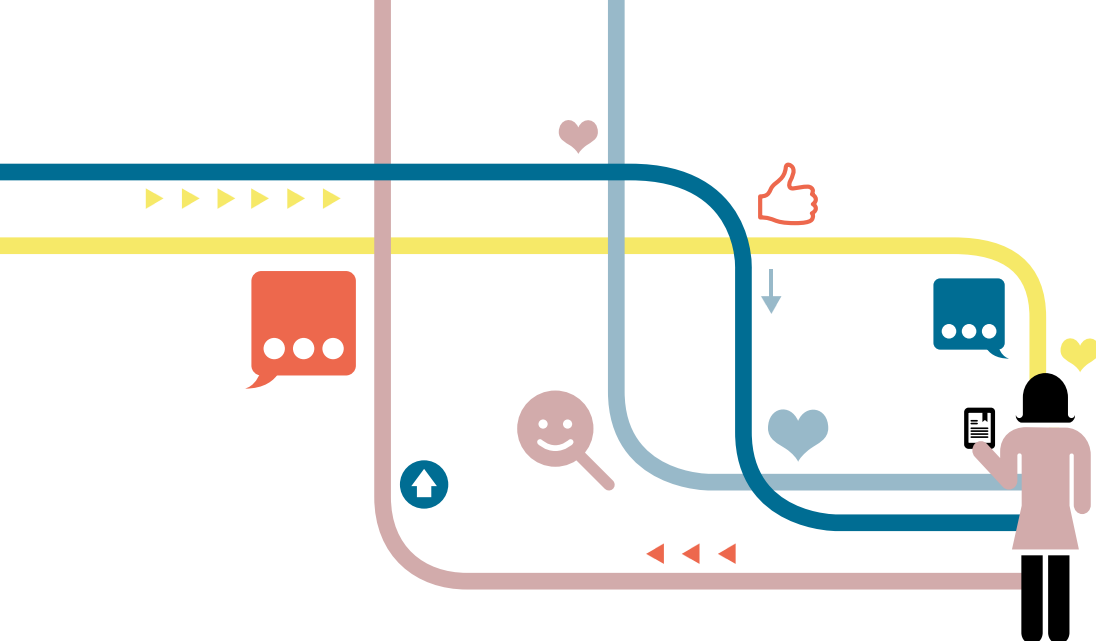
¹ Electronic Frontier Foundation. [E-reader Privacy Chart](#), 2012

² Entretien, et sur son blog [La Feuille](#) ainsi que régulièrement sur [InternetActu](#) sur ce thème.

OÙ EN SOMMES-NOUS ET OÙ VA-T-ON ?

La transformation numérique du monde de l'édition a été relativement précoce et s'est accomplie sur deux axes : les modes de distribution en ligne - avec des offres larges soutenues par des moteurs de recommandation efficaces et une logistique globale sans faille - et les nouveaux supports de lecture (les liseuses, pour la première génération). Dans les deux cas, Amazon a été le grand précurseur de cette transformation, son *Kindle* devenant le symbole des liseuses même si quelques concurrents existent à l'instar des *Kobo* distribués en France par la *FNAC*.

Jusqu'à maintenant, les seuls chiffres étaient ceux des ventes ; désormais l'usage est suivi, mesuré, analysé, sans que le lecteur en ait forcément conscience. Les fabricants de liseuses, *Amazon* au premier chef, enregistrent toutes les traces d'activité (les *logs*). Les liseuses sont des objets fermés et la connaissance des données en circulation dépend de ce que les fabricants veulent bien livrer au public. En 2012, l'[Electronic Frontier Foundation analysait la collecte de ces données](#) à partir de l'analyse des politiques de confidentialité¹ et notait que les plateformes prennent connaissance de la liste des ouvrages achetés, des ouvrages ouverts, du temps passé sur chaque page, des passages surlignés ou des annotations. Comme le rappelle Hubert Guillaud², il est possible de déduire des informations sensibles à partir de ces données (orientation politique ou état de santé, par exemple) ; dès lors, les utilisateurs devraient avoir la possibilité de les partager ou non. Connaître la nature des échanges entre ces entreprises et les éditeurs est également impossible : revendent-ils les données ? Cela fait-il partie des contrats ? Beaucoup de mythes courent en ce domaine. Les offreurs de technologies ont en effet du mal à convaincre les éditeurs qui conservent une approche plus traditionnelle de leur métier. Le responsable des relations avec les éditeurs de *Kobo* témoignait ainsi : « nous aidons les éditeurs en leur disant "ce livre est perdu au milieu de votre liste mais il est lu plus rapidement que vos meilleures ventes. Peut-être devriez-vous investir en relations publiques et dans les réseaux sociaux derrière lui". » Dans le même article, Micah Bowers, le CEO de *Bluefire*, affirme que les éditeurs « ne pensent traditionnellement pas au contenu après qu'ils »



TENDANCES CLÉS

- Évolution du livre numérisé vers les expériences d'écriture numérique
- Fin de la lecture linéaire : possibilité de proposer des modes de lecture à différentes échelles (extraits choisis selon les goûts de l'utilisateur, trames narratives construites pour le lecteur...).
- Plus grande hybridation des supports de lecture.
- Pilotage de la production de certains contenus par les données en provenance des utilisateurs
- Captation des émotions dans la lecture pour favoriser l'engagement : *neuromarketing, nudge...*
- Création de nouvelles pistes de socialisation autour des contenus écrits, par des lieux ou des services.

DONNÉES CLÉS



AUJOURD'HUI

- Les données techniques sur la lecture (passages préférés, vitesse de lecture, profondeur, etc.) et l'analyse sémantique des contenus.



DEMAIN ?

- Les émotions et les métadonnées pour l'écriture algorithmique, combinées au contexte capté.



l'aient publié. C'est une nouvelle manière de penser pour eux de commencer à analyser comment leur contenu est réellement utilisé.»³

Amazon fait sienne cette logique et rémunère certains auteurs à la page lue depuis le 1^{er} juillet 2015⁴. L'intégration des données d'usage de lecture dans le modèle économique, qui n'est pas sans rappeler le *pay as you drive* du monde de l'assurance automobile⁵, va d'abord concerner les auteurs autoédités sur Amazon avec le programme Kindle Direct Publishing⁶. L'annonce du géant du commerce a suscité de l'inquiétude quant aux conséquences d'un mode de rémunération à la page tournée sur les choix d'écriture. Comment les auteurs et éditeurs vont-ils réagir à ce que le chercheur Olivier Ertzscheid désigne comme « l'angoisse de la page non lue »⁷ ? Certains craignent une perte de créativité et une aversion croissante aux risques des éditeurs, qui ne voudraient produire que des contenus dont les données prédisent, garantissent pourrait-on dire, le succès. Le fait est que les éditeurs utilisent déjà beaucoup de données pour « deviner » quel auteur va rencontrer le succès⁸.

Les données d'usage sont en effet appelées à jouer un rôle majeur dans l'accompagnement de la création. L'innovation peut d'abord passer par une meilleure connaissance des réactions des lecteurs, très facile à transformer en *insights*⁹ pour le marketing des éditeurs : savoir comment les lecteurs lisent un livre, savoir quel passage ou quelle phrase ils ont surlignés, connaître leur rythme de lecture... sont des moyens très forts

pour encourager l'achat d'une suite ou faire la promotion du livre auprès d'autres lecteurs potentiels. L'usage de ces données peut aller bien plus loin. Il peut influencer directement sur le type d'écriture et sur la création. C'est déjà le cas pour certains *best-sellers*, qui sont savamment découpés en épisodes plutôt que vendus en un seul tome une fois pour toutes. Les modèles économiques émergents basés sur les usages effectifs de lecture peuvent conduire à une multiplication des *cliffhangers*, c'est-à-dire au maintien d'un suspense permanent pour garder le lecteur captif - au détriment de la nuance et de la complexité. L'impact des données sur la production de contenus devrait même aller plus loin : personnalisation du contenu à partir des habitudes du lecteur, en fonction de profils de lecteurs prédéterminés, sont des hypothèses recevables. Sont-elles nécessairement négatives ? Selon Hubert Guillaud, si auteurs et lecteurs sont présents, actifs et respectés, une soumission pure et simple de l'écriture à l'industrie du profilage est évitable. Les données sont en effet un formidable miroir pour les auteurs (l'usage émergent de Wattpad, voir encadré « usages émergents », point 2, p.31) et si les lecteurs sont aussi destinataires des données, ce qui peut en être fait par d'autres acteurs sera peut-être plus compréhensible et moins inquiétant.

Pour l'instant les auteurs ne sont pas les acteurs principaux de cette transformation. Aurélien Bellanger, romancier¹⁰, souligne ainsi, lors d'un entretien avec les auteurs de ce cahier, qu'en tant qu'auteur, il n'a pas réellement accès à de telles données. Pour lui « le monde de l'édition, en particulier autour du "roman littéraire", vit toujours avec un grand fantasme artisanal » et la mesure du taux d'engagement passe par exemple par les salons du livre et les rencontres en librairies. Conscient de l'opacité pour les auteurs des outils utilisés par les librairies pour le suivi des ventes, il va plus loin : « Le droit d'auteur devrait peut-être inclure une sorte de droit d'accès aux données... »

Pour le lecteur, la confidentialité de l'acte de lecture est questionnée. Ses habitudes de lecture, la vitesse à laquelle il termine (ou abandonne) un ouvrage, ses passages favoris, etc., toutes ces informations qui forgent le caractère intime de la lecture se voient dotées de nouvelles fonctions possibles, en dehors du lien tissé entre le lecteur et son livre. De quelles marges de manœuvre dispose-t-il pour adapter ses pratiques de lecture à ses exigences d'intimité ou une volonté plus ponctuelle de partage ? ■

³ HARVEY, Ellen. « [What Book Publishers Can Glean From Reading Behavior Data](#) », *Book Business*, février 2015

⁴ Voir le détail sur <https://kdp.amazon.com/help?topicId=A2J855CR326YIG> (visité le 9 septembre 2015)

⁵ Le conducteur paie une prime d'assurance automobile en fonction de son utilisation du véhicule.

⁶ Kindle Direct Publishing permet aux auteurs de publier eux-mêmes leur ouvrage en version électronique : <http://kdp.amazon.com>

⁷ ERTZSCHEID, Olivier. « [Lecture à la découpe : Amazon paiera les auteurs à la page lue](#) », *Affordance*, 21 juin 2015

⁸ Dossier *BookBusiness*, « [Could An Overdependence on Data Hurt The Book Industry?](#) », février 2015.

⁹ En marketing, le terme *insight* renvoie à l'expression d'un besoin par le consommateur.

¹⁰ Auteur de : *L'Aménagement du territoire* (Gallimard, 2014) et de *La Théorie de l'information* (Gallimard, 2012).

¹¹ WALLOP, Harry. [The Kindle is dead, the book is back. Or is it?](#), *The Telegraph*, 9 janvier 2015.

¹² CROUZET, Thierry. [Wattpad : la machine à best-sellers](#), février 2015.

¹³ Voir la vidéo de présentation du projet : <http://www.my-os.net/archives/?p=1156>.

¹⁴ [Interview de son fondateur sur BFM business le 23 mai 2015](#), (mai 2015)



USAGES ÉMERGENTS

1 NOUVELLES PRATIQUES D'ÉCRITURE NUMÉRIQUE

► Le numérique offre aux écritures de grandes opportunités de renouvellement des expériences proposées. À cet égard, dans le contexte des contenus créés pour les jeunes enfants, sur tablettes notamment, [Nosy Crow](#) se focalise sur l'édition jeunesse et considère que le livre numérique doit aller au-delà du « scan de la page sur un écran », pour intégrer de l'interaction. Pour Kate Wilson, managing director de cette entreprise : « la lecture ne doit pas être la chose la plus ennuyeuse qu'ils font sur un iPad ». Le futur de l'écriture doit répondre à la question « comment créer des objets qui continuent de stimuler les enfants, qui leur fassent penser que lire est une activité à laquelle ils veulent consacrer du temps ? »¹¹

2 « CONNAÎTRE SES LECTEURS, C'EST ÉCRIRE POUR EUX »

► Thierry Crouzet raconte dans un billet de blog¹² son expérience de [Wattpad](#). Dans un contexte où émergent de nouveaux intermédiaires - des acteurs qui sortent du modèle traditionnel centré sur l'éditeur et la verticalisation - l'application [Wattpad](#), plateforme d'auto-publication jouant le rôle de point de rencontre entre auteurs et lecteurs, permet aux auteurs de bénéficier directement de données sur leurs lecteurs. La création *data-driven* n'est plus l'apanage des éditeurs travaillant avec [Amazon](#) ou [Kobo](#) : les auteurs peuvent accéder immédiatement à des données leur permettant (ou pas) d'adapter leur création à la réaction des lecteurs. Les métriques jouent un rôle majeur au sein de l'application : les interactions entre auteurs et lecteurs, sous la forme de commentaires ou de suggestions, permettent d'initier un dialogue, et les données fournies par [Wattpad](#) aux auteurs (chiffres d'audience, profil et modalités de lecture du lectorat) permettent de mesurer l'engagement ou d'adapter sa création à son public. Thierry Crouzet prend l'exemple de l'âge de son lectorat adolescent, qui l'incite à créer des personnages jeunes.

3 INFLUENCE DU CONTEXTE SUR LES MODES DE LECTURE

► En mettant l'accent sur le nombre de minutes nécessaires à la lecture de chaque article, l'application mobile du groupe Le Monde « La Matinale » permet à ses utilisateurs de personnaliser leur édition quotidienne. La personnalisation du contenu en fonction du « temps de cerveau disponible » peut aussi être automatisée, à travers le croisement des données de géolocalisation et de déplacement. Déjà en 2013, l'équipe de [TimeReader](#) remportait le *hackathon* organisé avec la RATP #Opendatalab. La web app proposait aux usagers des transports parisiens des contenus texte et audio adaptés à leur temps de trajet, en se basant notamment

sur les positions géographiques des stations du réseau RATP. La connaissance très fine des mobilités peut également alimenter des usages plus ludiques : lors d'un atelier organisé à l'HEAD (Haute école d'art et de design) de Genève par le designer Étienne Mineur, un groupe d'étudiants a proposé une adaptation d'une bande dessinée de Riad Sattouf tenant compte du contexte de lecture. Les anecdotes de *La vie secrète des jeunes* se déroulant dans le métro sont accessibles au lecteur équipé d'un smartphone uniquement sur la ligne de métro où elles se déroulent¹³ !

4 SOCIALISER LE LIVRE NUMÉRIQUE ET LA LECTURE

► C'est l'ambition du service [Glose](#)¹⁴ qui se présente comme une « plateforme numérique de lecture sociale » et qui outre la lecture d'ouvrages permet leur annotation, l'archivage de citations et le partage de ces informations avec la communauté des lecteurs de l'application. L'objectif de ce type de service est de « socialiser » le livre et plus globalement les ouvrages écrits. La plateforme envisage également d'offrir la possibilité aux auteurs d'annoter leurs propres livres pour se rapprocher de leurs lecteurs et lever le voile sur le processus de rédaction (à la manière des *making of* pour les films). Le modèle économique est original puisqu'il consiste à offrir gratuitement les 10 premiers pourcents de chaque nouvel ouvrage référencé par la plateforme. Cette tactique tarifaire permet de traiter en partie de la problématique « bien d'expérience » (voir partie III) à laquelle est confronté un livre, et pose en même temps la question de son impact potentiel sur la phase de création si elle venait à se généraliser (i.e. : mettre en place des mécaniques d'écriture qui maximisent la probabilité d'achat après la lecture des 10 premiers pourcents).

5 MÉLANGER LA FICTION ET LA VIE DU LECTEUR

► L'agence de design [IDEO](#) s'est penchée sur le [futur du livre](#) par la mise en prototypes de grandes tendances. Leur concept *Alice* décrit une expérience de lecture interactive qui engage le lecteur dans la création de l'histoire. Pour cela, il s'agit de mélanger la fiction et la vie du lecteur, à l'aide de la géolocalisation, en incitant le lecteur à communiquer directement avec les personnages de l'histoire via son téléphone... L'objectif étant de favoriser au maximum l'immersion du lecteur.

DE LA PLAYLIST À LA CONTEXTUALISATION ET À LA MUSICALISATION DU QUOTIDIEN

“ Ma maison est faite de bois, de verre et de pierre. Elle est aussi faite de logiciels. Si vous venez me rendre visite, vous serez probablement surpris en entrant. Quelqu'un vous donnera un badge électronique à porter. Ce badge dit à la maison qui vous êtes et où vous êtes. La maison utilise cette information pour vous donner ce dont vous avez besoin. Quand il fait sombre dehors, le badge allume la lumière la plus proche de vous, et l'éteint quand vous vous éloignez. La musique bouge avec vous également. Si la maison connaît votre musique préférée, elle la joue. La musique semble être partout mais en fait, d'autres gens dans la maison entendent d'autres musiques. ”

Bill Gates *La route du futur, 1999.*
(traduction par l'équipe de rédaction)

► “ Dis-moi où et quand tu écoutes et je te proposerai des contenus contextualisés ”

► La musique écoutée a une fonction identitaire très forte, qui va se matérialiser par la revendication de certains contenus (partage, publication, recherche de visibilité sociale) et à l'opposé par des consommations plus discrètes ou cachées.

► La consommation musicale passe par beaucoup de *zappings*, de pratiques « boulimiques », de changements de style, de mécanismes de saturation... Elle est souvent liée au contexte (je suis seul, je suis au travail, je suis à une fête, je fais du sport, je suis en voiture, etc). Le « parcours utilisateur » évolue tout au long d'une journée avec des modes d'accès à la musique très variés via différents dispositifs, medias... La musique peut ainsi être très présente lors de notre visite dans un magasin, chez le coiffeur... ou dans un ascenseur.

► Il y a un intérêt évident pour les diffuseurs à comprendre les multiples séquences de la consommation de musique au quotidien. La consommation musicale, moins accaparante, permet d'accompagner l'utilisateur dans des moments clés du quotidien.

¹ Racheté par Apple, Beats Music est devenu Apple Music. Son centre de gravité a changé, même si l'annonce d'une radio « universelle » reste caractéristique de cette proposition éditoriale forte. Pour en savoir plus sur cette radio, voir partie 3 et <http://alireailleurs.tumblr.com/post/121338188203/apple-music-une-nouvelle-forme-de-gentrification>. Apple Music semble en effet s'attaquer aux acteurs du *streaming* comme Spotify mais également aux services plus communautaires comme SoundCloud au travers de sa fonction Connect.

² http://fr.wikipedia.org/wiki/Curation_de_contenu

³ Exemples ci-dessous issus des pages d'accueil des services.

⁴ « Ce logiciel utilise le microphone du téléphone pour capturer un échantillon de musique jouée. Une empreinte acoustique est créée à partir de cet échantillon, elle est comparée à la base de données centrale de la société. En cas de correspondance, les informations de la base de données comme le nom de l'artiste, le titre de la chanson, et l'album sont retournées à l'utilisateur. ». Source: Wikipédia, consulté le 15 juin 2015.

⁵ PEGORARO, Rob. « Pandora's "Music Genome Project" explores the cold hard facts of how we interact with music », *BoingBoing*, mai 2014.

OÙ EN SOMMES-NOUS ET OÙ VA-T-ON ?

« Plus d'accès, moins de propriété » caractérise la pénétration croissante des services de *streaming* pour l'accès au contenu audio. Spotify est aujourd'hui le leader mondial et reste la plateforme emblématique de *streaming* musical, principalement concurrencé par Deezer qui dispose d'une forte pénétration en France, à l'instar de Pandora aux États-Unis. D'autres acteurs tentent de se différencier : Soundcloud est beaucoup plus orienté vers les contenus mis en ligne et partagés par les utilisateurs (*user-generated contents*), ou plus récemment Beats Music¹ et Tidal qui promettent de faire la part belle à la « curation » et aux contenus exclusifs².

Ces plateformes formulent toutes peu ou prou la même promesse : proposer la bonne musique au bon moment³ :

- Spotify « Every time you fire up with Spotify, we'll suggest some great music for right now »
- BeatsMusic « What if you could always have the perfect music for each moment? »
- Pandora « We have a single mission: to play only music you love »
- Deezer « Your flow is your music: 100% customized. It know your library and selects the songs you want to hear »

Accompagner la découverte à la manière de ce que propose Shazam, qui permet d'identifier une musique écoutée par l'intermédiaire du micro du smartphone⁴ et recommander des contenus pertinents par rapport à la personne, sont deux dimensions incontournables des services de *streaming*. Pour atteindre ces buts, il faut d'abord décoder ce qui caractérise une chanson. Pandora a ainsi lancé depuis plusieurs années un projet de taxonomie musicale dénommé « Music Genome », dont l'ambition est de décrypter ce qui forme le cœur d'une chanson par l'analyse de chaque morceau au travers de [450 caractéristiques](#)⁵. Début 2014, Spotify a fait l'acquisition



Capture d'écran du site The Echo Nest, 09/2015

de *The Echo Nest*⁶, une société spécialisée dans l'analyse statistique et la suggestion musicale – solution jusqu'alors utilisée également par certains de ses concurrents comme *Deezer*. *The Echo Nest* revendique analyser des milliards de données issues des chansons pour les convertir en « compréhension musicale » : [pour reprendre certains de leurs concepts](#), *The Echo Nest* mesure ainsi l'« acousticité » ou la « danceabilité » des chansons⁷.

Comme [l'a annoncé Spotify](#)⁸ en mai 2015, l'objectif est de personnaliser davantage la recommandation en proposant des *playlists* en fonction des habitudes et des humeurs des utilisateurs ; ceci permet de moins s'appuyer sur des analyses de genres et de sous-genres qui restent beaucoup trop subjectives et peu précises. Paul Lamere, directeur du développement de *The Echo Nest*, a indiqué [lors de la conférence SXSW en mars 2015](#) que sur les 100 intitulés de *playlists* les plus utilisés sur le service, 17 sont liés à un genre (rock, rap...) et 41 à un contexte (fête, sport, détente, voyage, météo, amour, moment de la journée, concentration au travail...)⁹.

Un tel mécanisme favorise la découverte, très importante dans la consommation musicale. Selon Shiva Rajaraman, vice-président chargé du produit chez *Spotify*, il s'agit de transformer l'utilisation du service en un rituel : « Vous commencerez à l'utiliser avec vos habitudes et nous l'alimenterons en contenu pour chaque moment de votre journée. »¹⁰

Cette approche paraît d'autant plus pertinente que la consommation de musique obéit à des règles très particulières (voir typologie du chapitre II-1). Pour [Romain Bigay du Centre d'information et de ressources pour les musiques actuelles](#) (IRMA) « la portabilité de la musique entraîne une plus grande "musicalisation" du quotidien, renforcée par le caractère pervasive de la musique : celle-ci ne requiert pas une attention unique, l'écoute peut se superposer à toute autre activité »¹¹. Le numérique a donc amplifié des pratiques déjà existantes, en particulier la sélection des goûts musicaux en fonction du moment, du lieu, du contexte. Finalement, peu importeront les styles ou l'époque, ce qui compte aujourd'hui est l'expérience d'écoute : la musique sert de bande-son du quotidien et le contexte en est la clé.

Les systèmes de recommandation existants sont-ils en mesure de cibler aussi précisément leurs prescriptions pour suivre de tels besoins de contextualisation de l'expérience ? La compréhension des comportements et des activités

de l'utilisateur prend alors le pas sur le recours au filtrage collaboratif ou à l'analyse du contenu musical en lui-même (voir partie III). Comme le souligne Julie Knibbe, chef de produit chez *Deezer*¹², il faut disposer de tout un ensemble de paramètres et de variables hétérogènes pour produire une recommandation efficace, tels que le profil sociologique, la curiosité, les habitudes et goûts de l'utilisateur. Ces données servent à établir un profil d'écoute des utilisateurs allant de « indifférent » à « érudit » pour adapter le niveau d'originalité des contenus qui seront recommandés : des plus populaires pour les premiers aux plus pointus pour les utilisateurs avancés.

Avec toutes ces données, ces plateformes ne finissent-elles pas par connaître les utilisateurs mieux qu'ils ne se connaissent eux-mêmes ? Walt Hickey, du site spécialisé en *data science* *FiveThirtyEight*, en a fait l'expérience et [raconte ce qu'il a ressenti après avoir subi une telle analyse](#) : « avant de faire cette expérience, je croyais qu'il y avait deux grandes catégories de musique : celle que vous écoutez et celle que vous dites aux autres écouter. (...) Cette analyse m'a appris qu'il y a trois catégories de musique : celle que vous dites aux autres écouter, celle que vous croyez écouter, celle que vous écoutez réellement »¹³.

Par ailleurs, la capacité de ciblage ne sert pas qu'à enrichir les expériences des utilisateurs. Comme dans les autres secteurs, le marché publicitaire existe. Ainsi, [Spotify a annoncé en avril 2015](#) étendre son offre à destination des marques grâce à l'outil « *playlist targeting* » : les marques pourront cibler des utilisateurs qui sont en train d'écouter une *playlist* « sport » ou « détente ».

ET DEMAIN ?

Dans la durée, ces services construisent une connaissance fine des utilisateurs qui n'est pas nécessairement partagée ou mise à la disposition de ces derniers. La masse de données issue de nombreuses années d'écoute reste pour l'instant enfermée dans les serveurs du service, hors de quelques tentatives pour restituer une information comme « [Une année en musique](#) » chez *Spotify*. Dans le quotidien *The Guardian*, [un journaliste a pu analyser 6 années de ses consommations musicales chez Spotify](#). Confronté à ces données, il a pu constater deux choses : premièrement, que ce qu'il consomme autrement que sur *Spotify* n'apparaît évidemment pas et qu'il est donc seul à pouvoir compléter ces données, pour le moment en tout cas ; deuxièmement

⁶ Signe que *Spotify* mise peut-être plus que ses concurrents sur les données, en juillet 2015, la société a également racheté *Seed Scientific* un spécialiste de l'*analytics* (voir par exemple : <http://techcrunch.com/2015/06/24/pulling-the-data-rug-out-from-under-apple/>)

⁷ DREDGE, Stuart, « [Pop music is louder, less acoustic and more energetic than in the 1950s](#) » *The Guardian*, 25 novembre 2013.

⁸ La plateforme a également annoncé des velléités de diversification vers la vidéo en ligne avec des contenus exclusifs issus de partenariats. Voir par exemple : DUPONT-CALBO, Julien, « [Avec des podcasts et de la vidéo, Spotify se rêve en juke-box absolu](#) », *Les Echos*, 20 mai 2015

⁹ LAMERE, Paul « [How we listen to music](#) », *SXSW 2015*, mars 2015.

¹⁰ MANJOO, Farhad, « [Spotify Wants Listeners to Break Down Music Barriers](#) », *The New-York Times*, 3 juin 2015.

¹¹ BIGAY, Romain, « [Le streaming modifie-t-il les goûts musicaux ?](#) », *IRMA*, septembre 2014.

¹² KNIBBE, Julie et MOUSSALAM, Manuel. [Big Data as a Streaming Service](#), présentation *SlideShare*, octobre 2014

¹³ HICKEY, Walt, « [Spotify Knows Me Better Than I Know Myself](#) », *FiveThirtyEight*, septembre 2014.

que la question de la portabilité des données n'est pas seulement technique puisqu'il est difficile de passer d'un service à l'autre sans devoir recommencer tout l'apprentissage des goûts. Comme le conclue Stuart Dredge dans cet article : « comme de plus en plus de gens se rendent compte des avantages liés à une meilleure compréhension de nos goûts par les services de musique numérique, je pense que nous allons voir plus de discussions sur la façon dont (ou même si) nous pouvons exporter les données pour les utiliser ailleurs. »¹⁴. L'absence de tels outils est tellement troublante que [le designer](#)

[Brett Goldstein a imaginé](#) à quoi pourrait ressembler l'extension vers les utilisateurs de l'outil « Spotify Analytics », pour le moment réservé aux artistes¹⁵. En regardant ainsi de l'autre côté du miroir, chaque utilisateur pourrait en savoir plus sur ses propres goûts et sa manière d'expérimenter la musique.

Cet exemple illustre la valeur du droit à la portabilité : il doit à la fois favoriser la concurrence sur les marchés du numérique, mais aussi ouvrir enfin les portes d'une nouvelle valeur d'usages des données personnelles, [pour les individus eux-mêmes](#)¹⁶. ■



USAGES ÉMERGENTS

Si la personnalisation du service devrait rester la tendance majeure dans les prochaines années, elle devrait en revanche évoluer à la fois dans sa forme, en nécessitant de moins en moins l'intervention de l'utilisateur, et sur le fond en agissant sur le contenu musical lui-même.

1 SERVICE SANS INTERFACE

► [Zero Button Music Player](#) (ou lecteur sans bouton) est un concept de service présenté sur le blog du développeur de la plateforme d'*Echo Nest*. Selon lui le bouton « play » a vocation à disparaître du fait qu'il constitue la dernière barrière entre l'utilisateur et la musique. Les *playlists* pourraient être générées automatiquement en fonction du contexte perçu (ex. : musique dynamique pour accompagner une séance de sport, calme pour se concentrer sur une lecture), sans action spécifique de l'utilisateur.

2 MUSIQUE CONTEXTUELLE

► [Prizm](#) est un projet français présent sur la plateforme de *financement participatif Kickstarter*. Il propose un boîtier dont l'ambition est de comprendre les habitudes musicales des utilisateurs pour contextualiser davantage la recommandation. En sondant les différents réseaux et appareils qui l'entourent (comme des smartphones, bracelets ou autres objets connectés) le boîtier est capable d'en déduire par exemple si la personne est seule ou non, et de proposer un contenu musical adapté en mixant les goûts musicaux du propriétaire et son environnement.

3 INTERACTIONS « LIVE » ENTRE LE PUBLIC ET UN ARTISTE

► C'est le type de scénario que peut proposer le bracelet connecté [Lucie Hive destiné à apporter une expérience immersive et interactive lors](#) de festivals ou de concerts. Au-delà de la finalité de contrôle d'accès, le bracelet embarque une série de capteurs (accéléromètre, RFID...) et est équipé de LED¹⁸ qui peuvent être contrôlées de

manière centralisée afin que les festivaliers puissent faire partie intégrante du spectacle (i.e. : possibilité pour le groupe ou DJ de synchroniser la couleur des bracelets avec la musique). Pourrait-on aller encore plus loin dans la personnalisation des interactions ? C'est déjà le cas en amont de l'événement « Rock En Seine » lorsque l'application pour smartphones du Festival propose une connexion au compte Spotify de l'utilisateur afin de lui recommander ce qui devrait lui plaire dans la programmation du festival.

4 MUSIQUE PLUS COLLABORATIVE SUR LA BASE D'ŒUVRES INACHEVÉES

► Pour Tod Machover, inventeur et chercheur au Medialab du MIT, la tendance marquante du monde de demain sera l'avènement d'une musique plus collaborative¹⁹. Dans sa version la plus révolutionnaire, il imagine même des compositeurs qui proposeraient des œuvres inachevées, à contre-courant de ce qui a, selon lui, toujours fait le succès de la musique populaire, à savoir « la découverte par le compositeur du noyau qui touche tout le monde de la même manière ». Il fait ainsi référence à une expérience commune partagée par tous. À l'avenir, chaque auditeur pourrait compléter, personnaliser et finaliser en fonction de sa psychologie ou de sa personnalité. L'auditeur ne serait plus dans une posture passive mais vivrait ainsi une expérience totale et immersive.

¹⁴ DREDGE, Stuart. « [Spotify has six years of my music data, but does it understand my tastes?](#) », *The Guardian*, 6 janvier 2015.

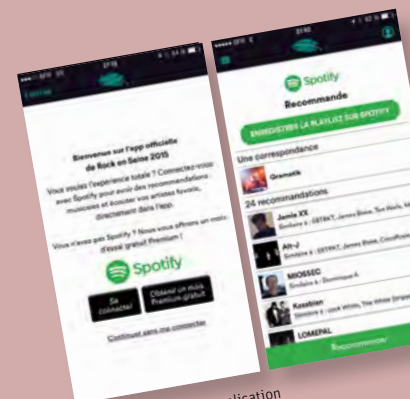
¹⁵ GOLDSTEIN, Brett. « [Spotify, we need to talk about your data](#) », *Medium*, 19 août 2014.

¹⁶ Sur ce thème, voir le projet Mes Infos de la FING. FING, [Self Data, cahier d'exploration Mes Infos](#), mai 2015.

¹⁷ Focalisation sur le développement d'une identité propre, à la manière d'une marque.

¹⁸ Diode électroluminescente

¹⁹ « [Le monde de demain selon Tod Machover](#) », *SoonSoonSoon*, mars 2014.



Capture d'écran de l'application mobile Spotify, 07/2015

DE LA VOD AUX CONTENUS ADAPTÉS ET ADDICTIFS

“ Personne, personne – ni maintenant, ni jamais – n’a la moindre idée de ce qui va ou ne va pas marcher au box office. ”

William Goldman
auteur, scénariste et dramaturge américain

- ▶ « Dis moi ce que tu regardes et je te proposerai de nouveaux contenus immersifs et addictifs. »
- ▶ L'apparition d'offres viables de vidéos à la demande sur abonnement (SVOD) bouleverse les équilibres de la diffusion et les pratiques de consommation : dans ce domaine, tout a changé d'un coup.
- ▶ Pour vivre et prospérer la SVOD a besoin de fidéliser les utilisateurs, de leur donner un sentiment de facilité et de confort par la prescription.
- ▶ La recommandation est un enjeu majeur du secteur, et Netflix, leader mondial, en a fait son cheval de bataille à la fois technologique et marketing.
- ▶ Les données remontent rapidement la chaîne de valeur, pour devenir aujourd'hui un élément non négligeable de la prise de décision de production des contenus : les films et séries, très chers à produire, sont de plus en plus calibrés par des algorithmes pour être des succès.

OÙ EN SOMMES-NOUS ET OÙ VA-T-ON ?

La vidéo omniprésente et l'ogre YouTube

La vidéo est devenue un type de contenu incontournable sur internet. Les médias et journalistes s'en servent de plus en plus et le marketing ne jure que par son pouvoir d'engagement – certaines études montrent que les contenus promotionnels sur Facebook ont un taux d'engagement 10 fois plus fort s'ils sont accompagnés d'une vidéo¹. Le phénomène de l'hébergement de vidéos personnelles, incarné par YouTube et le français DailyMotion, a lui aussi changé le rapport des individus à la vidéo. Autrefois, les personnes étaient de simples consommateurs de vidéos ; aujourd'hui, des millions de vidéos créées par les internautes sont

¹ Social Media Club. *Vidéo en ligne, la nouvelle cash machine*, octobre 2013.

² C'est-à-dire des acteurs qui peuvent atteindre directement le spectateur sans passer sous les fourches caudines de réseaux de diffusion classiques comme ceux des opérateurs

³ « n'importe quand, n'importe où, n'importe quel terminal, n'importe quel contenu »

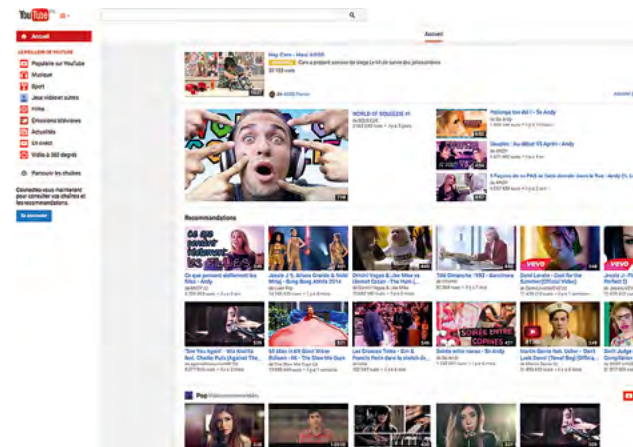
⁴ Sources : Netflix et Statista

diffusées, vues, partagées. YouTube a désacralisé la création vidéo, transformant des milliers d'anonymes en producteurs (parfois avec un grand succès) et même en chaînes de télévision. Les professionnels ont inventé nombre d'expressions étranges pour rendre compte de cette nouvelle réalité. Que l'on parle d'acteurs « OTT » (Over-the-top²) ou de la tendance « ATAWADAC » (any time, anywhere, any device, any content³), la transformation numérique a d'ores et déjà rompu beaucoup de digues dans ce secteur très encadré et très organisé.

La démultiplication des écrans confirme cette impression, smartphone ou tablette s'imposant comme second écran incontournable. Ces transformations interviennent dans un environnement massivement financé par la publicité, laquelle connaît d'ailleurs des innovations radicales. Par exemple, certains annonceurs testent des outils permettant au spectateur d'acquiescer directement un produit vu dans un programme, tel qu'un vêtement porté par un acteur.

La révolution SVOD (Subscription Video on Demand)

L'histoire de Netflix est désormais bien connue : créée en 1997, l'entreprise était spécialisée dans la location de DVDs expédiés par la poste. Très rapidement Reed Hastings, le fondateur, a fait le pari de la vidéo en ligne à travers un modèle d'abonnement illimité payant. Avec près de 60 millions d'abonnés payants dans le monde⁴, Netflix est aujourd'hui une plateforme dominante au niveau international de la « SVOD » (vidéo à la demande sur abonnement), et en est même devenu le symbole car l'entreprise n'a pour le moment pas d'autre activité ni d'autre modèle d'affaires, contrairement à ses concurrents les plus connus. Le succès du modèle de Netflix s'explique par son large



Capture d'écran site YouTube, 08/2015

catalogue ; les consommations constantes sont à la base d'un modèle payant qui fonctionne, même si certains analystes restent dubitatifs quant à son équilibre à moyen terme.

Au cœur des différents enjeux de lutte contre le *churn*⁵, la donnée trouve toujours une place de choix. Selon le site [GigaOM](#), *Netflix* peut en particulier savoir pour tous ses utilisateurs les lectures, pauses, retours et avances rapides, les notes qu'ils attribuent aux programmes, les recherches effectuées, le moment de la séance, l'appareil utilisé, etc.⁶ Dès 2011, [ReadWriteWeb](#) indiquait que la plateforme de vidéo en streaming stockait 50 fichiers différents pour chaque média, incluant 10 ans de notations par les utilisateurs, des données étendues concernant les comptes des utilisateurs l'ayant visionné, les droits de propriété intellectuelle, les fichiers audios, les fichiers de log ou encore les sous-titres⁷.

Avant tout, ces données servent à remporter ce que *Netflix* appelle les « moments de vérité », c'est-à-dire ces moments où un utilisateur a le choix entre utiliser *Netflix* et faire autre chose. Selon la [vision de long terme de l'entreprise](#) publiée en janvier 2015⁸, ces moments se gagnent « quand le membre s'attend à ce que *Netflix* soit un choix plus satisfaisant que les autres options, selon ses expériences précédentes. Cette satisfaction vient de la facilité de choix (...) et de la disponibilité de contenus qui correspondent aux goûts et à l'humeur de tout le monde dans le foyer ». De ce point de vue, [selon l'auteur et réalisateur Jason Lange](#), *Netflix* a construit une infrastructure qui relie les données sur ce que les gens regardent avec les équipes qui décident de ce qui est produit⁹.

La donnée comme guide du développement du catalogue : la production de contenus « *data driven* ».

Le premier enjeu pour un service de SVOD est de fournir un catalogue satisfaisant à ses abonnés, et donc pour cela de négocier les droits de diffusion de contenus pour alimenter ce catalogue. L'autre voie est de produire ses propres contenus originaux, comme le font la majorité des diffuseurs. *Netflix* dit utiliser (en tout cas se vante d'utiliser) massivement les *data sciences* pour orienter les créations de contenus. C'est en fait probablement dans ce secteur de la vidéo que la production est d'ores et déjà la plus orientée par les données. En effet, si pour le moment *Spotify* ou *Deezer* produisent peu de musique eux-mêmes, les diffuseurs télévisuels traditionnels

(comme Canal+ ou HBO) ou nouveaux (comme *Netflix* ou Amazon) sont des acteurs majeurs de la production de contenus originaux. L'histoire est souvent rappelée : *House of Cards*, l'immense succès de production de *Netflix* et premier contenu original de l'entreprise serait un pur produit du big data et des data sciences. En effet, selon des dizaines d'articles en ligne¹⁰, *Netflix* aurait décidé de financer le projet *House of Cards* car il combinait trois ingrédients à succès selon des algorithmes maison : un remake d'une ancienne série britannique, David Fincher à la réalisation et Kevin Spacey comme acteur principal. Dès lors, la décision de financer la production de la version américaine d'*House of Cards* devenait rationnelle... et *data-driven*. Kevin Spacey lui-même renforce cette histoire lorsqu'il raconte les échanges avec *Netflix* : « *Netflix* a été le seul réseau à nous dire : nous croyons en vous. Nous avons fait tourner nos données, et elles nous disent que notre public regarderait cette série. Nous n'avons pas besoin d'un épisode pilote. Combien voulez-vous en faire ? »¹¹. Bien évidemment un tel récit peut laisser dubitatif, ne serait-ce que parce qu'il ne semble pas nécessaire de déployer un algorithme puissant pour deviner qu'une œuvre de David Fincher avec Kevin Spacey et un bon scénario d'intrigue politique risque de conduire à un succès. Toujours est-il que *Netflix* peut se vanter d'être en capacité de faire ce travail, en se basant sur la grande quantité d'informations récoltées sur les usages, qui leur permettrait « de savoir ce que les gens veulent voir avant qu'ils ne le sachent », selon une expression [utilisée par David Carr dans le New-York Times](#)¹².

Si certains doutent de la capacité du *big data* à deviner ce qui va marcher, d'autres s'inquiètent au contraire du risque que cela comporte pour la diversité culturelle. Le journaliste Andrew Leonard [se demandait ainsi sur le site Salon](#) ce qu'il adviendrait lorsque les producteurs pourront parler avec les scénaristes en sachant qu'une partie des abonnés sont opposés à certaines méthodes de tournage (comme les plans-séquences ou les « jump cuts ») ou sont plus intéressés par tel ou tel type de scène : « L'auteur peut-il survivre dans une époque où les algorithmes informatiques sont le *focus group*¹³ ultime ? »¹⁴.

Les excès de la personnalisation, que l'on peut déjà ressentir sur les sites internet de contenus médiatiques et de presse, sont peut-être un futur possible dans le domaine de l'écriture télévisuelle et cinématographique (voir partie III). Certes cela paraît encore bien prospectif, mais

⁵ Le *churn* est un anglicisme qui renvoie au taux d'abandon du service, au taux de désabonnement.

⁶ HARRIS, Derrick. « [Netflix analyses a lot of data about your viewing habits](#) », *GIGAOM*, juin 2012

⁷ KIRKPATRICK, Marshall. « [Netflix's big data plans to take over the world](#) », *ReadWriteWeb*, juillet 2011

⁸ Site officiel *Netflix* : <http://ir.netflix.com/long-term-view.cfm>

⁹ LANGE, Jason. [Netflix, big data and what Hollywood is missing out on](#), septembre 2012.

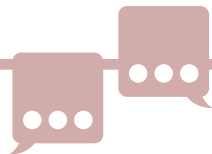
¹⁰ Un exemple parmi tous ceux-ci : ANAYA, Jeff. « [How big data and social media made House of Cards a hit](#) », Blog de *CISION*, février 2014

¹¹ AULETTA, Ken. « [Outside the box](#) », *The New Yorker*, février 2014

¹² CARR, David. « [Giving viewers what they want](#) », *The New York Times*, février 2013

¹³ En marketing, technique d'étude qualitative qui consiste à rassembler une demi-douzaine d'individus consommateurs ou prescripteurs pour comprendre en profondeur leurs attitudes ou comportements à l'égard d'une offre.

¹⁴ LEONARD, Andrew. « [How Netflix is turning viewers into puppets](#) », *Salon*, février 2013



3 QUESTIONS À... ÉRIC SCHÉRER, DIRECTEUR DE LA PROSPECTIVE, GROUPE FRANCE TÉLÉVISIONS

Comment qualifieriez-vous la transformation des usages des contenus audiovisuels actuelles ?

En Suède, un quart de la population a changé d'usages TV en deux ans ! Le tiers de la consommation TV y est déjà en streaming, le temps passé sur la TV linéaire recule très vite (-10% par an). Près de 30% des Suédois sont abonnés à une offre de SVoD. Au Royaume Uni, 20% de la consommation de TV se fait déjà en différé. Aux États-Unis, plus de 60% de la conso des jeunes adultes US se font en différé ! Pour l'ensemble de la population le ratio est déjà de 50/50 ! Un ratio qu'on attendait il y a peu encore pour 2020 !

Et du côté de l'offre, quels sont selon vous les tendances lourdes ?

La télé vit le « moment Ketchup » de sa transformation numérique : côté production, diffusion et surtout consommation d'infos et de divertissements, tout déboule très vite, et en même temps. Fiction, séries, magazines, jeux, sport : des millions de nouveaux concurrents sont apparus dans l'offre. Nous sommes au tout début d'une incroyable hyper-offre, d'un foisonnement de contenus vidéo sur tous les écrans. Et en streaming ! D'autant que de nouvelles plateformes se disputent notre attention, notamment les messageries instantanées, les objets connectés sur soi, demain l'Internet industriel et les voitures autonomes. Mais l'OTT libère la TV de ses contraintes historiques. Il faut aussi inventer d'autres modèles hybrides,

notamment ceux qui vont raccourcir la durée entre le concept et sa commercialisation (comme YouTube et Netflix le font).

Dans votre cahier de tendance « la TV de demain, 10 enjeux de transformation », vous affirmez le passage du PAF au PAP, le paysage audiovisuel personnalisé. Quelle est la place de la personnalisation et donc des données dans ce paysage ?

Dans ce nouveau monde de temps d'antenne illimité, les trois ressorts de la consommation à la demande sont la personnalisation, la simplicité d'accès et la qualité des contenus.

La télévision devient une expérience personnelle sur écrans individuels où les attentes, en matière de découverte, d'accès et d'expérience se transforment aujourd'hui très vite. Le web audiovisuel est en train de basculer du « search » (on cherche, mais on a du mal à trouver) au « push automatisé ». Qui l'emportera à l'heure de la programmation éditoriale, la recommandation sociale ou la prescription algorithmique ? L'idéal serait un mix du meilleur des trois mondes (éditeurs, amis et algorithmes) combiné à la liberté de la consommation à la demande et à la simplicité de la consommation en linéaire: en somme un « push personnalisé ».



Cindy Holland, responsable des contenus originaux pour Netflix, [a déjà pu dire en juin 2014](#) qu'il était possible de comprendre, grâce aux données, que les abonnés qui regardent en moyenne 2 épisodes et demi par soir, peuvent trouver répétitive l'utilisation de musiques similaires dans chaque épisode¹⁵.

Les contenus cinématographiques, riches en données mobilisables pour mieux cerner les goûts et envies de l'utilisateur

Une autre caractéristique des contenus cinématographiques est leur grande richesse en données descriptives (voir la typologie dans le chapitre II-1) : réalisateur, date, casting, lieu de tournage, genre et sous-genre, notes, critiques, anecdotes, musique originale... Autant d'informations qui peuvent être utilisées pour recommander des contenus aux spectateurs. Netflix revendique près de 77 000 micro-genres de films ou séries (i.e. : « films d'époque émouvants avec un personnage féminin » ou « films visuellement frappants sur les relations pères-fils »), dans une logique que [le journaliste Alexis Madrigal avait qualifié](#) de « rétro-ingénierie d'Hollywood »¹⁶. Mise en perspective avec les habitudes de consommation des clients, cette connaissance permettrait selon Netflix d'améliorer fortement

les mécanismes de recommandation. Canal+ adopte d'ailleurs une logique similaire. Pour Cécilia Frandjan (responsable recommandation au sein du Groupe Canal+), qualifier les contenus permet de recommander avec précision non seulement les catégories et les sous-genres, mais également les tonalités, les ambiances, les environnements, propres à chaque contenu et donc enrichir « Eureka », le moteur de recommandation lancé en 2011¹⁷, rebaptisé *Suggest*¹⁸. Mais Canal+ mobilise également ses équipes éditoriales et critiques de cinéma, qui sont chargées de réunir des contenus par thème et donc de jouer un véritable rôle de prescription experte. ■

SUGGEST

LA MEILLEURE FAÇON DE REGARDER LA TÉLÉVISION

SUGGEST est le nouveau moteur de recommandations du Groupe CANAL+. Il s'appuie sur le meilleur de la technologie mais également sur le savoir-faire d'experts éditoriaux dans chaque thématique (cinéma, séries, jeunesse, etc.).

Disponible à l'automne 2015 sur télévision via LE CUBE S, SUGGEST sera progressivement déployé sur second écran (MyCANAL) ainsi que sur la prochaine génération de décodeurs satellite du Groupe CANAL+.

En alliant la puissance de l'intelligence artificielle à la créativité de l'intelligence humaine, SUGGEST établit un graphique de corrélation de programmes d'une richesse éditoriale inégalée. Intuitif et immersif, il offre une "expérience télé" évolutive, à l'image de son utilisateur.

Extrait communiqué de presse "La nouvelle expérience Canal+", 06/2015.

¹⁵ ROSE, Lacey. « Netflix's Original Content VP on Development Plans, Pilots, Late-Night and Rival HBO », *The Hollywood Reporter*, juin 2014

¹⁶ MADRIGAL, Alexis. « How Netflix reverse engineered Hollywood », *The Atlantic*, janvier 2014

¹⁷ Social Media Club France, « La recommandation TV : comment nous sont proposées les vidéos sur le web ? », *ZDnet*, 13 janvier 2015

¹⁸ Canal+ a présenté en juin 2015 sa nouvelle stratégie numérique dont l'une des briques est la fonction de personnalisation *Suggest* qui alliera « technologies et savoir-faire d'experts éditoriaux » selon la chaîne cryptée. Pour plus d'informations, voir le [communiqué de presse](#).

The image shows a screenshot of the Netflix website. At the top left is the 'NETFLIX' logo. The main headline reads 'Profitez de films et séries TV où et quand vous le souhaitez.' Below this, it says 'Forfaits à partir de 7,99 € par mois.' and a blue button that says 'Démarrez votre mois gratuit'. The background of the page features a woman with long dark hair sitting at a table in a cafe, looking at a tablet. At the bottom of the page, there are three small icons with text: 'Tous les contenus sont disponibles sur PS3, Wii, Xbox, PC, Mac, smartphone.', 'Tous les contenus sont instantanément diffusés en streaming.', and 'Sans engagement, abonnement annulable en ligne à tout moment.'

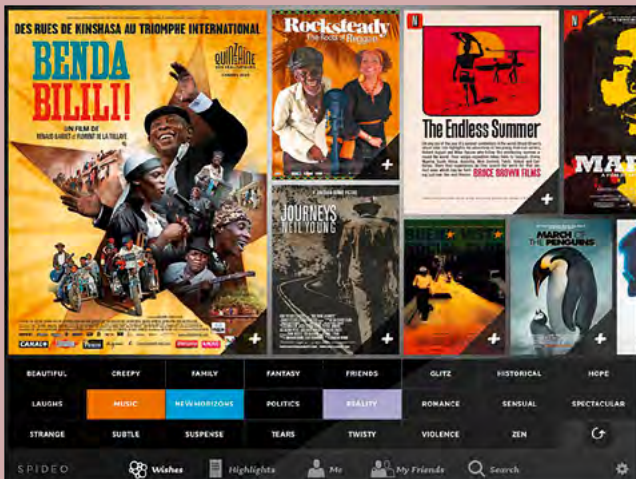
Capture d'écran site web Netflix, 06/2015



USAGES ÉMERGENTS

1 DES RECOMMANDATIONS PLUS INNOVANTES, PAR EXEMPLE SELON L'HUMEUR : [SPIDEO MOOD-DISCOVERY](#)

► Les français de la start-up Spideo n'hésitent pas à affirmer vouloir « battre Netflix » à son propre jeu¹⁸, en proposant de compléter la recommandation algorithmique de découverte basée sur l'humeur : leur application offre 20 « vœux » à l'utilisateur, qu'il va alors combiner à sa guise (par exemple « nouveaux horizons », « espoir » et « romance »).



capture d'écran site web Spidéo, 09/2015

2 DES CONTENUS TRANSMEDIAS EN CASCADE : WEBSÉRIES ([STUDIO 4.0](#) DE FRANCE TELEVISION), RÉSEAU MULTI-CHAINES SUR YOUTUBE ([AWESOMENESS TV](#)), PLATEFORMES D'AUTO-PRODUCTION/DIFFUSION ([TWITCH](#))

► Les contenus sont produits partout et par une multitude de créateurs, dans des contextes qui brouillent les frontières entre les médias. Ainsi, le succès de certaines webséries (*Le Visiteur du futur*, *Hello Geekette*, *Noob...*) est tel que France Télévisions a créé le Studio 4 pour travailler sur ces nouvelles formes narratives. La disparition des frontières se constate également à travers le succès des réseaux multi-chaines sur YouTube : Awesomeness TV, créée spécialement pour développer la visibilité de contenus à destination du jeune public, réunit pas moins de 90 000 créateurs et 100 millions d'abonnés pour plus d'un milliard de vues¹⁹. Enfin, d'autres formats encore plus disruptifs ont un grand succès, comme la diffusion en direct de parties de jeux vidéo sur la plateforme Twitch, acquise par Amazon ou sur le tout récent YouTube Gaming. Ces plateformes mélangent dans la même offre vidéo en live, commentaires par messagerie instantanée, contenus publicitaires...

3 DES HISTOIRES IMMERSIVES ET SENSORIELLES : [OCULUS STORY STUDIO](#) ET [SENSORY STORIES](#)

► La réalité virtuelle ou la réalité augmentée font leur grand retour grâce à plusieurs initiatives de terminaux (*Oculus Rift*, *Microsoft HoloLens*... voir chapitre II-5). Ces nouveaux médias interpellent les contenus audiovisuels autant que les contenus vidéo-ludiques. Selon [Eric Schérel](#)²⁰, ils font émerger « une grammaire, une syntaxe, un langage, des formes narratives inédites et toute une chaîne de production et de diffusion à construire ». Nancy Bennett (citée dans le cahier Méta-media #9) considère quant à elle que « l'immersion est le nouvel engagement ». Des formes adaptées de narration doivent être imaginées, ce que fait par exemple *Oculus* en créant en son sein un studio dédié, le « [story studio](#) ». Des manières nouvelles d'immerger le spectateur dans l'histoire peuvent également apparaître au croisement des contenus audiovisuels et des objets connectés comme le montre l'exposition [Sensory Stories du Museum of Moving Image](#) de New York. Dans cette exposition, le think-tank *Future of storytelling* présente des projets montrant comment, demain, les histoires pourront aller « au-delà de l'écran ». Antonin Lhote (conseiller éditorial et chef de projet web à la Direction des nouvelles écritures et du transmédia de France TV) indiquait d'ailleurs [dans le cahier Méta-media #9](#) que « de tous les sillons creusés par les nouvelles narrations digitales, celui qui s'appuie sur le développement de technologies propres aux sens semble regorger de possibilités »²¹. Et il regorgera également de données...

¹⁸ SPIDEO. [Beat Netflix at its own game](#), mai 2015

¹⁹ CHAZELLE, Barbara. « [Awesomeness TV : misez déjà sur la génération Z!](#) », *Méta-média #9*, juin 2015, pp. 70-71

²⁰ SCHERER, Eric. « [Plongeon dans l'immersion](#) », *Méta-média #9*, juin 2015, pp. 80-83

²¹ LHOPE, Antonin. « [Sensory stories : expériences sensorielles dédiées à l'innovation narrative](#) », *Méta-media #9*, juin 2015, pp. 88-91

DES JEUX EN LIGNE AUX INTERACTIONS HOMME-MACHINE AUGMENTÉES DANS DES ENVIRONNEMENTS IMMERSIFS

“ Le jeu le plus important des 5 prochaines années sera un jeu où vous ressentirez une forte empathie avec les personnages (...) ce sera un énorme succès, et résultat, il sera comme le « Titanic » [NDLR : le film, pas le navire...] de l'industrie du jeu (...) Parce que vous aurez une vraie relation avec les gens au lieu de juste leur tirer dessus. ”

George Lucas

Conférence à l'Université de Californie du Sud, juin 2013

- ▶ « Laisse-moi analyser comment tu te comportes et je te proposerai des contenus hyper-personnalisés et addictifs. »
- ▶ Paradoxalement, ce secteur « naturellement numérique » a connu une mutation plus lente. Elle s'accélère maintenant grâce à l'accès au haut débit).
- ▶ Les modèles économiques du jeu vidéo, qui étaient jusqu'alors relativement simples, sont en train de se diversifier et de s'hybrider à grande vitesse, avec d'un côté le spectre des jeux « gratuits » qui ressemblent de plus en plus à des jeux de casinos dans leur dynamique, et de l'autre côté des « hits » dont les coûts de production dépassent ceux des films hollywoodiens et rendent tout échec commercial impensable.
- ▶ Les joueurs passent plusieurs dizaines d'heures à être actifs « dans » le jeu (aujourd'hui au sens figuré, demain de plus en plus au sens propre grâce aux nouvelles interactions virtuelles ou augmentées) : les jeux vidéo sont donc de formidables outils d'analyses des réactions et des comportements.

¹ CHANTEPIE, Philippe, MICHAUD, Laurent, SIMON, Laurent et ZACKARIASSON, Peter. « [Video game business models and monetization](#) », *Digiworld Economic Journal*, n°94, juin 2014.

² RAYNA, Thierry et STRIUKOVA, Ludmila. « [Few to many : change of business model paradigm in the video game industry](#) », *Digiworld Economic Journal*, n°94, juin 2014.

OÙ EN SOMMES-NOUS ET OÙ VA-T-ON ?

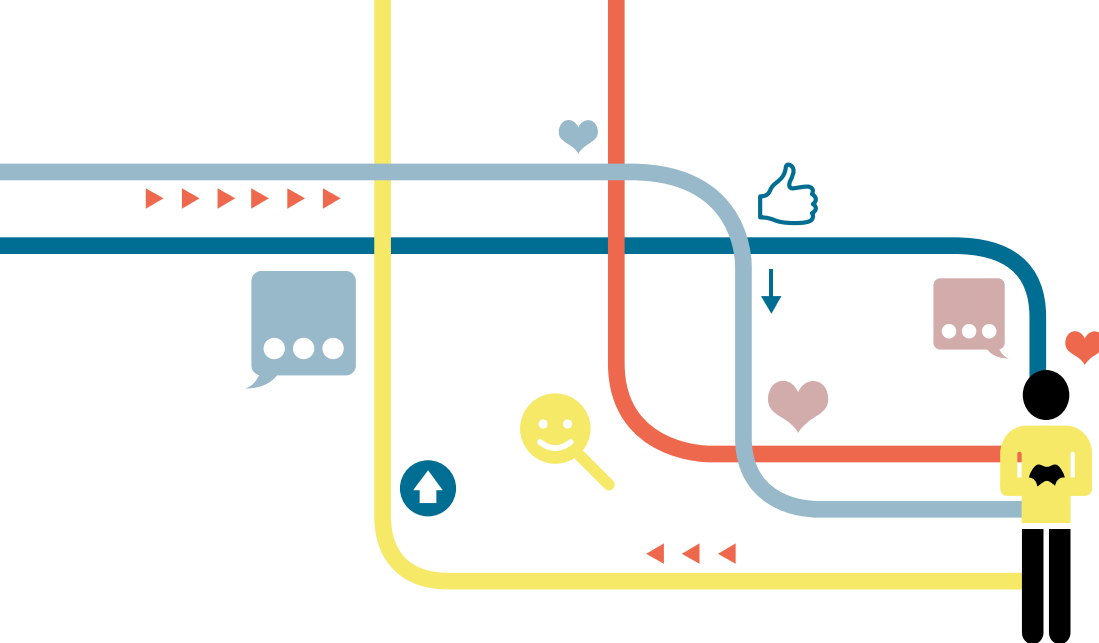
A première vue, des quatre secteurs qui retiennent notre attention dans ce cahier, le secteur du jeu vidéo est celui qui paraît le mieux armé culturellement et technologiquement pour accompagner la transformation numérique de la production, de la distribution et de la consommation de ses contenus. Pourtant, comme le souligne [l'introduction du numéro 94 de la revue Digiworld Economic](#)

[Journal](#), « on peut affirmer que paradoxalement, l'industrie du jeu vidéo, « native » du numérique et secteur le plus jeune et le plus dynamique des industries des loisirs, ne fait que rentrer dans la révolution numérique¹. En effet, la dématérialisation et la démocratisation de la distribution des produits commencent seulement à modifier la chaîne de valeur de ce secteur et à renforcer la tendance au *data-driven*.

Par ailleurs, cette industrie connaît une mutation des interfaces. La manette, le couple clavier-souris, ou encore les écrans sont aujourd'hui complétés par des nouveaux dispositifs d'interactions tels que les manettes avec capteurs de mouvement de la Wii de Nintendo (WiiMote, Balance Board...) ou la Kinect de Microsoft basée sur une caméra capable de capter les mouvements du corps ; en outre, les écrans tactiles sont évidemment un autre marqueur de cette tendance. Les précurseurs de cette voie, s'ils ont eu parfois du mal à convertir l'innovation en succès commercial dans la durée, annoncent toutefois la modification profonde des interactions entre l'humain et les ordinateurs dans l'espace vidéoludique. La tendance est en effet au rapprochement progressif des dispositifs et du corps, par des capteurs et des interfaces plus immersifs (retour de force, réalité virtuelle ou augmentée...)

Les données disponibles dans ce domaine permettent avant tout d'analyser le comportement des joueurs, de vérifier ce qui plaît et, dans le cadre d'une mise à jour du jeu, de conserver l'engagement. Le contenu des créations se trouve donc influencé par des goûts collectifs, identifiés par l'analyse d'une grande quantité de données. La mise au jour de ces goûts collectifs tend dorénavant à se doubler d'une connaissance unitaire de chaque joueur et donc d'une personnalisation possible de l'expérience en temps réel.

Certes, comme [le soulignent Thierry Rayna et Ludmila Striukova](#)², les techniques à la disposition des éditeurs de jeux leur permettent déjà en théorie de personnaliser de manière très avancée les offres, mais les grands éditeurs ont encore peu adopté ces pratiques. Les éditeurs de jeux sur les réseaux sociaux ou sur smartphone, comme Zynga ou King, ont saisi cette opportunité de manière plus rapide (voir encadré infra). En effet, dans le domaine du jeu sur smartphone et tablette, la tendance majeure ►►►



TENDANCES CLÉS

- L'essor des environnements immersifs (plus ou moins complexes)
- La diversité croissante des modèles économiques, des plateformes, des formats, etc.
- La personnalisation croissante des offres « additionnelles » (bonus, contenus spécifiques, offres spéciales, etc.)

DONNÉES CLÉS



AUJOURD'HUI

- Les statistiques du jeu, les choix de l'utilisateur (consentement à payer des options, des bonus, des contenus additionnels) dans un contexte le plus dynamique possible.



DEMAIN ?

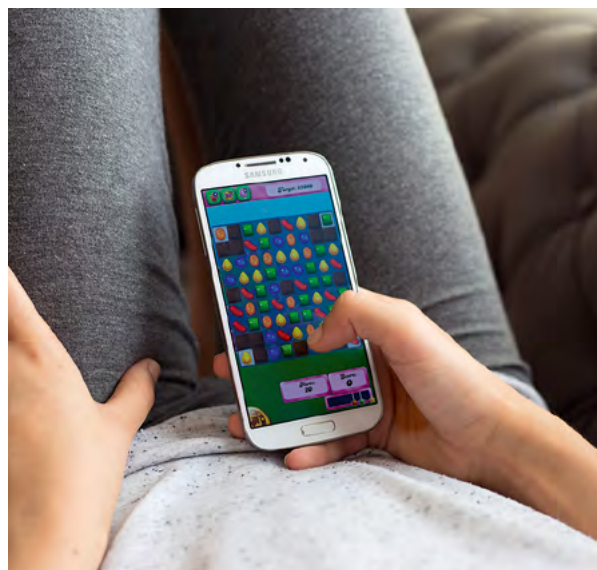
- Les comportements de l'utilisateur, ses « préférences », ses choix, ses émotions, en temps réel. Des informations venant directement de capteurs corporels (stress, rythme cardiaque, mais aussi ondes cérébrales).

est au jeu « *free to play* » (F2P), c'est-à-dire obtenu gratuitement. La monétisation, quelle qu'en soit la forme (contenus additionnels, bonus, version étendue, publicité, ...) a donc lieu pendant l'utilisation du jeu (voir chapitre I-2) et non pas lors de l'acquisition. Économiquement, l'équilibre de marché en est bouleversé : au lieu de pouvoir se fier à une demande globale, agrégée, les éditeurs doivent évaluer une demande individuelle extrêmement contextuelle et éventuellement

fugace. Selon l'expression de Myriam Davidovici-Nora³, le modèle « F2P » requière donc « une micro-gestion dynamique de chaque joueur et de son expérience de jeu » pour être rentable. Les éditeurs de jeu ont une forte incitation à garder le joueur connecté, afin d'accroître les chances de monétisation. La fine compréhension des ressorts psychologiques des joueurs est donc bien plus qu'auparavant un enjeu commercial : dans le monde du jeu vidéo, le *neuromarketing* n'est

LES JEUX « GRATUITS » SUR SMARTPHONES, SYMBOLES DE L'USAGE DES DONNÉES POUR FAIRE DÉPENSER DE L'ARGENT AU JOUEUR

Candy Crush Saga est un des jeux les plus populaires des dernières années avec des dizaines de millions d'utilisateurs quotidiens. Initialement développé pour le réseau social Facebook, le casse-tête à base de bonbons a été adapté sur *Android* et *iOS* et son succès est tel qu'il a largement contribué à l'entrée en bourse de sa société d'édition King, début 2014. *Candy Crush Saga* se caractérise aussi par sa capacité inédite de collecte des données de ses utilisateurs. Son modèle repose sur du Free-2-Play où la limitation à 5 vies par jour peut être débloquée par un paiement ou de manière plus originale en acceptant de visionner une publicité ou invitant des amis à jouer. Le carnet d'adresses fait alors office de chéquier. Étant par ailleurs relié au profil social du joueur (Facebook) tout un ensemble d'autres données sont mobilisables. Un [article du Wall Street Journal du 30 avril 2015](#)⁴ sur les développeurs de ce jeu souligne ainsi combien les développeurs prennent grand soin à analyser les données générées par les parties des utilisateurs afin de trouver le bon équilibre entre le défi qui procure une gratification et celui qui fera renoncer le joueur... En effet, la source principale de revenus de King - l'éditeur - provient des bonus permettant d'avancer dans le jeu plus vite. Le consentement à payer des utilisateurs sera donc très directement corrélé aux sentiments de frustration et de gratification qu'ils recevront du jeu. C'est pourquoi, selon Jens Hansegard, les designers utilisent le nombre moyen de tentatives nécessaires pour passer d'un niveau au suivant afin d'ajuster la difficulté des niveaux anciens et d'éviter de reproduire les caractéristiques des niveaux qui ont vu un grand nombre de joueurs abandonner.



Un succès plus récent mais tout aussi fulgurant est le jeu *Clash of Clans* qui [pourrait rapporter jusqu'à 2 milliards de dollars à SuperCell de recettes](#), son éditeur, en 2015. Là aussi, le modèle économique est d'attirer par un jeu d'apparence traditionnel des utilisateurs vers ce qui utilise en fait ensuite les mécanismes psychologiques du jeu d'argent : « maintenir l'addiction, – en reprenant les recettes utilisées par d'autres incontournables (tels *Candy Crush Saga*, *Puzzle & Dragons*, *Angry Birds*), et distiller régulièrement des moments de satisfaction pour le joueur »⁵.

³ DAVIDOVICI-NORA, Myriam. « [Paid and Free Digital Business Models Innovations in the Video Game Industry](#) », *Digiworld Economic Journal*, n°94, juin 2014.

⁴ HANSEGARD, Jens. « [The drama behind Candy Crush Saga : creating new levels](#) », *Wall Street Journal*, 30 avril 2015.

⁵ RAYNAL, Adeline. « [Clash of Clans, l'art de devenir un blockbuster... et une machine à cash](#) », *FrenchWeb*, 15 juin 2015.

plus vraiment une prédiction ou une hypothèse, c'est une réalité émergente que tous les acteurs comprennent, sans toutefois l'aborder avec la même urgence.

Au-delà du simple repérage de ce qui plaira et de ce qui conduira le joueur à consentir à payer, va-t-on faire un usage des techniques de neuro-marketing pour faire de l'incitation douce ? Dans un article intitulé « [comment le jeu vidéo vous manipule](#) »⁶ le *game designer* Oscar Barda explique ce passage du A/B testing au A/B profiling. Le test A/B est une technique, répandue dans les services numériques, qui permet de tester plusieurs versions⁷ qui diffèrent d'un seul critère d'un même produit ou de sa présentation commerciale. On peut tester par exemple plusieurs formules d'incitation à l'achat ou plusieurs agencements de la page d'achat sur un site de commerce auprès de différents échantillons des visiteurs, afin de choisir empiriquement la plus efficace. Ces méthodes sont également très utilisées dans le jeu vidéo : plutôt que de cantonner les choix aux murs d'une salle réunissant les concepteurs du jeu et au mieux quelques « clients tests », des tests grandeur nature sont organisés auprès des joueurs. Une décision peut ensuite être prise en fonction des objectifs testés. Avec le *A/B profiling*, il ne s'agit plus seulement de tester quelques variantes et de trouver la meilleure, mais bien de « fournir une formule à chacun et de considérer que chaque formule est la bonne pour la personne en question, explique Oscar Barda. Dès lors que le joueur est connu personnellement au travers de données sociales, de préférences ou de comportements, le jeu pourra s'y adapter : « Vous êtes une femme, 34 ans, habitant en Californie qui aime [sur les réseaux sociaux] les pages des restaurants végétariens autour de chez elle ? Très bien, notre ordinateur nous dit que vous dépenserez plus d'argent avec un bouton vert, et hop ! ». Ces informations peuvent également être utilisées pour chercher des profils similaires au joueur analysé, et ainsi reproduire les conditions au sein du jeu qui les ont amenés à effectuer des achats.

Comme l'expliquent Thierry Rayna et Ludmila Striukova⁸, les éditeurs commencent donc à considérer leurs utilisateurs en tant qu'individus, et non plus en tant que segments : « (...) bien que Sony collecte des données concernant des modèles d'usages via le *Playstation Network central*, cela n'a pas encore abouti à des offres augmentées ou personnalisées. ». Certes, tous les acteurs du secteur ne se sont pas encore approprié tous ces possibles mais de plus en plus de concepteurs de jeux mobilisent la

connaissance individuelle de leurs joueurs pour leur proposer des publicités personnalisées, au sein du jeu, s'appuyant sur la localisation des joueurs ou leur comportement.

Du point de vue de l'analyse économique, le marché du jeu vidéo est un formidable « terrain d'expérimentation » pour reprendre l'expression de Yanis Varoufakis, ancien économiste chez Valve, l'éditeur de la plateforme dominante de la vente de jeux vidéo en ligne sur PC (*Steam*)⁹. Sur une plateforme comme *Steam*, les prix sont aisés à adapter, la demande aisée à segmenter et il existe même des marchés de biens virtuels très dynamiques. Comme le souligne Martin Quinn, doctorant en économie à Telecom Paris¹⁰, *Steam* a une stratégie de promotions temporaires nombreuses et agressives, liées à un système de recommandation personnalisée de plus en plus présent dans l'interface. Cela permet à Valve d'étudier très finement le consentement à payer de ses clients et d'affiner les recommandations en augmentant la base de jeux installés par l'utilisateur. Ces données restent sous le contrôle de la plateforme de distributions et échappent donc aux éditeurs des jeux, selon un modèle se rapprochant de celui des autres secteurs des contenus culturels : ici encore, la plateforme de distribution se place au cœur de la valorisation des données.

Cette évolution n'est probablement qu'une première étape de ce qu'il paraît imaginable de faire. L'exploration de nouvelles sources de données devrait permettre de cerner au plus près les émotions et l'état physiologique du joueur, comme en témoigne [le projet de recherche](#) « FUN II » associant *Ubisoft Québec*, l'École de psychologie de la Faculté des sciences sociales de l'Université Laval et le Conseil de recherche en sciences naturelles et en génie du Canada¹¹. Peu de détails concernant ce projet croisant psychologie, neurosciences et analyse de facteurs physiologiques sont disponibles, mais une promesse ambitieuse : utiliser les émotions du joueur pour influencer en temps réel son expérience au sein du jeu en créant des jeux « adaptatifs » en fonction de l'état d'esprit des joueurs. Ce projet s'inscrit dans le contexte d'autres recherches sur la [modélisation des comportements des joueurs](#)¹². Futuriste ? Après tout, [dès 2007, Google a commencé à déposer ce type d'idées](#), évoquant par exemple l'idée que les données issues des jeux vidéo seraient intéressantes pour le ciblage publicitaire car les joueurs interagissent entre eux et prennent des décisions qui reflètent probablement leurs comportements dans la vie réelle.¹³ ■

⁶ BARDA, Oscar. « [Comment le jeu vidéo vous manipule](#) », *Rue89*, 17 mars 2015.

⁷ Par exemple, vous pouvez présenter un bouton vert « achetez ! » ou le même bouton en rouge, et si le bouton rouge obtient 5% de clic en plus, peu importe la raison, il est cohérent de le généraliser.

⁸ RAYNA, Thierry et STRIUKOVA, Ludmila. « [Few to many : change of business model paradigm in the video game industry](#) », *Digiworld Economic Journal*, n°94, juin 2014.

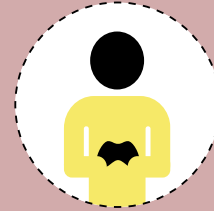
⁹ Et ancien ministre de l'économie en Grèce. Voir sur le [blog « Valve Economics »](#) la recension par Varoufakis de son arrivée dans l'entreprise, où il qualifie les travaux économiques traditionnels, qui se basent sur des modèles abstraits, « d'astrologie informatisée ».

¹⁰ Entretien avec les auteurs.

¹¹ Université Laval, « *Ubisoft Québec investit pour créer des jeux vidéo qui s'adaptent aux émotions des joueurs* », 29 mai 2014 ([Communiqué de presse](#))

¹² BAKKES, Sander, SPRONCK, Pieter et VAN LANKVELD, Giel, « [Player behavioural modelling for video games](#) », *Entertainment Computing*, janvier 2012.

¹³ ADAM, David et JOHNSON, Bobby. « [Google may use games to analyse net users](#) », *The Guardian*, mai 2007, cité dans CHAMPEAU, Guillaume. « [Google peut vous connaître à travers vos jeux vidéo](#) », *Numerama*, mai 2007.



USAGES ÉMERGENTS

Le jeu vidéo étant un média où les réactions psychologiques et émotionnelles des utilisateurs ont une importance cruciale. La détection directe ou indirecte des émotions deviendra probablement à l'avenir un objectif commun aux différentes catégories de jeux. Ces analyses permettraient d'adapter en temps réel certains éléments du jeu ou de ses contenus additionnels pour correspondre aux préférences et aux goûts du joueur. Dans ce domaine, le profilage émotionnel des joueurs permet, selon Pieter Spronck, d'établir automatiquement un profil psychologique ou sociologique par des mesures directes ou indirectes (en inférant un comportement par rapport aux actions du joueur dans le jeu).

L'arrivée croissante d'objets connectés adaptés ou dédiés au jeu permettront d'imaginer des environnements immersifs nouveaux. C'est d'ailleurs la prédiction de Steven Spielberg lui-même, qui a indiqué [lors d'une conférence en 2013](#) que jusqu'ici, les jeux n'ont pas été capables de créer la même empathie avec les personnages à l'écran que l'ont fait les formes narratives, « à la seconde où vous prenez la manette, quelque chose s'éteint dans votre cœur, et cela devient un sport »¹⁴.

¹⁴ COHEN, David. « [George Lucas & Steven Spielberg: Studios Will Implore; VOD Is the Future](#) », *Variety*, juin 2013.

¹⁵ GREENE, Kate. « [Connecting your brain to the game](#) », *Technology Review*, mars 2007

¹⁶ Les versions actuelles sont des versions de tests et de développement, plutôt destinées à des professionnels.

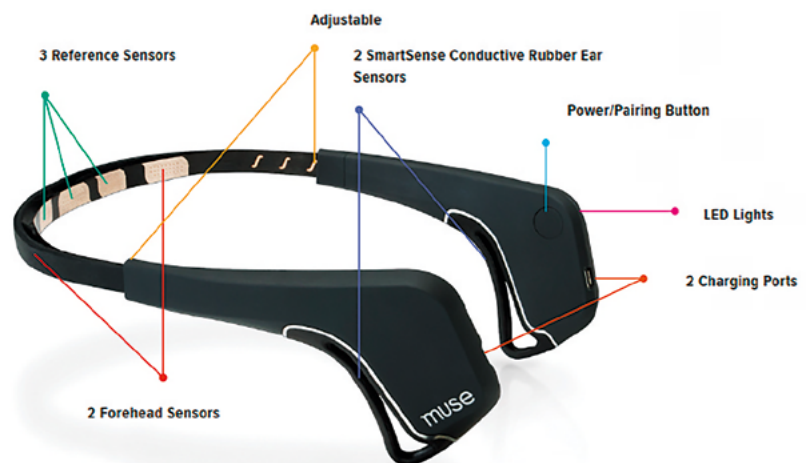
¹⁷ C'est-à-dire offrant des sensations de toucher et de perception kinesthésique - retour de force, résistance, ...

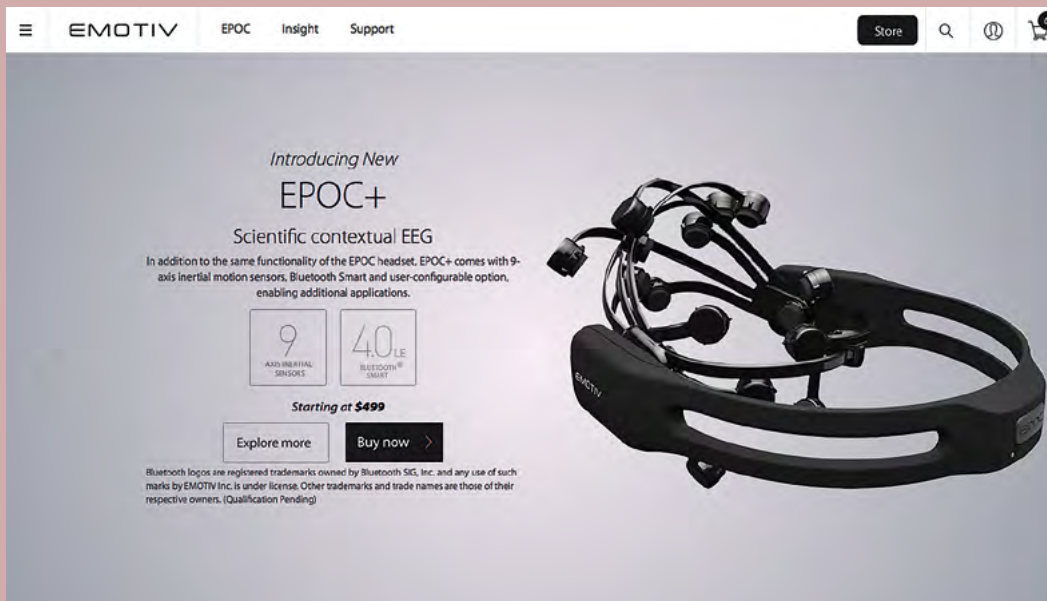
¹⁸ *SXSW 2015*, session : « [Explore the Future of Virtual Reality with Oculus](#) »

¹⁹ Video *Google ATAP* « [Welcome to project Soli](#) »

What does it measure?

Muse: the brain sensing headband is a brain fitness tool that **measures brain signals** much like a heart rate monitor senses your heartbeat. Muse's 7 finely calibrated sensors – 2 on the forehead, 2 behind the ears plus 3 reference sensors – detect and measure the activity of your brain.





Capture d'écran site web Emotiv, 06/2015

1 CAPTATION DES ONDES CÉRÉBRALES MUSE ET EPOC+ D'EMOTIV

Les casques d'analyse des ondes cérébrales (casque dit EEG, pour électroencéphalographie) commencent à sortir des hôpitaux et laboratoires pour se diriger vers le grand public. *Muse* commercialise ainsi un casque (voir illustration) qui mesure des signaux du cerveau afin de favoriser la relaxation, le tout pour environ 300\$. Le projet français *MeloMind* surfe sur une proposition similaire. *Emotiv* produit des casques plus futuristes d'aspect, mais son modèle EPOC+ est pensé pour servir d'interface pour des jeux vidéo, notamment en détectant les états émotionnels (l'excitation, le calme...). Selon Randy Breen d'*Emotiv* cité dans la [Technology Review](#), ces indices inconscients peuvent être utilisés pour modifier l'ambiance musicale d'un jeu ou pour changer la manière dont les autres personnages interagissent avec le joueur¹⁵.

2 ENVIRONNEMENTS IMMERSIFS OCULUS RIFT, STEAMVR, HOLOLENS, MAGIC LEAP

Oculus Rift est un projet de casque de réalité virtuelle qui, après avoir fait les beaux jours de *Kickstarter*, a été acheté par *Facebook*. Ce rachat pose beaucoup de questions, concernant la stratégie future de la société, mais l'objectif couramment réaffirmé par son créateur est la sortie en 2016 d'une version grand public¹⁶ du casque de réalité virtuelle. Si pour le moment ce casque n'est pas centré sur la captation de données, il contient déjà beaucoup de capteurs, en particulier

pour mesurer la position et les mouvements de la tête. Ses développeurs insistent sur l'importance du développement des accessoires permettant les *inputs*, comme ceux permettant le suivi de mouvements de la main. Le fondateur d'*Oculus*, Palmer Luckey, affirmait ainsi [lors du festival SXSW en mars 2015](#) que, même sans parler d'une échelle de temps où nous aurons des implants neuronaux et de parfaits dispositifs haptiques¹⁷, il considère comme accessible à court terme « le suivi des mouvements basiques des doigts et des interactions basiques avec des doigts »¹⁸. Google a d'ailleurs dévoilé récemment un projet de recherche avancé dénommé *Soli*¹⁹ concernant la captation par radar (un radar sur une puce électronique) de mouvements des doigts très fins et précis afin de créer des interactions hommes-machines d'un nouveau genre. *Oculus* n'est pas le seul projet en cours : Sony a présenté son projet *Morpheus* dès 2014 et le constructeur de terminaux *HTC* s'est associé avec Valve pour le projet *Viver*, qui intègre gyroscope, accéléromètre et capteurs lasers dans un casque pour suivre les mouvements de la tête et repérant par laser les mouvements des personnes dans la pièce. Microsoft explore quant à lui la réalité augmentée (donc par des matériels dans lesquels l'immersion n'est pas complète) avec son projet *HoloLens*. Enfin, la très mystérieuse entreprise *Magic Leap* promet une technologie révolutionnaire de réalité augmentée, sans que personne n'ait vraiment pu voir ce que cette entreprise, dans laquelle Google a investi plus de 500 millions de dollars, propose réellement.

0.3



Partie 0.3

LE GRAAL DE LA RECOMMANDATION ET DE LA PERSONNALISATION

POURQUOI RECOMMANDER EST-IL SI IMPORTANT ?	50
COMMENT RECOMMANDE-T-ON ?	51
MYTHES & RÉALITÉS DE LA RECOMMANDATION	52
DE LA RECOMMANDATION À LA PERSONNALISATION : QUELLE EXPÉRIENCE POUR L'UTILISATEUR ?	56

LE GRAAL DE LA RECOMMANDATION ET DE LA PERSONNALISATION

“ J’étais chez *eBay* la semaine dernière, et ils m’ont dit que 90% de ce que les gens achètent chez eux vient des recherches. Nous sommes le modèle opposé. La recommandation est énorme, et notre option de recherche est ce que les gens utilisent quand nous ne sommes pas capables de leur montrer quoi regarder. ”

Xavier Amatriain
Netflix, *TechRadar*, 2014

- ▶ Recommander sert à attirer de nouveaux clients, à les fidéliser et à optimiser les revenus. C’est un impératif pour tous les acteurs économiques.
- ▶ Techniquement, la recommandation peut combiner des outils et approches très différentes, ce qui la rend d’autant moins transparente et compréhensible pour l’utilisateur.
- ▶ Cela peut nuire paradoxalement à son expérience d’utilisateur, et parfois... certaines frictions sont désirables pour recréer de l’humain, du contact.
- ▶ De ce point de vue, être explicite sur l’usage fait des données concernant la personne peut paraître en contradiction avec l’objectif d’invisibilité de la technologie dans l’expérience. En réalité, ces « imperfections » sont sources de confiance et sont des opportunités d’innover et de se différencier de la concurrence.

¹ L’*up selling* ou montée en gamme est la pratique qui consiste à proposer un produit ou service légèrement supérieur et plus cher que celui auquel s’intéresse le prospect, par exemple en améliorant l’exposition de contenus complémentaires, des offres complémentaires, des services « premiums »

² SMITH, Chris. “How Spotify, Netflix and Amazon control your online habits”, *TechRadar*, janvier 2014

³ KEMBELLEC, Gérald, CHARTRON, Ghislaine et SALEH, Imad (sous la direction de). *Les moteurs et systèmes de recommandations*, ISTE éditions, 2014, pp. 17.

POURQUOI RECOMMANDER EST-IL SI IMPORTANT ?

Pourquoi recommander ?

- **Attractivité.** Recruter de nouveaux clients et accroître le taux de conversion des visites en achats, en se différenciant de la concurrence à travers une expérience optimisée,
- **Fidélité.** Garder ses abonnés (en développant l’usage et la satisfaction) et donc diminuer le « churn » (ou taux d’attrition),
- **Optimisation.** Augmenter les revenus par abonné, stratégie d’*upselling*¹.

Face au catalogue, recherche ou recommandation ?

Face à un catalogue vaste, un service en ligne propose généralement deux manières de trouver des produits intéressants, soit en offrant un espace de recherche, soit à travers des recommandations. La recommandation consiste à faciliter l’appariement de l’offre et de la demande sur des marchés caractérisés par de fortes asymétries d’informations.

La plupart des offreurs utilisent les deux outils pour accompagner l’utilisateur. Ceci se matérialise notamment sur le site *Amazon* puisque la barre de recherche est centrale et permet de rechercher toute sorte de produits et le reste de la page est largement consacrée à des recommandations, personnalisées ou non.

La recommandation permet de concilier la taille du catalogue avec la facilité d’accès à des contenus, comme l’expliquait Donovan Sung de *Spotify* à *TechRadar* : « ... il y a une frontière fine entre offrir trop de choix et en offrir trop peu dans cette société de surcharge informationnelle. Les utilisateurs ne veulent pas 100 propositions, ils préféreraient une ou quatre bonnes propositions »². Les travaux de recherche en marketing montrent en effet que plus le choix est vaste, plus le consommateur devient hésitant : recommander est une manière de résoudre les problèmes de surcharge cognitive et de surabondance des contenus, et donc de lutter contre la pénurie d’attention. Les utilisateurs ont « besoin d’aide pour trouver le choix le plus approprié parmi de multiples produits ou informations disponibles, en leur fournissant des recommandations issues du filtrage des alternatives possibles »³.

Un impératif économique et marketing pour les services sur abonnement et une réponse à l’incertitude sur les marchés de la création

La réduction du « *churn* » est le combat le plus important pour un service sur abonnement. Pour que les personnes restent abonnées, il est nécessaire qu’elles ressentent de la satisfaction à utiliser le service, ce qui peut passer par des recommandations donnant à l’utilisateur l’impression qu’il y a une valeur ajoutée à rester chez son fournisseur de contenus.

Recommandation et personnalisation sont donc devenues des impératifs marketings du fait qu’elles sont des moyens de valoriser son service par rapport à celui de ses concurrents et donnent un pouvoir de marché importants aux acteurs qui investissent le rôle de prescripteur.

DES TRACES DE LA *LONG TAIL*

La *long tail* ou longue traîne en français est un concept mis en évidence par Chris Anderson dans [un article de Wired en 2004](#)⁷ selon lequel le succès d'acteurs comme Amazon proviendrait de la vente d'une grande diversité de produits en petite quantité, plutôt que de la vente massive de produits stars.

L'hypothèse sous-jacente est que la numérisation du commerce, en s'affranchissant en partie des contraintes physiques de distribution (surface des espaces de vente, coûts de stockage, de transport,...), permet un allongement de la traîne, c'est-à-dire une augmentation du nombre de références accessibles aux consommateurs.

Appliquée aux industries culturelles, cela signifierait que la numérisation de la consommation devrait être favorable à des œuvres de niche. Or, sur les marchés culturels, le numérique semble au contraire accentuer le succès des popstars⁸ pour la musique, des best-sellers pour le livre⁹ et des blockbusters dans le cinéma. Le numérique produirait en réalité des consommateurs encore plus mimétiques.

La recommandation est un élément essentiel pour que l'effet de longue traîne joue à plein. Comme le décrit Lionel Maurel¹⁰, il est nécessaire que les contenus les plus populaires soient d'une manière ou d'une autre (y compris par de l'intermédiation sociale) reliés à ceux qui le sont moins. Il appuie notamment son raisonnement sur une étude¹¹ mettant en évidence les conséquences de la fermeture de *MegaUpload* sur la diversité des œuvres consommées. Indépendamment du caractère illégal de la plateforme, un système de recommandation sociale (réseaux P2P, newsgroup) s'était constitué autour du site, favorisant ainsi l'accès à des contenus de niche. Il en déduit qu'à l'inverse, les recommandations des grandes plateformes seraient par nature biaisées pour favoriser les contenus ayant nécessité le plus d'investissements – en supprimant le risque lié aux aléas de la rencontre avec les goûts du public. L'effet longue traîne serait donc en partie contraint par les stratégies des plateformes numériques de distribution. Ce rôle pivot des prescripteurs interroge nécessairement sur leur neutralité en matière de recommandation.

L'importance économique de la recommandation tient aux caractéristiques singulières des marchés de la création :

■ L'incertitude

Il est par nature difficile d'anticiper la demande pour un contenu donné et donc d'en prédire le succès. À la différence d'autres marchés, cette incertitude est partagée par l'ensemble des acteurs des industries culturelles puisque ni le producteur, ni le consommateur ne savent vraiment définir la qualité d'une œuvre ; l'ignorance est symétrique⁴. En ce sens, le produit culturel est un bien d'expérience particulier car son appréciation, même après « consommation », reste subjective (voir partie I-2). Il n'est également pas possible de porter réclamation pour un morceau de musique qui décevrait, à la différence de certains biens plus classiques qui ne répondraient pas aux attentes. La recommandation a donc vocation à réduire les risques liés à l'aléa de production.

■ Le mimétisme et la contagion

De manière moins intuitive, un bien culturel a davantage de probabilité d'être adopté lorsqu'il l'a déjà été par d'autres individus. Sa consommation est ainsi caractérisée par des externalités de réseaux⁵ qui découlent de comportements sociologiques. Ils sont le résultat

de phénomènes de mimétisme et de contagion sociale, traduisant une stratégie des individus pour minimiser le niveau d'incertitude sur la qualité des biens qu'ils consomment. La recommandation peut ainsi amorcer la contagion.

■ Le cycle de vie des biens culturels

La phase d'amorçage est d'autant plus cruciale que la période durant laquelle se réalise l'essentiel des ventes d'un bien culturel est généralement courte⁶. La recommandation peut donc ralentir le cycle de vie, en allongeant la qualité et la durée d'exposition (voir encadré ci-dessus).

COMMENT RECOMMANDE-T-ON ?

Il existe différentes manières de proposer de nouveaux contenus à partir d'informations concernant un catalogue.

Il ne faut pas négliger la plus simple des méthodes : recommander ce qui vous arrange en tant que vendeur, ou recommander ce qui est programmé pour marcher (*best-sellers*, *blockbusters*, etc.). Dans le cas d'un vendeur, ce type de recommandation « basique » est d'ailleurs souvent très efficace, surtout s'il dispose de peu d'information sur les goûts du visiteur : proposer le dernier album à la mode ou le dernier film à gros budget d'Hollywood est un bon moyen >>>

⁴ PELTIER, Stéphanie.

« [L'économie de la culture, les industries culturelles : une exception économique](#) », *Cahiers Français* n°312, 2003

⁵ Effet que l'on retrouve traditionnellement pour les technologies de communication où la valeur augmente avec le nombre d'individus qui y participe : le meilleur exemple étant le téléphone ou de manière plus contemporaine les réseaux sociaux.

⁶ Cet effet joue aussi à plein dans l'univers des magasins d'applications mobile, où l'effet « hit parade » - cercle vertueux où les mieux classées sont les plus téléchargées – et où la phase de lancement sont cruciaux.

⁷ [The death of the long tail](#), *Midia Consulting*, 2014

⁸ 1% des artistes capterait 77% des revenus générés

⁹ Les achats se concentrent sur les contenus phares au détriment de la diversité selon [l'Institut GfK](#).

¹⁰ MAUREL, Lionel. « [Le contrecoup le plus négatif de la guerre au partage et son véritable objectif](#) », *S.I.Lex*, mars 2014

¹¹ PEUKERT, Christian, CLAUSSEN, Jörg et KRETSCHMER, Tobias. [Piracy and Movie Revenues: Evidence from Megaupload: A Tale of the Long Tail 2](#), août 2013

« L'HUMAIN VA INEXORABLEMENT DISPARAÎTRE DE LA RECOMMANDATION AU PROFIT D'ALGORITHMES TOUT-PUISSANTS »

Probablement pas. La recommandation est avant tout produite par des humains. Qui plus est, les recommandations sociales (“votre ami X a aimé tel contenu”) sont jugées comme assez efficaces par certains services. Surtout dans le domaine des contenus culturels, le rôle des experts reste crucial. Pour la vidéo, le groupe *Canal+* mise fortement, avec son nouveau moteur « *Suggest* », sur sa crédibilité au plan éditorial pour se différencier de *Netflix*, par l'adjonction du travail de ses experts éditoriaux à la puissance de l'algorithme. Dans le domaine de la musique, de nombreux services, dont le tout nouveau *Apple music*, proposent des contenus sélectionnés par des musiciens ou des journalistes. Selon *TechRadar*, l'application de streaming *Slacker Radio*, populaire aux États-Unis, avait indiqué que 84% de ses clients payants passaient plus de temps à écouter des playlists sélectionnées par des humains que celles générées algorithmiquement à partir d'une chanson ou d'un artiste.¹

« LES ALGOS SATISFONT LES CLIENTS ET SONT DES MÉTHODES DE PRESCRIPTION EFFICACES »

Difficile de répondre à cette question. Finalement, est-ce que la recommandation fonctionne ? Les utilisateurs en sont-ils réellement satisfaits ? Les utilisateurs sont souvent saisis par un sentiment de déception face aux propositions des algorithmes de recommandation que le marketing des plateformes vante pourtant à grands coups de superlatifs. Lors d'un entretien, le romancier Aurélien Bellanger a ainsi indiqué qu'il était toujours un peu déçu de voir qu'Amazon s'obstinait à lui recommander d'acheter le livre dont il était l'auteur²...

Fact Checking

La magie parfois bien décevante des algorithmes

« LES ALGOS RECOMMANDENT... CE QUI ARRANGE LE VENDEUR »

C'est possible et... cohérent avec les impératifs économiques. Le marché de la recommandation peut être analysé comme un exemple type de marché biface. Le prescripteur s'adresse en effet à deux segments constitués des utilisateurs et des producteurs de contenus, proposant du contenu aux premiers et des consommateurs aux seconds. Les intérêts des deux groupes ne sont donc pas naturellement alignés, les producteurs pouvant être soucieux avant tout de rentabiliser les contenus les plus coûteux. Quid de la neutralité de la recommandation dans ce contexte ? Se fait-elle dans l'intérêt du consommateur pour lui offrir le contenu le plus en phase avec ses goûts, ses attentes ou dans l'intérêt des producteurs de contenus pour favoriser un blockbuster ou prolonger le cycle de vie d'un bien culturel ? Le service ne favorise-t-il pas ses propres productions ou les contenus pour lesquels sa marge est la plus importante ?

Enfin, de manière plus subtile, les acteurs ne recommandent-ils pas des contenus... afin d'améliorer l'algorithme ? Une équipe de chercheurs du MIT de Boston a ainsi pu constater que certains produits d'un catalogue comme celui d'Amazon ou de Netflix embarquent une quantité disproportionnée d'informations à propos des goûts des utilisateurs. Si le fournisseur de service recommande ces contenus à tous ses utilisateurs, alors, sur la base des notations qui en résulteront, il sera bien plus en mesure de classer les utilisateurs dans des groupes de profils³. Les utilisateurs sont donc un rouage essentiel de tout mécanisme de recommandation. Sans leur travail dématérialisé et inconscient, le système perd en qualité (voir encadré sur le *digital labor* ci après).

« LES ALGOS RISQUENT D'APPAUVRIR LE CHAMP DE LA CRÉATION »

Peut-être... mais le pire n'est jamais certain ! Le risque est réel d'appauvrir la découverte culturelle en sacrifiant la sérendipité sur l'autel de l'efficacité algorithmique : les créations qui ne semblent pas *a priori* coller avec un profil ou ses goûts supposés auront moins de chance de lui être proposés. Cette tendance inquiète les créateurs et les autorités chargés de la promotion de la diversité culturelle (voir encadré chapitre I-2). *Spideo critique ainsi la spirale* pouvant conduire un système de recommandation à proposer en permanence les mêmes contenus...⁴ (voir encadré « 3 questions à Olivier Ertzscheid »). L'algorithme n'est pas non plus une entité immanente, mais le fruit d'un calcul basé sur des critères, comme tout autant de choix éditoriaux. Différents critères ou diverses sources d'analyse créeront des chemins différents pour un même utilisateur. La menace sur la diversité culturelle réside donc moins dans l'hégémonie des algorithmes que dans celle d'un acteur monopolistique. Enfin, dans la tête de certains entrepreneurs, artistes ou chercheurs, les algorithmes sont loin d'être les ennemis de la création, au contraire ! Ils y perçoivent l'émergence d'un nouveau mode de création, algorithmique.

« LES ALGOS RECOMMANDENT CE QUE TOUT LE MONDE AIME, ILS NE SONT PAS BONS POUR RECOMMANDER DES NOUVEAUTÉS OU FAIRE DÉCOUVRIR DES CHOSES INATTENDUES »

Parfois. Certaines méthodes de recommandation (par exemple le filtrage collaboratif) ont des effets puissants de rétroaction et vont donc renforcer la notoriété d'un contenu populaire au départ. Dans ce modèle la popularité engendre de la popularité, et finalement l'algorithme finit par évaluer l'ancienneté d'un « hit » plus que sa qualité réelle. L'effet pervers le plus connu concernant les outils automatisés de recommandation est celui dit du « départ à froid » : un nouveau contenu n'est pas évalué, n'est pas populaire, et le système dispose de peu d'informations sur lui, alors il ne sera donc pas recommandé. Dans ce cas l'algorithme ne produit que des prophéties auto-réalisatrices.

« LES ALGOS RENDENT LA PRESCRIPTION MOINS TRANSPARENTE »

Souvent. Si cette question s'est finalement toujours posée quant à l'indépendance des critiques, des jurys qui remettent des prix et même plus récemment avec les « blogueurs influents »⁵, elle rejaille sous un autre jour avec l'automatisation toujours plus grande de la recommandation.

Bien que la recommandation demeure toujours une proposition que l'utilisateur peut accepter ou non, contester ou admettre, les acteurs de la recommandation disposent d'un pouvoir important accentué par l'opacité des algorithmes utilisés. L'utilisateur ne saura jamais vraiment comment il est calculé, pas plus qu'il est peut-être enfermé dans des bulles de contenus... Demain, si les recommandations s'appuient sur des données captées par des objets connectés (voir partie IV), la transparence sur les raisons de la recommandation sera sans doute encore plus difficile à obtenir... C'est pourquoi certains ont envisagé de légiférer sur la transparence des algorithmes.⁶

¹ SMITH, Chris. "How Spotify, Netflix and Amazon control your online habits", *TechRadar*, janvier 2014

² Entretien avec l'équipe de rédaction, novembre 2014.

³ Voir le communiqué de presse sur le site du MIT, novembre 2014.

⁴ SPIDEO, *Beat Netflix at its own game*, mai 2015

⁵ Ces différentes interfaces ayant vocation à jouer un rôle de prescription, à réduire l'asymétrie d'informations.

⁶ voir par exemple CONSEIL D'ETAT. *Le numérique et les droits fondamentaux*, septembre 2014.

d'obtenir rapidement des conversions de visites en achats.

Dans le cas des services sur abonnement, la stratégie doit en revanche être plus sophistiquée pour fonctionner dans la durée. Cela n'exclut pas qu'elle soit également soumise avant tout aux intérêts du vendeur, au contraire : un tel service peut avoir plus de latitude pour influencer son utilisateur afin de rendre la recommandation plus orientée vers les besoins du service (par exemple en favorisant ses propres contenus au détriment de contenus produits par les concurrents).

Sur le plan technique, les plateformes mobilisent des algorithmes dont les logiques sont variées. Les approches de recommandation peuvent être basées sur :

- une analyse d'experts (la curation par des équipes de critiques par exemple) ;
- l'analyse automatique des contenus (donc liée à des questions de similarités), qui s'appuie surtout sur des métadonnées ;
- l'analyse de l'utilisateur (mots-clés, historique d'achat...) et de son réseau social (recommander ce que les amis aiment) ;

■ le « filtrage collaboratif », approche qui enrôle les individus dans le travail de création de la recommandation¹² : si Alice aime les contenus *a*, *b*, *c* et *d*, et que Benoit aime *a*, *b*, et *c*, il est cohérent de recommander *d* à ce dernier. Ces systèmes utilisent « le volume d'évaluations communautaires pour proposer des prescriptions culturelles personnalisées, basées sur la statistique d'évaluation et la corrélation de profils d'utilisateurs »¹³.

Souvent, les spécialistes considèrent que le filtrage collaboratif est à la fois la méthode la plus traditionnelle et la plus robuste lorsqu'elle n'est pas associée à d'autres procédés¹⁴.

Mais les approches les plus modernes sont hybrides et s'appuient sur l'apprentissage automatique (*machine learning* en anglais). Sont particulièrement prisées les stratégies dites de

« learning to rank », centrales pour les moteurs de recherche et la publicité ciblée. Les critères de choix pour ordonner les résultats sont alors déterminés automatiquement, et non par un choix *ex ante* (par exemple entre la popularité et la prédiction liée aux goûts de la personne).

Ainsi, lorsque le journaliste du *Guardian* Stuart Drudge a demandé à *Spotify* d'explicitier son système de recommandation, il lui a été répondu que le processus se découpait en 3 parties :

- d'abord, une analyse des contenus écoutés (style, tempo), soit une approche plutôt basée sur les contenus ;
- puis un regroupement de genres musicaux les plus appréciés par comparaison à ceux des autres utilisateurs, soit une approche de filtrage collaboratif ;
- enfin, l'utilisateur en lui-même est catégorisé par rapport à ses comportements et son profil (type de consommation, fréquence d'écoute, écoute des artistes en profondeur ou en surface¹⁵).

La recette utilisée par *Spotify* serait donc en phase avec le discours de Jeffrey Dean¹⁶, ingénieur en charge du projet *Google Brain*, pour qui il est nécessaire de disposer de 3 éléments essentiels pour produire des recommandations efficaces :

- le contexte – comprendre ce que fait l'utilisateur, ce qu'il recherche, son environnement ;
- les comportements :
 - comprendre les comportements passés de l'utilisateur (au moyen d'un historique de recherches ou d'actions par exemple) ;
 - agréger les comportements passés d'un grand nombre d'utilisateurs (en particulier lorsqu'il n'y a pas encore d'historique exploitable concernant un utilisateur en particulier).
- la compréhension sémantique – au-delà de la simple compréhension textuelle par des mots-clés, comprendre les sentiments exprimés par l'utilisateur.

¹² Ces algorithmes sont des exemples très clairs de « *digital labor* », c'est-à-dire des activités numériques qui peuvent s'assimiler au travail parce que productrices de valeur alors qu'elles ne sont pas marquées par des relations de travail traditionnelles. Sur ce sujet, voir :

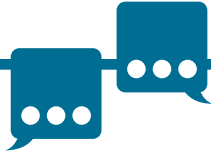
CARDON, Dominique et CASILLI, Antonio. [Qu'est ce que le digital labor ?](#), 2015 et encadré ci-après

¹³ KEMBELLEC, Gérald, CHARTRON, Ghislaine et SALEH, Imad (sous la direction de). *Les moteurs et systèmes de recommandations*, ISTE éditions, 2014, pp. 23-24.

¹⁴ KEMBELLEC, Gérald, CHARTRON, Ghislaine et SALEH, Imad (sous la direction de). *Les moteurs et systèmes de recommandations*, ISTE éditions, 2014, 227 p.

¹⁵ GUILLAUD, Hubert, « [Spotify : comment fonctionne la recommandation musicale ?](#) », *Internet actu*, février 2015

¹⁶ [Keynote](#) sur *Large Scale Machine Learning for Predictive Tasks*.



3 QUESTIONS À... OLIVIER ERTZSCHEID,

MAÎTRE DE CONFÉRENCES EN SCIENCES DE L'INFORMATION
À L'UNIVERSITÉ DE NANTES, LABORATOIRE DICEN-IDF (UNIV.
PARIS OUEST) - BLOG: [HTTP://AFFORDANCE.TYPEPAD.COM/](http://affordance.typepad.com/)



© André Gunthert

Du point de vue de la diversité culturelle, la recommandation est-elle vraiment un danger ?

La question clé de la recommandation, qu'il s'agisse de contenus culturels ou d'opinions politiques ou religieuses exprimées par nos amis sur les réseaux sociaux, est celle de la bulle de filtre d'Eli Pariser : les algorithmes favorisent-ils la diversité ou au contraire les phénomènes d'homophilie, c'est à dire le renforcement de nos propres goûts, croyances, opinions et affinités ? Toutes les études scientifiques sur le sujet, depuis celle de [Fleder et al. en 2007*](#) jusqu'à celle de *Facebook* en Mai 2015 montrent que 2 niveaux d'interprétation se superposent : on observe, à l'échelle de l'individu, une augmentation de la diversité des produits culturels et/ou opinions auxquels il est exposé ; mais cette diversité diminue

dès que l'on passe à un niveau plus global. Il y a donc, au niveau micro, une augmentation de la diversité, et une diminution de cette même diversité au niveau macro.

Comment les systèmes de recommandation ont-ils évolué ?

Nous sommes sortis d'une simple recommandation statistique (basée sur l'historique de nos achats ou de nos consultations) pour aller vers des logiques de plus en plus fines mais aussi de plus en plus contraintes de prescription capables de s'affranchir de toute historicité pour mettre en place une approche multifactorielle. D'abord l'algorithme a progressé et la simple analyse statistique intègre désormais une part de stochastique, d'aléatoire, qui permet au consommateur de ne pas voir défiler éternellement les mêmes catégories de produits. Ensuite nos comportements et nos usages se sont également diversifiés, à la fois du point de vue du matériel (smartphones, tablettes), du temps et de la vitesse de l'accès, mais aussi du point de vue des usages : plus nous consommons de contenus en ligne au travers par exemple d'offres d'abonnement ou de streaming et plus il est facile de qualifier ces comportements et de produire des recommandations plus fines. Enfin la collecte massive de données personnelles permet de faire varier le cycle et les logiques de prescription en fonction d'un nombre quasi-illimité

de critères : heure de connexion, historique des sites visités, panier moyen d'un consommateur, goûts littéraires ou musicaux, composition du foyer (nombre d'enfants, âge, sexe, etc.),

Ces acteurs économiques sont donc de vrais industriels de la recommandation ?

C'est cette alliance de la quantification statistique et de la qualification documentaire qui fonde la puissance des industries de la recommandation. Le cycle de la prescription algorithmique, beaucoup plus renseigné, plus complet et plus paramétrable que celui de la prescription classique, fonctionne par percolation grâce aux régies publicitaires partagées : le produit que vous avez regardé sur le site de la FNAC hier s'affichera pendant une semaine en lien sponsorisé sur vos sessions *Facebook*, vous enverra des notifications sur votre smartphone, etc. Le problème est qu'il est pour ainsi dire impossible de reconstruire *a posteriori* ce processus de prescription algorithmique : la prescription n'étant pas incarnée, le consommateur n'a en retour aucune possibilité de la recontextualiser et se contente trop souvent de la subir. La traçabilité fonctionne à sens unique.

* FLEDER, Daniel et HOSANAGAR, Kartik. « Blockbuster Culture's Next Rise or Fall : The impact of recommender systems on sales diversity », *Management Science*, septembre 2007

DE LA RECOMMANDATION À LA PERSONNALISATION : QUELLE EXPÉRIENCE POUR L'UTILISATEUR ?

Individualisation des flux et des contenus

Recommandation sociale ou algorithmique, mise en avant de contenus similaires, populaires ou adaptés à votre humeur ou à l'heure de la journée... nos environnements numériques se façonnent jour après jour à notre image. Les interfaces permettent à l'utilisateur d'évoluer dans un environnement balisé et les services collectant les données de leurs utilisateurs au fil de l'eau pour affiner leurs recommandations en profitent parfois pour leur tendre un miroir.

¹⁷ MALTER, Jordan. « [Spotify's vision for the future of music](#) », *CNN*, décembre 2013 (video).

Spotify a ainsi développé une fonctionnalité « *a Year in Music* » (« une année en musique »), qui montre une synthèse personnelle des écoutes de l'année écoulée et propose, sur cette base, une *playlist* pour l'année à venir.

La personnalisation de l'expérience utilisateur commence même à s'immiscer au cœur même des œuvres. Daniel Ek, fondateur de *Spotify*, [rêve qu'un même chanson](#) puisse « être différente pour chaque personne qui l'écoute ? »¹⁷. Tod Machover, compositeur, inventeur et professeur au MediaLab du MIT [imagine un futur de la musique dans lequel](#) « les compositeurs vont proposer des œuvres inachevées que chaque auditeur pourra compléter, personnaliser et finaliser en fonction de

Compliquées, les techniques de recommandations ?



sa psychologie ou de sa personnalité. L'auditeur ne sera plus dans une posture passive mais vivra ainsi une expérience totale et immersive¹⁸ (voir chapitre II-3).

Il en va de même pour la lecture. En fonction des préférences de lecture d'un utilisateur, ou du temps qu'il a à consacrer à la lecture (lecture dans les transports,...), le contenu proposé pourra s'adapter (voir chapitre II-2).

Aurélien Bellanger, romancier auteur de *la théorie de l'information*, nous expliquait définir une œuvre par son caractère «achevé». Pour lui, «... un label de qualité et de complétude est vendu avec un livre papier, [...] la littérature étant l'un des derniers endroits de contrôle total de l'œuvre.». Dans un contexte où les acteurs des industries culturo-ludiques cherchent à proposer à leurs clients une expérience globale et intime, l'œuvre n'existe plus uniquement par elle-même mais s'inscrit dans une mise en scène cohérente présentant un univers au plus proche de l'usager, familier, regroupant ses amis, ses goûts, des informations contextuelles (en fonction de l'heure de la journée, de sa localisation) et aussi les propensions ou préférences qu'il est susceptible de développer.

« Recherche frictions pour dialogue »

De la collecte des données à la consommation du service, en passant par le mécanisme de recommandation, les services «sans couture» finissent par se fondre avec l'utilisateur en fonctionnant comme une tâche de fond, omniprésente et ubiquitaire.

**Peut-on réellement accepter et comprendre le fonctionnement d'un service automatisé et invisible ?
Faut-il imaginer, prévoir des mécanismes de transparence qui fonctionneraient comme des points de contact avec l'utilisateur, des moments où le service rendu devient explicite ?**

Des possibilités de prise en main ponctuelles pourraient être proposées à l'utilisateur, avec

par exemple un bouton *off* pour les situations sensibles, la possibilité de repartir à zéro voire d'être anonyme. En bref, y'a-t-il des frictions désirables ?

Les services de recommandation les plus avancés tentent aujourd'hui d'être plus explicites sur les raisons de la recommandation, pour éviter les soupçons de manipulation et se rendre plus humains, plus compréhensibles. Sur [le blog technique de Netflix, à l'issue d'une longue explication sur la manière dont l'entreprise gère la recommandation](#), l'auteur explique qu'un élément majeur de la personnalisation réside dans la sensibilisation : « nous voulons que les membres soient au courant de la façon dont nous nous adaptons à leurs goûts. Cela permet non seulement de promouvoir la confiance dans le système, mais encourage les membres à donner un retour qui se traduira par de meilleures recommandations¹⁹. Une autre manière de promouvoir cette confiance en la composante de personnalisation est de fournir des explications quant à la raison pour laquelle nous décidons de recommander un film ou une émission donnée. Nous ne recommandons pas ce contenu pour nos intérêts commerciaux, mais parce que cela correspond aux informations que nous avons sur vous : vos préférences explicites, vos goûts, vos notes, votre historique ou même des recommandations de vos amis. »²⁰

L'opacité des modes de fonctionnement des différents services est en tout cas à l'origine de nombreux discours inquiets quant à la place laissée à l'individu dans cet environnement de plus en plus autodéterminé. Au-delà de la stricte problématique de la maîtrise par les individus de leur environnement numérique, l'article [la culture à la merci des algorithmes](#)²¹ résume les craintes de l'imposition d'œuvres aux spectateurs qui les enfermeraient dans des bulles (les fameuses [filter bubbles](#) d'Elie Pariser²²) et feraient que « l'amateur de polars scandinaves ne mettra jamais le nez dans un roman américain ». « À long terme, c'est extrêmement dangereux », analyse le philosophe Bernard Stiegler. « La représentation que se fait un individu de lui-même est construite sur une somme de contradictions et d'idéaux. Nous avons la possibilité de changer en permanence. Or ces algorithmes détruisent notre singularité en créant

¹⁸ « [Le monde de demain selon Tod Machover](#) », *SoonSoonSoon*, mars 2014.

¹⁹ Une nouvelle forme de *digital labour*

²⁰ AMATRIAIN, Xavier. « [Netflix Recommendations: Beyond the 5 stars](#) », *Netflix Tech Blog*, avril 2012.

²¹ BORDIER, Julien et HANSEN-LOVE, Igor. « [La culture à la merci des algorithmes ?](#) », *L'Express*, janvier 2015.

²² Pour plus d'information sur ce concept, voir la page *Wikipedia* « [bulle de filtres](#) »

un double de nous-mêmes qui n'est qu'une image statistique».

Selon *Spideo*, cet enjeu est crucial pour l'amélioration des systèmes de recommandation : « du point de vue de l'expérience utilisateur, l'amélioration des recommandations sera la conséquence d'une idée simple : les technologies doivent révéler le raisonnement derrière la recommandation intelligente – et ce en langage naturel, aisé à comprendre - afin de fournir des suggestions de confiance, toujours personnellement pertinentes et simples à comprendre »²³.

Les utilisateurs sont, de leur côté, en phase avec cette évolution : loin d'être opposés à la personnalisation, ils la souhaitent et l'acceptent même si leurs données sont utilisées à cette fin. En bref, la finalité de personnalisation et de recommandation est légitime à leurs yeux.

Les chercheurs Ting Li et Till Unger ont essayé de déterminer si l'absence de confidentialité, corollaire d'un service de recommandation de qualité, avait un impact négatif significatif sur son adoption par la clientèle. Les résultats montrent que des utilisateurs sont non seulement prêts à divulguer des informations personnelles lors de l'utilisation de services qu'ils jugent bons qualitativement, mais qu'ils sont également susceptibles de payer pour un tel service²⁴.

Loi d'être étonnante, cette étude montre que les individus ne se résignent pas à un monde numérique où le seul choix qui s'offre à eux est entre le renoncement à des services et l'usage sans limite et caché des données pour payer un accès à des services gratuits (« si c'est gratuit, c'est vous le produit »).

La réalité est que les utilisateurs souhaitent des services à forte valeur ajoutée, éventuellement payants, qui utilisent des données les concernant de manière certes intensive mais responsable, transparente, à leur profit et sous leur contrôle.

²³ SPIDEO. *Beat Netflix at its own game*, mai 2015

²⁴ LI, Ting et UNGER, Till. « Willing to pay for quality personalization? Trade-off between quality and privacy » *European Journal of Information Systems*, p. 1-22, mars 2012



3 QUESTIONS À... DOMINIQUE CARDON ET ANTONIO CASILLI

SOCIOLOGUES ET AUTEURS
DE *QU'EST CE QUE LE DIGITAL
LABOR ?*

INA ÉDITIONS, COLLECTION : ÉTUDES ET
CONTROVERSES, JUILLET 2015

Antonio Casilli, comment qualifiez-vous le lien entre digital labor et données personnelles ?

Antonio Casilli : Le *digital labor* en lui-même renvoie à ces activités numériques quotidiennes assimilables au travail car productrices de valeur, faisant l'objet d'une forme de prescription et soumises à des métriques de performance. Plus précisément, il s'agit de regarder les contributions numériques quotidiennes des utilisateurs de service, en particulier celles à faible intensité et à faible expertise, mises à profit via des algorithmes et des fouilles de données. On a par trop insisté sur les contenus générés par les utilisateurs, mais les recherches les plus récentes sur les critiques, recommandations ou productions amateurs semblent indiquer que ces activités sont peu valorisées par les plateformes. C'est la donnée, voire la métadonnée, les clics, les mesures de connexion qui comptent. Avec le *digital labor*, l'accent porte sur l'extraction pure et simple de données de la masse des usagers, et pas seulement des plus spécialisés et plus qualifiés d'entre eux. Ce régime de production n'a pas besoin de présupposer une compétence ou une formation préalable. Être connecté suffit, à plus forte raison quand l'entreprise analyse les requêtes insérées dans un moteur de recherche, les consommations de contenus, ou une

synchronisation de dispositifs mobiles tracés en permanence, le tout venant nourrir des bases de données et calibrer des algorithmes. En tant que producteurs des données ensuite vendues aux annonceurs, les utilisateurs sont en même temps des marchandises et des travailleurs. En fait, le célèbre dicton « si c'est gratuit, c'est que tu es le produit » devrait plutôt s'énoncer « si c'est gratuit, c'est que tu y travailles ». Dans un contexte aussi asymétrique entre les utilisateurs et les plateformes de capture de données/métadonnées, ces dernières sont donc au cœur des problématiques d'exploitation de la « *data as labor* ».

Dans le domaine des industries créatives et culturelles, les systèmes de recommandation / personnalisation sont cruciaux. Y'a-t-il selon vous une dimension de *digital labor* particulière à ce domaine ?

Antonio Casilli : Certes malgré l'effort des plateformes pour restreindre leur autonomie et les forcer à une production standardisée de contenus, les publics numériques ne se laissent jamais complètement réduire à la condition de simples récepteurs passifs. Ils aspirent à être des contributeurs actifs, qui maîtrisent les résultats de leurs activités, les conditions d'utilisation (lesquelles sont en fait des conditions de travail). Il y aurait donc un *digital audience labor*, un travail du public des médias numériques par l'ensemble d'activités de consommation et en même temps de production de contenus des industries innovantes de la culture et de la communication. Même les intelligences artificielles les plus avancées d'aujourd'hui ne sont pas vraiment adaptées, par exemple, pour créer des *playlists* à partir d'un ensemble de morceaux de musique ou pour donner une note à un film.

Finalement, pour les publics, une conscience exacte du moment où le

dispositif numérique est en train de commander leur travail, ou de leur accorder un moment d'oisiveté est impossible à entretenir. Les utilisateurs des plateformes n'arrivent pas à distinguer un moment où ils seraient simplement en train de profiter de services extrêmement performants (et gratuits), d'un autre dans lequel ils seraient en train d'ajouter eux-mêmes des données et des calibrages aux services en question, pour participer à leur optimisation et permettre à l'entreprise d'extraire de la plus-value. Cette prise de conscience d'un loisir qui s'avère en fait un travail « implicite » correspond en fait à des explosions de conflictualité sur les plateformes, aux révoltes des utilisateurs organisés qui émaillent l'histoire des géants d'internet, de AOL à YouTube, ou Facebook.

Dominique Cardon, vous êtes plus réservé quant à cette idée de *digital labor* en particulier par rapport à sa réalité empirique. Quel équilibre existe selon vous entre la valeur créée pour les individus et celle créée pour les plateformes, au travers en particulier des algorithmes de recommandation / personnalisation ?

Dominique Cardon : À ne pas prendre au sérieux l'importance cruciale prise aujourd'hui par les procédures d'intermédiation (hébergement, agrégation, classements, algorithmes, etc.), on risque toujours de penser qu'individuellement, les productions des internautes auraient de la valeur. Il suffirait alors de s'émanciper des plateformes, ou d'en créer de nouvelles d'un claquement de doigts, pour se libérer de la captation indue de plus-value par les GAFA (Google-Apple-Facebook-Amazon). Sans effet réseau, sans artefact de mise en relation, sans algorithme de recommandation, sans métrique de visibilité, nos productions numériques ne seraient plus du travail extorqué, mais elles auraient aussi perdu leur visibilité,

leur réputation, les gratifications de l'échange et les honneurs du commentaire, bref, elles seraient sans aucune valeur, ni monétaire ni symbolique. Les productions individuelles n'ont de valeur qu'à travers la base de données qui les associe à une masse considérable d'autres productions individuelles. S'articule ainsi un travail individuel et l'émergence d'un bien commun, une utilité produite par l'agrégation des activités individuelles.

Le déplacement du débat intellectuel de la notion d'« intelligence collective » vers celle de « gouvernance algorithmique », a contribué à défaire le lien ténu et fragile qui faisait du web des communautés « des origines » un espace dans lequel les contributions individuelles entretenaient un lien avec des formes collectives. Désormais, les parties peuvent se sentir autorisées à demander des comptes individuels, parce que le tout n'est plus que l'expression des intérêts économiques des firmes qui façonnent des algorithmes trompeurs.

D'ailleurs, concernant la personnalisation des profils, il se joue me semble-t-il un paradoxe du même ordre que le fameux « paradoxe de la vie privée ». En interview, les internautes y sont très hostiles ; en pratique, ils la réclament. Le débat sur la personnalisation est structuré par la question de l'information politique (ne pas voir l'information « de droite » quand on est « de gauche », et inversement ; ne pas voir ce qui se passe en Syrie, quand nos amis ne s'intéressent qu'à Justin Bieber, etc.) Il est évident que cela pose problème et ce problème est identifié par les utilisateurs dans les interviews comme une menace sur leur libre accès à l'information. En revanche, tous les autres mécanismes de personnalisation, c'est-à-dire en fait le principal, porte sur des questions de goûts et de centre d'intérêt. Et là, les fans de John Cage n'ont pas envie de voir Beyoncé apparaître dans leurs recommandations...

POUR SORTIR DE L'OPPOSITION ENTRE INNOVATION ET PROTECTION DES DROITS, INTÉGRONS ÉTHIQUE ET CONFIANCE À L'EXPÉRIENCE UTILISATEUR

Une fois la logique de valeur d'usage des données prise en compte, il est enfin possible de sortir de l'opposition stérile entre innovation et protection des droits. Les grands principes de protection des données personnelles ne sont alors plus seulement des contraintes, mais des éléments fondateurs voir refondateurs de la relation de confiance entre les utilisateurs et ces services. Cette position n'est pas seulement celle de la CNIL, elle est aussi très présente dans l'étude annuelle 2014 du Conseil d'État « [le numérique et les droits fondamentaux](#) »²⁵ et dans le rapport « [Ambition numérique](#) » du Conseil national du numérique²⁶ : il faut assurer concrètement le droit à l'autodétermination informationnelle des individus et la loyauté des plateformes et de leurs outils. La polémique qui a eu lieu dans certains médias et sur les réseaux sociaux en août 2015 à propos d'une nouvelle politique de confidentialité de Spotify montre que toute incompréhension sur ces questions peut ouvrir une crise entre les services et certains utilisateurs. Spotify a été obligé de réagir en quelque jour pour expliciter ses intentions, alors que celles-ci devraient être claires.²⁷

Dans le contexte de services personnalisés d'accès à des contenus créatifs et culturels, quatre axes semblent, de ce point de vue, prioritaires :

► INFORMATION ET CONSENTEMENT DES UTILISATEURS

Les services doivent dépasser la vision traditionnelle des conditions générales d'utilisation et du recueil d'un consentement peu libre et peu éclairé par l'adhésion à un contrat qui n'est jamais vraiment lu. Par exemple, il faut encourager le recours à des écrans d'installation, des écrans d'autorisation d'accès, des notifications contextuelles... qui ne sont pas pensées pour induire l'utilisateur en erreur ou pour une application formelle et minimale du droit en vigueur, mais bien pour lui expliquer réellement ce qui sera fait et utilisé. Une autre voie féconde serait d'explorer la fourniture de tableaux de bord indiquant quelles données réellement utilisées.

► DROIT À LA PORTABILITÉ

Corollaire à ce besoin d'explication et d'information intelligente des utilisateurs, les données, dont on a pu voir le rôle économique, ne doivent pas devenir un trésor de guerre capté par les entreprises et utilisé par celles-ci pour empêcher le client de changer de fournisseur de services. Les données doivent être récupérables et réutilisables par l'individu, dans des formats standardisés. Ce droit à la portabilité, qui est inscrit dans le projet de règlement européen sur la protection des données, est aussi une garantie contre le fait que les données personnelles deviennent une restriction de la concurrence au profit de ceux qui les ont monopolisées, au détriment des individus et des autres acteurs du marché. Il favorise

l'émergence de nouveaux acteurs tout en donnant plus de contrôle aux utilisateurs.

► TRANSPARENCE DES ALGORITHMES

Aujourd'hui, une entreprise doit être transparente sur ses prix, ses processus industriels, ses pratiques de ressources humaines, ses finalités, les données qu'elle collecte ou son impact environnemental. Demain, elle pourrait devoir être transparente sur ses algorithmes, sur leurs règles et leurs réglages. Aussi, une certaine transparence des algorithmes de recommandation est souhaitable. Dans le domaine culturel, il s'agit avant tout de garantir la neutralité des plateformes vis-à-vis de la diversité culturelle et d'éviter que l'utilisateur ne se trouve face à des recommandations incompréhensibles et obscures. À cet égard, ce secteur est peut-être en avance sur les autres, qui auront eux aussi, à l'avenir, intérêt à penser leurs outils de prescriptions et de recommandations pour qu'ils ne soient pas vécus comme des augures incontestables auxquels l'utilisateur doit se soumettre...

► LOYAUTÉ

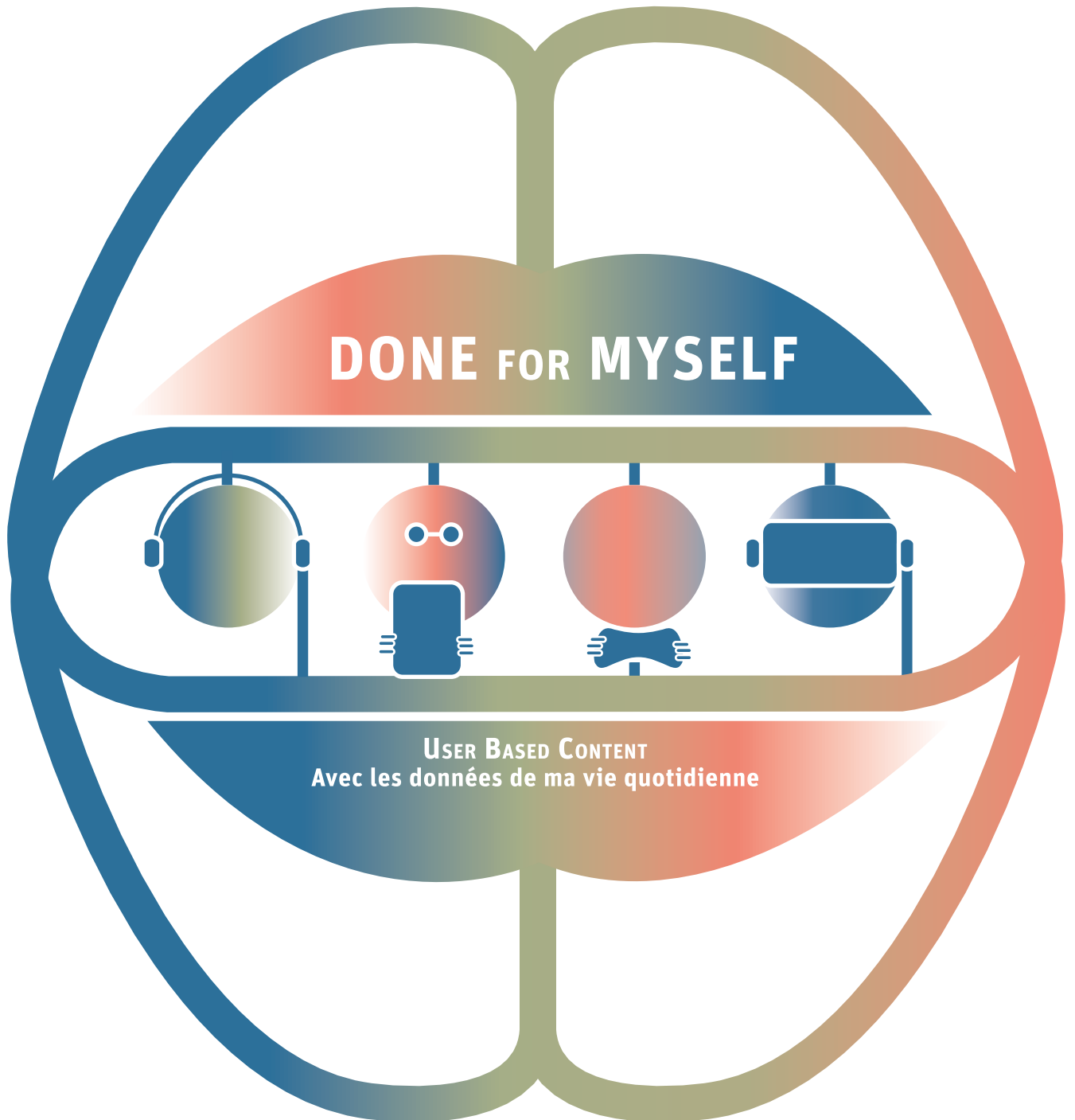
Enfin, la loyauté doit porter sur la finalité quant à l'utilisation des données qui concernent les individus. En effet, les données issues de la consommation de contenus culturels et ludiques peuvent s'avérer particulièrement sensibles en ce qu'elles révèlent par elles-mêmes sur la personne, sur ses goûts, ses intérêts, mais aussi en ce qu'elles peuvent permettre d'inférer sur son mode de vie. L'analyse d'un historique de consommation *Spotify* ou *Netflix* peut par exemple facilement permettre de déterminer à quel moment la composition d'un foyer a évolué avec l'arrivée d'un enfant (i.e : modification de la typologie des contenus consommés intégrant davantage de comptines et de dessins animés). Le caractère intime de ces données tient aussi au fait que généralement les individus écoutent, lisent, regardent et jouent à des choses qui leur plaisent et leur correspondent. Cette lapalissade a vocation à souligner qu'aujourd'hui peu d'individus mettent en place des stratégies d'obfuscation dans la consommation de ces contenus, contrairement à d'autres domaines où il existe des comportements plus stratégiques²⁸, en particulier sur les réseaux sociaux où les individus mettent en avant ce qui les valorise et où ils cachent ce qui les dérange. Les données de consommation en contenus culturels renvoient également à l'identité et la personnalité des individus (voir partie II-1 qui évoque la loi VPPA aux États-Unis). Pour ces raisons et pour préserver la confiance entre les utilisateurs et les plateformes, il est essentiel que ces données ne fassent pas l'objet de détournements en usages marketing ou publicitaires cachés mais qu'elles servent bien à la personnalisation du service dans l'intérêt de l'ensemble des parties et dans le respect de la loi Informatique et Libertés.

²⁵ CONSEIL D'ÉTAT. [Le numérique et les droits fondamentaux](#), septembre 2014

²⁶ CONSEIL NATIONAL DU NUMÉRIQUE, [Ambition numérique](#), juin 2015.

²⁷ Sur cette polémique, voir par exemple : BOLESSE, Cécile. "[Après la polémique, Spotify éclaire ses conditions d'utilisation](#)" O1net, septembre 2015.

²⁸ Cf. le chapitre « La révolution du web social : demain tous des people ? » dans CNIL. [Vie privée à l'horizon 2020](#), *Cahier IP n°1*, novembre 2012.



Pourquoi cette illustration ? Il fut un temps où le "Do It Yourself" était de mise, faire soi-même pour soi-même. Pour aujourd'hui et demain j'emploie l'expression "Done For Myself", l'homme étant considéré comme passif, tandis que ses faits et gestes au quotidien participent à une élaboration sur mesure qui lui serait directement destinée. Services ingénieux ou cercles vicieux ?

Mylène Podvin, Designer/Illustratrice du cahier



Réalisation : Mylène Podvin/CNIL



Partie 0.4

DEMAIN, QUELLES CRÉATIONS ET QUELS USAGES DATA-DRIVEN ?

QUELS SCÉNARIOS D'USAGE ET QUELS MODÈLES ÉCONOMIQUES À 5 ANS ?	64
HYPERMOVIE GENERATOR	66
LIFE TRACKS	68
TRAUMDATA	70
THE COLLECTIVE MOOD MASTER	72
DE QUOI CES SERVICES SONT-ILS LE NOM ?	74

Quels scénarios d'usage et quels modèles économiques à 5 ans ?

Un nouveau paysage lié à l'utilisation et l'exploitation des données est en train de prendre forme dans les marchés de contenus culturels et créatifs. La personnalisation de l'expérience y devient plus intime, l'intrication entre l'œuvre et l'expérience d'utilisation plus grande, la frontière entre création par l'auteur et contenus générés par et pour l'utilisateur plus floue. Certaines tendances lourdes, transversales aux différents secteurs, paraissent particulièrement robustes et dessinent les contours de ce qui pourrait advenir dans les prochaines 3 à 5 années.

SENSIBILITÉ ET ADAPTATION AU CONTEXTE (CONTEXT AWARENESS) + ALGORITHMIQUE = PRÉDICTION PLUS PERSONNALISÉE

Le service sait ce qui se passe et peut prédire et agir (recommander, adapter, etc) en fonction de ces éléments.

La prise en compte du contexte de manière automatique permet d'adapter d'abord la recommandation, puis le contenu aux utilisateurs (exemple : adapter le cadre d'un film au contexte culturel des spectateurs). L'objectif d'une application contextuelle est de délivrer un service à un utilisateur en minimisant les interactions entre celui-ci et la machine. Ainsi, l'application cherche à détecter automatiquement le contexte dans lequel il se trouve pour lui fournir un service adapté. Rand Hindi, fondateur de la startup SNIPS, espère que ces technologies permettront de «faire disparaître les technologies»¹ et réduire leur côté agaçant (notifications, demandes récurrentes d'inputs, etc). Pour lui, «proposer un contenu le plus pertinent possible a pour but de rendre la recommandation plus agréable, plus acceptable, moins intrusive»².

¹ HINDI, Rand. [How Artificial Intelligence Will Make Technology Disappear](#), juin 2015

² Entretien avec Rand Hindi, octobre 2014.

³ Le data mining est un processus qui permet d'extraire des informations pertinentes à partir d'une grande masse d'informations.

⁴ Le transmédia ou transmédia storytelling est la pratique qui consiste à développer un contenu narratif sur plusieurs médias en différenciant le contenu développé et les capacités d'interaction en fonction des spécificités de chaque média.

ANALYSE DES SENTIMENTS, DE L'HUMEUR ET DES ÉMOTIONS + OBJETS CONNECTÉS ET CAPTEURS = SERVICES AMBIANTS

Le service sait ce que l'utilisateur ressent et réagit (recommande, adapte) en fonction de ces analyses en temps réel.

Les technologies numériques se rapprochent du corps humain, de ses constantes, et de ses réactions, comme nous l'évoquions dans notre cahier IP2 «[Le corps, nouvel objet connecté ?](#)». Grâce aux objets qui se portent (*wearables*), grâce aux montres connectées et aux autres capteurs de la maison ou de la ville intelligente, les individus vont de plus en plus bouger, penser, agir et ressentir au cœur d'un réseau de capteurs embarqués dans leur environnement et capables d'analyse en temps *quasi* réel. En quelque sorte, nous pouvons faire l'hypothèse de l'extension du domaine du *data mining*³ aux émotions et ressentis des individus par l'intermédiaire des capteurs et objets connectés. L'émergence de nouveaux modèles de valorisation des données personnelles dans les industries culturelles repose ainsi en partie sur l'objet matériel, le *hardware*.

FICTIONS PLUS IMMERSIVES ET INTERACTIVES + TRANSMÉDIAS⁴ = ENGAGEMENT PLUS FORT ENTRE LES INDIVIDUS ET LES ŒUVRES

Il y a une histoire. Il y a plusieurs histoires. Tu la/les découvriras en fonction de ce que tu es, de ce que tu fais.

Les frontières entre les formes d'œuvres fictionnelles s'estompent, se transcendent, se renversent. L'immersion, les nouvelles interactions humains-machines, l'interactivité, sont autant de voies qui augmenteront l'intensité de notre rapport aux œuvres. De telles fictions, de telles œuvres, s'élaborent (marginale ou fondamentalement) au fur et à mesure des interactions entre l'humain et la machine, entre l'humain et l'œuvre.

Au sein de ces trois axes, les données concernant l'utilisateur jouent un rôle crucial. Elles rendent possible de nouvelles formes de narration et d'œuvre ; elles rendent aussi probablement viables de nouveaux modèles d'affaires. Elles sont d'ailleurs déjà devenues une brique centrale dans la stratégie des entreprises et les moteurs de la personnalisation des services, des recommandations et de la monétisation des profils des utilisateurs.

La technologie contribue quant à elle à développer de nouveaux usages autour d'une expérience individuelle forte où les données captées suivent un processus d'interprétation-compréhension : le contenu est-il suffisamment impactant pour engager le consommateur dans une attention durable ? Est-il possible de mieux comprendre les mécanismes d'attention en compilant des données sur les utilisateurs ?

UNE RÉFLEXION COLLABORATIVE BASÉE SUR DES MÉTHODOLOGIES ALLIANT LE *DESIGN THINKING* ET LES TECHNIQUES DE PROSPECTIVE, AVEC [WEDESIGNSERVICES](#) ET [FIVEBYFIVE](#).

Deux ateliers ont été organisés dans les locaux de la CNIL, avec la participation d'un petit groupe de personnes. Nous avons veillé à une répartition équilibrée des groupes de participants afin d'assurer une diversité de compétences et visions. Ces acteurs étaient répartis en plusieurs catégories, notamment :

- ▶ des *pure players* proposant de nouveaux modes de consommations culturelles,
- ▶ des acteurs plus établis des industries de contenus en cours de transformation,
- ▶ des créateurs et fournisseurs de nouvelles technologies (algorithmes, objets connectés, APIs, souvent incarné par des startups),
- ▶ des régulateurs,
- ▶ des chercheurs.

Si les ateliers collaboratifs sont au cœur du dispositif, d'autres modes opératoires (recherche documentaire ou entretiens qualitatifs) y ont été associés afin de générer les ressources et supports nécessaires à une réflexion partagée.

Ces ateliers ont eu un objectif très exploratoire : imaginer des usages et services dans l'industrie créative et culturelle des cinq prochaines années et poser les premières briques de leurs modèles économiques potentiels.

PHASE #1 ▶ Décrypter les modèles économiques actuels et leurs points de rupture

PHASE #2 ▶ Cartographier le présent pour mieux imaginer de futurs scenarios

PHASE #3 ▶ Prototyper des modèles économiques à 5 ans

EXEMPLE : SUPPORTS DE CRÉATIVITÉ REMPLIS LORS DU SECOND ATELIER

Support 1 - Project Board (disponible en annexe)

Ce canevas a permis aux participants de formaliser le service qu'ils ont eu à imaginer en tenant compte du marché investi, données nécessaires au fonctionnement du service et modèle des revenus.

Support 2 - Persona Board (disponible en annexe)

Pour que l'exercice soit le plus possible des attentes des utilisateurs de contenus culturels, les participants ont dû se mettre dans la peau d'un personnage type. Ils ont ainsi défini et analysé sa consommation de contenus culturels, ses besoins, ses freins et motivations au partage de données personnelles.

Project Board (exemple de remplissage)

Persona Board (exemple de remplissage)

PARTICIPANTS DU SECOND ATELIER

Julie Knibbe (Product Manager, Orange)
Thibault D'Orso (Fondateur, Spino, algorithmicien de recommandation pour les contenus vidéo)
Jana Moser (Investment Manager, Axel Springer)
Mael Primet (Chief Technical Officer et co-fondateur, SIBOS, modèles algorithmiques combinés et conseil médias personnalisés)
Numa Bourragué (Co-fondateur, E-cap Venture Capital, dans le secteur digital)
Thomas Beauvisage (Research Project Manager, Orange Labs)
Kevin Mellet (Chercheur, Orange Labs)
Hubert Saint Olive (Administrateur, J'ai Magazine, Impaqueur de référence sur le paysage électronique)
Sophie Peinado (Chargée d'étude, Roblox)
Antoine Durieux (Ingénieur de recommandation à la grande consommation)
Renaud Visage (Chief Technical Officer, Evernote, spécialiste de distribution et de gestion d'événement)
Gaetan Gachet (Head of Business Development, Apple Search Ads Service)
Trista Bridges (Partner, Rubelagette, premier support digital anglophone sur l'écosystème technologique français)

SCÉNARIO 1/

HyperMovie Generator

« Everyone can be a great movie creator »



HyperMovie Generator promet un monde « où tout le monde peut devenir un excellent créateur ».

Le service permet aux individus de créer des contenus vidéo originaux très facilement à partir d'une bibliothèque de contenus et des créations personnelles.

Le système permet d'explorer différentes pistes créatives en fonction d'un genre, d'un personnage, d'un sujet et/ou d'un public. D'un point de vue fonctionnel, le service est alimenté par des données sur les contenus et des analyses sur les préférences des publics auxquelles *HyperMovie Generator* a accès grâce à ses accords, ainsi que par des grandes bases de données agrégées sur les préférences des plateformes d'intermédiation de distribution de films et séries.

Le modèle économique est fondé sur une application gratuite, mais complétée d'une formule premium comprenant un abonnement donnant accès à plusieurs niveaux de services additionnels. D'abord, l'abonnement permet une utilisation illimitée. Ensuite, il donne la possibilité de partager plus largement les contenus grâce à des accords avec des « MCN », ces réseaux multi-chaînes sur les plateformes de vidéos (*YouTube*, *DailyMotion*) qui permettent à des centaines de milliers de créateurs de proposer leurs contenus à des centaines de millions de spectateurs. *HMG* a négocié avec plusieurs de ces MCNs afin de servir de « dénicheur de talents » pour eux. Un MCN va en effet enrôler des créateurs sélectionnés afin de permettre à chacun de gérer son propre canal de diffusion tout en utilisant leurs outils de promotion, marketing et surtout de monétisation. *HyperMovie Generator* sert en réalité de « premier niveau ». Si les *power users* de *HMG* parviennent à produire des contenus engageants et drainant des réactions positives, ils se voient proposer une chance sur un MCN adapté à leur style (intellectuel, jeunesse, reportage, vidéos musicales...). Même si très peu d'élus arrivent à vivre de leurs productions, les quelques *success stories* sont largement suffisantes pour attirer des créateurs en herbe sur *HyperMovie Generator*. *HMG* parvient même à attirer des créateurs qui sont loin d'être des amateurs mais qui trouvent dans cette voie une capacité à utiliser en toute liberté des moyens (de production comme de monétisation) qui imposeraient normalement de passer par les grands producteurs et diffuseurs.

Sont donc ciblés les auteurs et créateurs établis ou émergents. Les studios en quête permanente de blockbusters sont également une cible privilégiée, dans la mesure où *HyperMovie Generator* permet de tester des concepts ou des formules narratives. Le succès d'*HyperMovie Generator* vient de la passerelle qu'il semble

avoir réussi à créer entre le monde des contenus générés par les utilisateurs et les studios de production.

Par ailleurs, *HMG* est devenue spécialiste de l'adaptation dynamique des contenus et conseille dorénavant les grands studios pour qu'ils puissent intégrer à leurs blockbusters des éléments d'adaptation « discrète » de la narration ou des références contextuelles à différents publics. Par exemple, des technologies numériques permettront de changer la localisation de certaines scènes de films afin de les rapprocher de chaque public : un spectateur américain verra une scène se passant sur fond de *skyline* newyorkaise, alors qu'un spectateur chinois reconnaîtra la ville de Shanghai.



REVENUS

Les revenus de *HyperMovie Generator* sont essentiellement liés à la monétisation de l'audience de l'application et à la part des revenus ultérieurs des œuvres négociée par contrat entre l'auteur, *HMG* et le diffuseur. De plus en plus de marques se tournent directement vers *HMG* pour créer du « brand content », c'est-à-dire du contenu lié à la marque, qui est ensuite proposé comme boîte à outils « clés en main » pour réutilisation par les auteurs.



DÉFIS

Toute la difficulté consiste à capter et agréger massivement des données extrêmement granulaires sur les publics et leurs préférences. Il s'agira alors de définir le mécanisme d'incitation à la remontée d'information de la part des spectateurs pour mieux les connaître.



INSPIRATIONS

- ▶ Livres personnalisés / création automatique : [LostMyName](#)
- ▶ *L'Expansion* « [YouTube : les MCN ou réseaux de chaînes, accélérateurs d'audience](#) », 2014
- ▶ *Lagniappe* – [livraison dix-huit](#), newsletter de Nicolas Nova du 28 juin 2015
- ▶ *Popular highlights* - La pratique qui consiste pour le lecteur sur liseuse à surligner ses passages favoris et pour le fabricant à collecter et stocker ces courts extraits, peut être source d'inspiration pour imaginer un futur possible de la littérature. Le designer Nicolas Nova relaie le projet [Networked Optimization](#) porté par deux artistes, Silvio Lorusso et Sebastian Schmiege, aidés des « utilisateurs du Kindle Amazon ». Ils proposent des ouvrages papier dont l'originalité est de ne donner à voir que les phrases surlignées par leurs lecteurs. Ils espèrent ainsi faire réfléchir sur les modes d'optimisation de l'écriture et pointent qu'à travers le surlignage, « l'acte de lecture devient un processus de data-mining ».

www.hypermoviegenerator.com

Télécharger dans l'App Store

DISPONIBLE SUR Google play

Je me connecte

Créer mon compte

HyperMovie Generator

“Tout le monde peut devenir créateur!”

Qu'est ce que HyperMovie Generator ?

Créez des contenus vidéo originaux très facilement à partir à la fois d'une bibliothèque de contenus et des créations personnelles. Explorez différentes pistes créatives en fonction d'un genre, d'un personnage, d'un sujet et/ou d'un public.

Comment ça marche?

- Créer mon histoire**
- Utiliser les outils de production**
- Adapter le contenu aux goûts du public**
- Diffuser mes créations**

Vous aussi devenez la prochaine star de la vidéo en ligne !

- #PETflix
- #MusculArthur

Qui sommes nous? FAQ Assistance Vie privée & confidentialité Presse Contact Mentions légales

Credits images pour cette page de haut en bas : CC BY Kirill Khizhev - CC by Gabriel Garcia Marengo - CC by Juloshka - CC by CoachClub

HyperMovie Generator

“Hyper Movie Generator promet un monde où tout le monde peut devenir créateur”.
Prototype de la page d'accueil du site internet.

Réalisation : Mylène Podvin/CNIL



SCÉNARIO 2/

Life tracks

Life Tracks est une nouvelle génération d'assistant personnel dédié à la productivité et au bien-être de l'utilisateur. L'application *Life Tracks* innove en associant le profil génétique ainsi que l'état émotionnel pour proposer une écoute de contenus musicaux personnalisés, accompagner les activités de l'utilisateur, mais aussi agir sur son humeur, son stress... Ce service repose sur un objet connecté relié à des capteurs ouverts disséminés dans notre environnement. Il permet à l'utilisateur de gagner du temps, de réguler son humeur, d'apaiser ses troubles psychologiques, d'améliorer sa créativité tout en découvrant constamment de nouveaux morceaux de musique. Les dernières avancées scientifiques issues initialement de recherches en neuromarketing ne laissent désormais plus aucun doute sur les formidables capacités de la musique à influencer les humeurs, booster la créativité ou favoriser la concentration.

Life tracks tire pleinement profit de ces vertus en étant alimenté par des données particulièrement intimes de l'individu et en croisant son profil génomique, ses caractéristiques physiques ou psychologiques, ainsi que l'environnement dans lequel il se trouve (bureau, maison, transport) avec ses appétences musicales. La combinaison de ces variables génère des modèles permettant de produire l'ambiance musicale la plus adaptée. L'innovation provient aussi de la nature des algorithmes sur lesquels reposent *Life Tracks* : le système de recommandation délaisse la connaissance *a posteriori* des comportements sociaux et consommations culturelles au profit d'un système qui cherche à prédire *a priori* ces comportements. On passe ainsi d'un modèle centré sur la corrélation (si l'utilisateur aime ce produit, il pourrait aimer cet autre produit) à un modèle centré sur la causalité (si l'utilisateur aime ce produit, il est aussi « programmé » pour aimer cette autre offre).

REVENUS

Les revenus de *Life Tracks* proviennent essentiellement de la commercialisation de l'objet connecté – en format bijou d'oreille pour homme et femme dans sa première version. En cohérence avec le fait d'avoir complètement écarté la monétisation par la publicité (en raison de la sensibilité des données : génome, l'humeur etc.), les fondateurs proposent deux prix pour l'objet selon que l'utilisateur opte ou non pour l'option « contribution à la connaissance collective » qui signifie que ses données sont valorisées en mode anonymisé, pour améliorer le service. Il obtient une réduction de 10 % s'il opte pour cette option, dans le cas contraire ses données ne sont pas conservées à d'autres fins que de contextualiser son expérience. En B2B, les revenus proviennent de la vente de licences pour accéder et réutiliser le logiciel d'intelligence artificielle.

DÉFIS

Au niveau des algorithmes, la difficulté consiste à classifier les humeurs de l'individu, les caractéristiques de l'environnement et à établir le degré de pertinence d'un morceau sur la base du contexte, de l'humeur et du génome. Par ailleurs, il faudra identifier et extraire les éléments clés d'un morceau de musique pour évaluer son adéquation aux données de l'individu et mesurer la part de détermination du génome sur les goûts culturels.

INSPIRATIONS

Donovan Sung, gestionnaire de produits « découverte et recommandations » pour *Spotify* [anticipe ainsi d'utiliser les capacités étendues des capteurs](#) du smartphone : « peut-être qu'avec les capteurs de mouvements des téléphones nous pouvons commencer à deviner si les utilisateurs sont en train de courir, de faire du vélo ou de conduire ? Peut-être qu'avec le capteur de température, ou le capteur cardiaque, nous pourrions deviner si l'utilisateur est tendu ? »¹ Ce raisonnement peut évidemment s'étendre aux autres capteurs connectés (bracelets, montres...)². Daniel Ek, fondateur de *Spotify*, prédit même que « dans un futur pas si lointain, peut-être d'ici une dizaine d'années, nous serons tous connectés, tout le temps, avec des *Google Glass* par exemple, des capteurs intégrés dans nos vêtements ou d'autres dispositifs informatiques portables. Nous n'aurons plus à manipuler un écran pour écouter de la musique, celle-ci se lancera par un mouvement des mains ou juste par la voix des bandes musicales adaptées à nos émotions. »³ Cette vision est aussi partagée par Ken Hertz de *memBrain*, [qui pronostiquait à Wired UK en 2013](#) que la recommandation ne serait plus dictée par nos préférences précédentes ou nos cercles sociaux mais par des feedbacks en temps réel des terminaux connectés que nous porterons⁴.

Life tracks ambitionne ainsi de s'imposer comme un *Google Now* musical, en étant en mesure d'anticiper mieux que l'utilisateur lui-même l'ambiance qui lui sera la plus profitable. La génétique y joue un rôle fondamental en pré-déterminant les goûts et la personnalité de l'utilisateur. à la manière de *Genetrainer* qui propose des plans d'entraînements sportifs en fonction du profil génétique, *Life Tracks* vise « l'hyperpersonnalisation » du service. La compréhension du contexte qui entoure l'utilisateur n'est pas sans rappeler le fonctionnement de la pyramide *Prizm* (voir chapitre II-3) même si l'objectif n'est pas seulement le divertissement mais bien de placer l'individu dans des conditions de productivité optimale. De nombreuses startups investissent désormais le champ de la compréhension des émotions humaines, comme *Melomind* dont le casque permet de se connecter au cerveau pour réaliser des séances de relaxation, avec la promesse d'effets bénéfiques à long terme sur la gestion du stress.

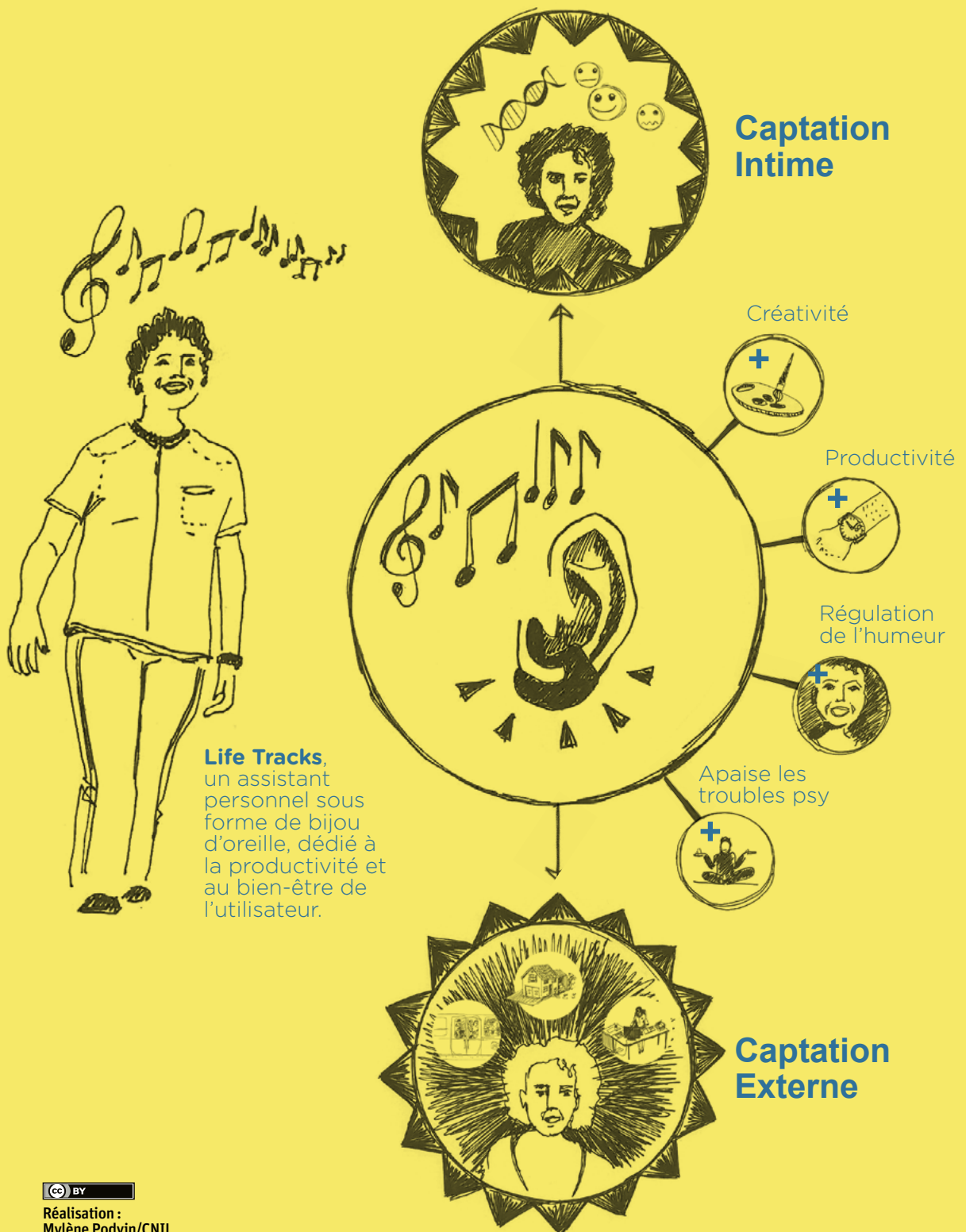
L'impact de la musique sur nos émotions fait d'ailleurs l'objet de nombreuses publications : GUILLOU, Philippe (2013) : « [Lettre Neuromonaco 86 : Les effets de la musique](#) ». (14 Oct 2013) *Neuromonaco*.

¹ TORKINGTON, Janel. « [How technology is changing the way we listen to music](#) », *The Daily Dot*, 25 mars 2014.

² Sur ce sujet, voir le cahier IP numéro 2 : CNIL. [Le corps, nouvel objet connecté ?](#), mai 2014.

³ MALTER, Jordan. « [Spotify's vision for the future of music](#) », *CNN Money*, décembre 2013 (Interview video)

⁴ COLLECTIF. « [The big question : "What is the future of music?"](#) », *Wired UK*, août 2013.



Réalisation :
Mylène Podvin/CNIL

Life Tracks

L'application Life tracks associerait le profil génétique et l'état émotionnel pour proposer une écoute de contenus musicaux personnalisés pour accompagner les activités de l'utilisateur.

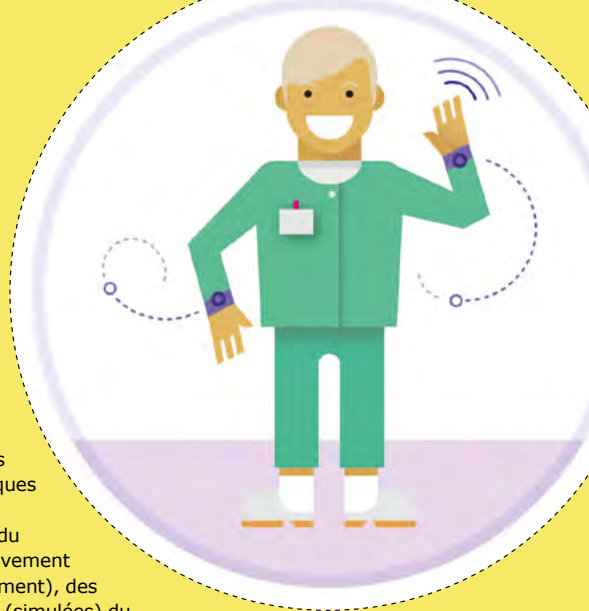
SCÉNARIO 3 /

TraumaData

TraumData est une entreprise allemande spécialisée dans le développement de contenus immersifs augmentés dans le domaine médical. Le nom de l'entreprise est né d'une double référence au mot «rêve» en allemand (*Traum*) et au mot anglais «trauma» qui évoque à la fois l'expérience traumatisante et la blessure physique. Progressivement, la réalité virtuelle et la robotisation ont pris une place de plus en plus importante dans les hôpitaux, et en particulier dans les salles d'opérations. Pour certains gestes extrêmement précis, les chirurgiens travaillent main dans la main avec des assistants robotisés auxquels sont de plus en plus souvent adjoints des terminaux de réalité augmentée ou virtuelle. Des systèmes complexes et coûteux ont été installés dans les hôpitaux, et les chirurgiens doivent aujourd'hui associer une parfaite maîtrise de ces outils informatiques et robotiques, dans des situations quotidiennes de stress, avec des risques importants pesant sur les patients en cas de mauvaise manipulation. TraumData propose des solutions complémentaires à ces installations lourdes et complexes, qui tirent partie des pratiques vidéoludiques pour faire de la formation et permettre aux chirurgiens de travailler sur l'amélioration de leurs compétences sans y penser. En effet, la compagnie a constaté que, comme dans toutes les strates de la société, un pourcentage croissant de professionnels de tous âges jouent, de manière plus ou moins régulière, aux jeux vidéo. TraumData édite donc une gamme de jeu «sérieux» de réalité augmentée ou virtuelle à destination des chirurgiens et leur propose de tester, d'améliorer et de comparer en continu leur dextérité et leurs aptitudes. Ce jeu fonctionne grâce à un dispositif qui capte les données (en particulier les mouvements) des joueurs en temps réel et les retranscrit dans des interfaces tactiles et haptiques réalistes, à l'intérieur d'un environnement immersif reproduisant fidèlement des situations professionnelles. Mais TraumData ne s'est pas arrêtée là, d'autres entreprises proposant déjà des produits de ce type. D'abord, le jeu est personnalisé : il est alimenté par des données historiques

¹ API : Interface de programmation d'application. Ces interfaces permettent de se «brancher» sur une application pour échanger des données.

de l'environnement de travail, des données de l'état initial de patients réels du praticien, des données encyclopédiques et académiques, des données analytiques du dispositif tactile (mouvement de l'utilisateur notamment), des réactions et réponses (simulées) du patient. Cette personnalisation de l'expérience explique une partie du succès de l'entreprise. Celui-ci s'explique aussi par l'intégration «sans couture» de moments de formation dans d'autres jeux, à vocation purement ludique en apparence. TraumData a noué en 2019 un partenariat avec plusieurs éditeurs de jeux vidéo et avec certaines plateformes de diffusion dématérialisée de jeux (tel *Steam* de Valve). Dans le cadre de cet accord, TraumData peut accéder à des informations sur les pratiques vidéoludiques de ses clients (quels types de jeux ont leur préférence, comment se comportent-ils dans le jeu ?), utiliser ses algorithmes pour analyser ces informations en les confrontant à ses propres données (sur les pratiques professionnelles de l'utilisateur) et enfin, à partir de cette analyse, diffuser par l'intermédiaire des APIs¹ proposées par les plateformes et les éditeurs de légères modifications du jeu, par exemple de la jouabilité ou de la dextérité requise pour certaines actions à accomplir afin de faire progresser les compétences de la personne d'une manière invisible, discrète mais évaluable. Par exemple, certains jeux vidéo peuvent être très utiles pour améliorer l'habileté, la précision de certains gestes de la main, la résistance au stress, la capacité à développer des réactions réflexes... En renforçant son système de formation et en l'intégrant non seulement dans son jeu sérieux mais aussi de manière plus douce dans d'autres expériences ludiques, TraumData a su devenir un service très prisé par les chirurgiens. Progressivement, son modèle s'étend à d'autres métiers.



REVENUS

Le montant des investissements en équipements robotiques et immersifs dans le secteur de la santé a permis à l'entreprise de commencer par ce domaine avec un modèle économique robuste, les revenus de TraumData reposant d'abord sur un modèle B2B (vente du système de réalité virtuelle à des hôpitaux et universités, formations des étudiants en médecine). L'entreprise gagne aussi de l'argent dans le cadre de ses contrats avec des éditeurs de jeux vidéo mais l'accès à leurs données lui coûte aussi beaucoup d'argent. TraumData explore de nouveaux modèles d'affaires, comme la réalisation d'études sur l'efficacité des gestes et l'analyse des bonnes pratiques professionnelles... L'entreprise songe surtout à étendre son modèle au-delà de ce premier marché des professionnels de santé.

DÉFIS

Superposer numérique et réalité pour mieux former et éduquer. Ce procédé de production, qui fait la part belle aux interactions entre mondes physique et numérique, transforme le jeu en outil de formation. A l'heure où les MOOCs (cours en ligne) se généralisent, on peut imaginer étendre cette logique à une large part de la formation professionnelle, voire au système éducatif. Or, la maîtrise de telles informations est un enjeu de pouvoir important car les informations sur les compétences d'un individu ont une très haute valeur stratégique, pour les entreprises comme pour les Etats (voir à ce sujet les analyses de [Claude Kirchner, Inria](#)).

INSPIRATIONS

ADEÏDA, Kenza. «[Jean Chaoui met la réalité virtuelle au service de la chirurgie](#)», *L'Atelier BNP-Paribas*, 2014. «Et pour finir, Jean nous confie qu'il est envisageable d'imaginer une utilisation de cet outil pour la formation des apprentis chirurgiens.»
 GUILLAUD, Hubert. «[Vers des technologies de l'empathie](#)», *InternetActu*, 2015 et site <https://www.crystalknows.com/> «When you looked up ***'s personality profile, Crystal looked through all of the answers provided directly by people who know him, as well as thousands of publicly available online data sources to find information written by or about Rand. Crystal then runs a personality detection analysis on the text to match them with one of 64 different personality types within a few seconds. (...) The insights below are recommended strategies for communicating with someone that has ***'s specific personality type, based on several widely-accepted personality assessments, like the Five-Factor Inventory, DiSC, and True Colors.»
 KIRCHNER, Claude. *Big data and education*, 2014
 MONTJOYE (DE), Yves-Alexandre. [What Can Your Phone Metadata Tell About You?](#)
 ROUSSEAU, Rémi. [Virtual surgery gets real](#), 2014 et site <http://www.surgevry.com/> «Virtual surgery gets real. What the Oculus Rift could mean for the future of medicine?»
 SCIFO, Guillaume. «[Les industriels commencent à s'intéresser à la réalité virtuelle pour la médecine](#)», *L'Atelier BNP-Paribas*, 2015. «De mon côté, je travaille à utiliser les technologies pour permettre aux personnes soit de récupérer des capacités, soit de faire face à des problèmes. Globalement, cela correspond à la prise en charge du handicap. (...) Mais il y a aussi un autre volet : la formation des médecins avec des simulations.»
 WABER, Ben. *People Analytics*, 2013

WIZED

BUSINESS DESIGN ENTERTAINMENT GEAR SCIENCE SECURITY

SHARE

PLAY 118

TWEET 524

SING

COMMENT

EMAIL

SPONSORED CONTENT BY TRAUMDATA

LES JEUX VIDÉO AU SERVICE DU « TALENT MANAGEMENT » GRÂCE À TRAUMDATA.

14 MAI 2021



« Aurélien est très analytique, plutôt sceptique et concentré sur les résultats. Il sera plus à l'aise face à des sollicitations brèves (3 phrases ou moins) contenant le sujet principal dans la première phrase, et écrites dans un style simple. »

« Sylvain s'épanouit dans un environnement peu structuré, aime explorer de nouvelles idées. En vous adressant à lui, faites de petites blagues et utilisez un langage expressif du point de vue émotionnel (et des *emojis*). »

Les consommateurs se sont habitués à l'idée que les vendeurs disposaient d'outils d'analyse de la personnalité offrant ce type d'informations, et plusieurs entreprises ont développé des outils de personnalisation des messages publicitaires ou de prises de contact commerciales optimisées selon la personnalité prédite du récepteur. Aujourd'hui, les internautes les plus avisés savent que leur navigation web, leurs messages instantanés sur certains services gratuits et surtout leurs publications en ligne, par exemple sur les réseaux sociaux, risquent d'être analysés par des acteurs intrusifs pour affiner leur portrait vis-à-vis des fameux « *Big Five* », les 5 grands traits de personnalité mis en évidence et affinés empiriquement depuis les années 1990. Êtes-vous ouvert à l'expérience nouvelle ? Conscientieux ? Extraverti ? Agréable et coopératif ? Stable émotionnellement ? Le modèle OCEAN (Ouverture, « Conscientiosité », Extraversion, Agréabilité, « Neuroticisme »²) est devenu un outil omniprésent au cœur des modèles de tracking publicitaire.

Mais ces analyses manquent souvent d'une matière première fiable en quantité et en qualité : il faut inférer beaucoup d'informations pour passer d'une liste de *tweets* à des traits tranchés de personnalités au-delà de quelques dizaines de combinaisons courantes du modèle OCEAN. Pour cette raison, les entreprises ont continué à avoir recours en interne à des questionnaires ou des bilans plus traditionnels plutôt qu'à des analyses basées sur le big data.

Progressivement, des *startups* ont proposé des outils nouveaux, basés sur l'analyse de capteurs dans les locaux professionnels ou sur l'analyse sémantique des messages échangés sur les outils professionnels (messagerie, réseaux sociaux internes, etc).

Problème : en restant confinée au comportement professionnel quotidien, l'analyse de personnalité qui transparaît peut être incomplète. Comment résoudre ce problème ? *TraumData* investit massivement en R&D pour faire sauter les frontières de ces analyses, comme nous l'explique son *chief data and future-vision officer*, Gregor Haus : « cela fait plusieurs années que nous mobilisons, pour nos produits de formation, les données issues des comportements des joueurs de jeux vidéo. Ces données se sont avérées extrêmement utiles dans ce contexte. Nos clients souhaitent aujourd'hui aller plus loin et utiliser ces informations pour faire du *talent management*. *TraumData* va analyser le comportement des joueurs et les classer selon les méthodes les plus utilisées de classification de la personnalité (big5, MBTI). Ces classifications seront ensuite utilisées par les services de ressources humaines, de *people analytics* et de *talent management*. » Pour le moment, ce n'est qu'un prototype : *TraumData* va, avec l'accord d'un éditeur de jeu vidéo, intégrer des événements significatifs du point de vue de la personnalité dans des jeux (par exemple, dans des passages où le joueur doit choisir une réponse ou une action parmi plusieurs possibles). L'entreprise classera ensuite les joueurs très précisément dans plusieurs centaines de types de personnalité, et proposera à l'employeur une stratégie adaptée de gestion des talents. À l'avenir, *TraumData* souhaite proposer ses jeux de données dans les stratégies de *people analytics*, et ainsi améliorer l'efficacité des stratégies de ressources humaines *data driven*. « Nous espérons offrir à nos clients une amélioration de 15% de l'efficacité de leur algorithme de recrutement, de rétention (ceux qui prédisent la probabilité qu'un employé quitte l'entreprise) et probablement aussi de leurs algorithmes de diversité et de design des espaces de travail », affirme Haus.

GET WIZED
CC by Michael Praetorius
Recyclez dès maintenant vos lunettes connectées et repartez avec les toutes nouvelles lentilles de contact Lens3000.

LATEST NEWS

HYPERMOVIE GENERATOR
ZOOM SUR CETTE NOUVELLE PÉPITE QUI PROMET À TOUS UN FUTUR DE CRÉATEUR !

UN BIJOU DE PRODUCTIVITÉ
LIFE TRACKS EST UN ASSIANT PERSONNEL DÉDIÉ À LA PRODUCTIVITÉ ET AU BIEN-ÊTRE QUI PREND LA FORME D'UN BIJOU D'OREILLE. NOUS L'AVONS TESTÉ.

Traum Data

Réalisation : Mylène Podvin/CNIL

Extrait d'une campagne fictive de *native advertising* de TraumData dédié à l'analyse comportementale.

SCÉNARIO 4/

The collective mood master



The Collective Mood Master a été le projet phare de la *Disrupt conférence* organisée par *TechCrunch* à Lisbonne fin 2019 – désormais la capitale européenne des cultures numériques. Les fondateurs n'ont cessé de rappeler les origines du projet né il y a maintenant 2 ans avec l'intuition que les commerces de proximité pouvaient personnaliser leur relation sur la base des goûts culturels de leurs clients plutôt qu'en cherchant à leur proposer [des offres promotionnelles basées sur leurs beacons...](#)

« Imaginez-vous aujourd'hui que vous puissiez entrer chez votre coiffeur sans que vos goûts musicaux ne soient intégrés à la playlist du salon ? » s'amuse Jeffrey Cross le CEO de la plateforme. L'offre de *The Collective Mood Master* est finalement assez simple, il s'agit de connecter les goûts des clients aux offres des commerçants. Les services de la plateforme se décomposent ainsi en :

► **une offre dédiée au marché B2C** permettant aux utilisateurs de vivre une expérience personnalisée dès lors qu'ils entrent dans les commerces partenaires, reconnaissables à leur macaron « Collective Mood Ready », comme des VTC, des coiffeurs ou dans certains espaces privatifs des bars et restaurants les plus à la mode qui diffusent leurs playlists favorites ;

► **une offre (B2B) en direction des marchands** qui souhaitent mieux connaître les goûts et habitudes culturels de leur zone de chalandise. C'est sur la base de ce service que *Netflix* a développé la franchise *YourCineflix*, la première chaîne de cinémas de quartier dont la programmation est basée exclusivement sur les préférences locales des utilisateurs du service. Ainsi, certaines salles ne diffusent presque que des films de science fiction, quand d'autres sont spécialisées dans les polars des années 80.

Pour les fondateurs, *The Collective Mood Master* est le premier service « VRM collectif » qui se base sur les préférences locales des utilisateurs pour créer une offre pertinente et personnalisée. Dans ses cartons, la startup travaille actuellement sur un projet très ambitieux visant à mesurer en temps réel le niveau d'émotion généré par la diffusion d'une chanson ou par certaines scènes lors de la projection d'un film. L'objectif de l'entreprise est d'enrichir son offre B2B autour de cette fonctionnalité, en assurant un retour en temps réel aux marchands sur le niveau de satisfaction de leurs clients...

REVENUS Côté utilisateur final, les revenus de *The Collective Mood Master* proviennent des abonnements premium permettant de vivre des expériences personnalisées chez les commerçants affiliés au programme « Collective Mood Ready ». Les utilisateurs gratuits de l'application peuvent acquérir des points leur offrant des minutes premium s'ils acceptent de partager leur localisation, les commerces qu'ils fréquentent et en synchronisant leurs comptes relatifs aux contenus culturels. Toutes ces informations alimentent la base de données à laquelle peuvent s'abonner les commerces pour mieux comprendre leur zone de chalandise et adapter leur ambiance aux socio-styles locaux.

DÉFIS Le défi principal sur lequel travaillent les ingénieurs de la firme est de définir des styles « médians » aussi appelés « home staging » en interne (pour passe partout). Il s'agit en effet de mieux couvrir les zones géographiques caractérisées par une forte hétérogénéité des styles qui compliquent l'atteinte d'une masse critique – même si ce seuil reste bas, selon les fondateurs.

INSPIRATIONS Le projet HNL pour « Heart Never Lies » (le cœur ne ment pas) vise à développer une approche combinatoire des mesures émotionnelles. Dérivé d'un algorithme développé par un médecin pour définir le niveau de douleur, l'objectif est aujourd'hui de mesurer le niveau d'émotion à des fins marketing. Parmi les exemples d'application où l'émotion est prépondérante dans l'acte d'achat figure la musique. « *Un chanteur à succès ou en devenir, a enregistré 25 chansons en studio. Pour l'album il doit en choisir 12 et 2 pour les singles (l'usage veut que d'un album, l'artiste pourra extraire deux singles). Une fois de plus, HNL pourra mesurer l'effet émotionnel comparatif des 25 chansons enregistrées sur un panel propriétaire correspondant à la cible visée par le producteur.* »

Discontrol Party est une performance artistique réalisée en 2011 à l'initiative de Samuel Bainchini à la Gaité Lyrique. Chaque spectateur possède un harnais récoltant des informations sur ses déplacements, sa consommation d'alcool,.... Ces données alimentent un dispositif de surveillance intelligent qui modifie les caractéristiques du concert.

Dans son billet sur les données de *Netflix* « [Netflix, Big Data, and what hollywood is missing out on](#) », le blogueur Jason Lange imagine les services qui pourraient être conçus sur la base d'une analyse des données prises localement. Il imagine ainsi des cinémas de quartier qui utiliseraient les données de consommation de leur code postal pour orienter la programmation des films.

The Business Model Canvas

Designed for:

Designed by:

Date:

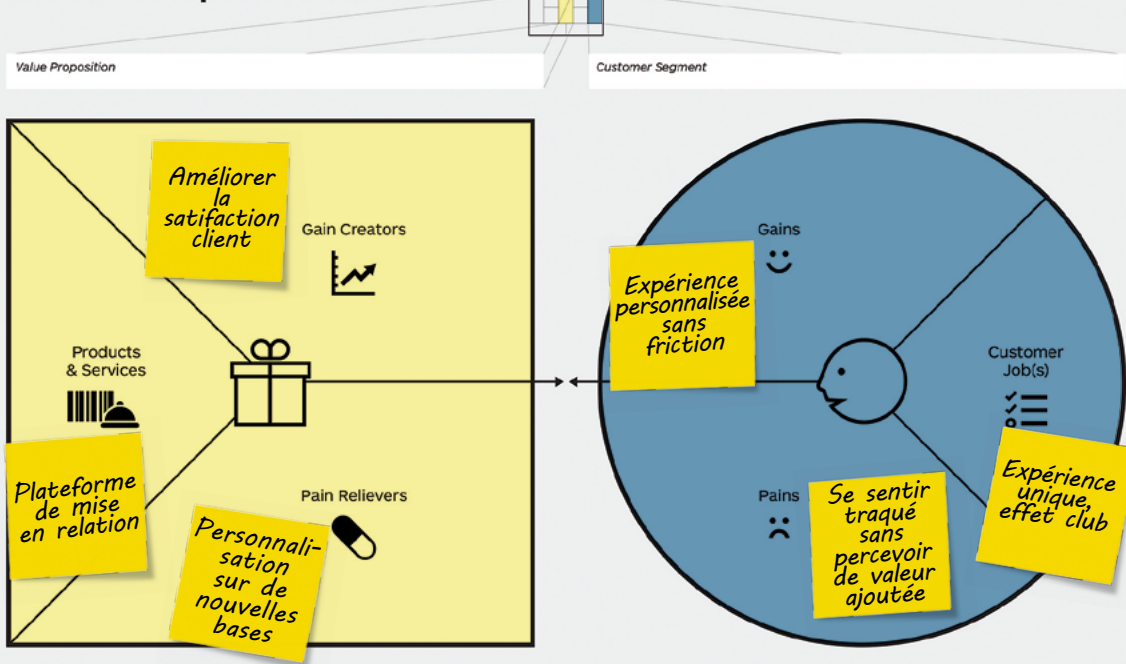
Version:



DESIGNED BY: Strategyzer AG

Strategyzer
strategyzer.com

The Value Proposition Canvas



COPYRIGHT: Strategyzer AG

Strategyzer
strategyzer.com

Réalisation : Mylène Podvin/CNIL

The Collective Mood Master

“La première plateforme qui connecte les goûts des clients aux offres des commerçants.”
Représentation de la création et de la captation de valeur.

De quoi ces services sont-ils le nom ?

Essai de synthèse sur les enseignements de cet exercice de prospective de services

LA COMBINAISON DE LA DONNÉE ET DU HARDWARE BOULEVERSE LA CHAÎNE DE VALEUR DE SECTEURS TRADITIONNELS ET FAIT ÉMERGER DE NOUVEAUX MODÈLES ÉCONOMIQUES.

En collectant de la donnée, l'objet permet d'affiner la compréhension de comportements humains, bien au-delà de la simple consommation de contenus culturels. Celle-ci devient un point d'entrée vers la captation de données contextuelles extérieures, variables indispensables dans la compréhension des comportements ou phénomènes sociaux.

La valeur est donc double que l'on soit du côté du fournisseur ou de l'utilisateur, et cela ouvre la voie à des modèles économiques alternatifs au ciblage publicitaire (« On croit au service, pas à la publicité »).

Pour l'utilisateur du service :

- ▶ Une valeur d'usage qui permet d'optimiser son quotidien à partir des données collectées par l'objet,
- ▶ son positionnement au centre d'un écosystème de services innovants dès lors que l'infrastructure est ouverte et qu'il a la maîtrise sur les échanges de données. Les données peuvent en effet être diffusées à travers des services tiers. L'avantage comparatif de l'objet est alors décuplé par le fait qu'il s'intègre à une myriade d'applications.

Pour le fournisseur du service :

- ▶ Une valeur financière à la croisée du B2C et du B2B. Côté B2C, il s'agit de la vente d'un objet au design simple et unique. Côté B2B, il s'agit de la monétisation de jeux de données anonymisés (sur ce sujet voir l'avis d'avril 2014 du G29, le groupe des CNILs européennes, sur les techniques d'anonymisation) et de services de *data scientists* dans les processus de décision des acheteurs, organismes ou entreprises qui cherchent à optimiser leurs opérateurs et leurs offres avec une meilleure connaissance du marché.

L'ÈRE DE LA COLLECTE MULTI-SOURCES DE DONNÉES ET DE L'« HYPERCONTEXTUALISATION » POSE LA QUESTION DE LA NEUTRALITÉ ET DE LA PERTINENCE DES ALGORITHMES.

Les technologies permettent aujourd'hui de collecter et traiter des millions d'enregistrements en temps réel, et par ailleurs le stockage de la donnée n'est plus un problème insurmontable. L'ensemble des scénarios explorés lors de l'atelier révèle que les services des prochaines années pourront collecter et croiser des données de différentes natures : données de santé, données sociales (comportement de consommation de biens culturels par exemple), données contextuelles, données historiques, etc. Le croisement de l'ensemble de ces informations place le traitement algorithmique dans l'hypercontextualisation : il est possible de rétablir finement le contexte qui fait que nous adoptons tel ou tel comportement. Mais qui sélectionne les données alimentant ces modèles ? Qui leur accorde leur valeur mathématique, choisit leur pondération ? Absence de neutralité et de loyauté et risque de reproduction de préjugés sur les comportements et goûts « minoritaires » sont deux des questions éthiques les plus brûlantes que cette *algorithmisation* du quotidien soulève (voir partie III, encart innovation / protection des droits).

AVEC LA COLLECTE ET LA VALORISATION CONTINUE DE DONNÉES PERSONNELLES, LE SERVICE SE FONDRA-T-IL AVEC L'UTILISATEUR... AU POINT DE FAIRE DISPARAÎTRE L'EXPÉRIENCE UTILISATEUR ?

Les scénarios révèlent une tendance à la disparition de toute action de l'utilisateur pour faire fonctionner le service. Cela soulève de nouvelles problématiques quant aux modalités d'acceptation du service et de connaissance des processus techniques nécessaires à son fonctionnement : peut-on accepter et comprendre le fonctionnement d'un service dont le fonctionnement est automatisé et invisible et dans lequel l'expérience utilisateur n'est plus un sujet car elle a tout simplement disparu ? (voir partie III)

Cette évolution est d'autant plus vraie lorsque l'on prend en compte la source des données. L'utilisation de données de santé (physique et physiologique), voire de l'ADN — symbole ultime de l'identité de l'individu — est à ce titre révélatrice. Dans le scénario d'un croisement continu des données ADN et contextuelles traitées par du *machine learning*, on pourrait ainsi imaginer un service dont la recommandation se substitue à l'autonomie de décision de l'individu.



« DE GRANDS pouvoirs impliquent de grandes RESPONSABILITÉS »

CC BY ND CNIL & Eventypo by Geoffrey Dorne

LES TECHNOLOGIES PRÉDICTIVES PERMETTENT AUX DISTRIBUTEURS DE CONTENUS CULTURELS D'ÉQUILIBRER L'OFFRE ET LA DEMANDE, PREMIÈRE ÉTAPE VERS LA CONNAISSANCE UNITAIRE DE CHAQUE UTILISATEUR

Historiquement, le centre de gravité du secteur des biens culturels est concentré au milieu de chaîne : par sa position, le producteur concentre le capital et maîtrise à la fois la réalisation et la distribution. Mais avec le numérique et les technologies de personnalisation, le distributeur peut connaître très finement les attentes des utilisateurs, et donc peut maîtriser la recette du *hit*. Dès lors, les distributeurs de contenus culturels remontent la chaîne de valeur jusqu'à réaliser et produire des contenus, ce qui est le cas de *Netflix* avec les séries télévisées. Le distributeur joue ainsi un rôle central dans l'équilibre de l'offre avec la demande. En découle deux

externalités : pour l'utilisateur, une amélioration de la satisfaction ; pour le distributeur, une augmentation des revenus.

Au-delà du secteur des biens culturels, le développement des technologies de prédiction invite à reconsidérer notre approche du traitement des données. Avec le *big data*, on passe en effet d'une approche statistique à une approche de la connaissance unitaire des individus. C'est une question de granularité de la donnée, mais elle a des implications au regard du potentiel du *machine learning* et de l'intelligence artificielle : elle vient éclater le fonctionnement existant des algorithmes actuellement fondé sur la base de groupes types.

Plusieurs scénarios de services montrent même que nous pourrions aller vers un système de recommandation qui délaisse la connaissance *a posteriori* des comportements sociaux et consommations culturelles, au profit d'un système cherchant à prédire *a priori* des comportements à partir des critères supposés objectifs de la génétique.

ANNEXES



ACCOMPAGNEMENT MÉTHODOLOGIQUE ET D'INNOVATION : WEDESIGNSERVICES ET FIVEBYFIVE



Merci à Chloé, Kat, Latifa, Christophe,
Mehdi, Rafael et Thibault.

PERSONNES RENCONTRÉES

Stéphane Allaire - Directeur services
et contenus, [Bouygues Télécom](#)

Aurélien Bellanger - Romancier,
auteur de *La Théorie de l'information*
et *L'Aménagement du territoire*

Virginie Berger - Fondatrice,
[Don't Believe the Hype](#)

Thibault d'Orso - Co-fondateur et COO,
[Spideo](#)

Brett Goldstein - UX designer,
<http://brettgoldstein.com/>

Hubert Guillaud - Journaliste
[InternetActu](#)

Rand Hindi - Co-fondateur et CEO, [SNIPS](#)

Janine Langlois - Présidente, [Forum des
médiias mobiles](#)

Christophe Montague - Ex Directeur
NRJ Europe - co-fondateur et CEO, [5M
Ventures](#)

PARTICIPANTS DU PREMIER WORKSHOP (23 octobre 2014)

Louis Aubert - Fondateur, [Tracktl](#)

Reda Berrehili - [Mixo TV](#)

Johanna Castel - [Cap digital](#)

Barbara Chazelle - [Direction de la
prospective](#), Groupe [France TV](#)

Florence Icard - Chargée de mission
études, [Hadopi](#)

Philippe Le Guern - Sociologue,
[Université de Nantes](#)

Raphael Lebas - Fondateur, [AËDLE](#)

Catherine Ramus - Ingénieur et designer,
Orange Labs et [artiste](#)

Pierre Rudolf Gerlach - CTO, [Qleek](#)

Edward Silhol - co-fondateur, [Stample](#)

PARTICIPANTS DU SECOND WORKSHOP (7 NOVEMBRE 2014)

Thomas Beauvisage - Sociologue,
[Orange Labs](#)

Numa Bourragué - Co-fondateur et COO,
[eCap](#)

Trista Bridges - Partner, [Rude Baguette](#)

Thibault D'Orso - Fondateur, [Spideo](#)

Antoine Durieux - Co-fondateur et CEO,
[Alkemics](#)

Gaëtan Gachet - Head of Business
Development, [Algolia](#)

Julie Knibbe - Product Manager, [Deezer](#)

Kevin Mellet - Chercheur, [Orange Labs](#)

Jana Moser - Investment Manager, [Axel
Springer](#)

Sophie Peinado - Chargée d'études,
[Hadopi](#)

Maël Primet - Co-fondateur et CTO, [SNIPS](#)

Hubert Saint-Olive - Administrateur,
[Trax Magazine](#)

Renaud Visage - Co-fondateur et CTO,
[Eventbrite](#)

nom / name _____

profession / job

age
 ..

citation / quote _____

“ _____

_____ ”

	MUSIQUE MUSIC <input type="checkbox"/>
	VIDÉO VIDEO <input type="checkbox"/>
	JEU VIDÉO VIDEO GAME <input type="checkbox"/>
	CONTENU ÉCRIT WRITTEN CONTENT <input type="checkbox"/>

HEURES PAR JOUR
HOURS PER DAY

MOTIVATIONS DRIVERS

communauté / community		reconnaissance sociale / esteem		découverte / discovery		respect de la vie privée / privacy		coût / cost		
------------------------	--	---------------------------------	--	------------------------	--	------------------------------------	--	-------------	--	-----	--	-----	--

VOLONTÉ À PARTAGER SES DONNÉES
WILLINGNESS TO SHARE DATA

1
2
3
4

En échange de / In exchange for _____

appareil
device

atelier Prospective #2

FAVE
BY
FAVE

PERSONA BOARD

CC BY

78 | CNIL | CAHIERS IP - LES DONNÉES, MUSES ET FRONTIÈRES DE LA CRÉATION

EXPLORATION PROSPECTIVE DE LA PLACE DES DONNÉES PERSONNELLES DANS LES MARCHÉS DES CONTENUS CULTURELS ET LUDIQUES

CNIL · We design services

FIVE
BY
FIVE

ip@cnil.fr

Debriefing du workshop #1

Le premier atelier de prospective organisé par la CNIL en collaboration avec WeDesignServices a permis de réfléchir collectivement à l'utilisation des données dans les marchés des contenus culturels et ludiques.

Les données sont devenues une brique centrale dans la stratégie des entreprises, en rendant possible une personnalisation des services, des recommandations et une monétisation des profils des utilisateurs.

Dans leur recherche permanente d'audience, les industries culturelles sont ainsi de plus en plus naturellement consommatrices et créatrices de données et adoptent des stratégies variées d'engagement du public et de capture de l'attention.

Ces stratégies mèneront-elles à l'émergence de nouveaux modèles économiques autour des données des utilisateurs ? C'est autour de cette question que nous animerons un second atelier.

Les 13 participants du premier workshop

-Louis	Aubert	(Start-up	Tracktl)
-Raphael	Lebas	(AËDLE	SAS)
-Pierre	Rudolf	Gerlach	(Qleek)
-Philippe	Le Guern	(Sociologue,	Univ.Nantes)
-Barbara	Chazelle	(Groupe	France TV)
-Reda	Berrehili	(Mixo	TV)
-Catherine	Ramus	(Orange	lab)
-Johanna	Castel	(Cap	digital)
-Edward	Sithol	(Stample)	
-Vincent	Castaignet	(Musicover)	
-Florence	Icard	(Hadopi)	
-Hélène	Merillon	(Youboox)	
-Johanna	Castel	(Cap	digital)



L'écosystème des contenus écrits à 5 ans se présente comme :

«Dis-moi comment tu lis et j'adapterai des contenus à ton style de lecture»

— Le livre qui lit le lecteur ?

Pour cette industrie, un nouveau type de lecture est en train d'apparaître avec des contenus et des formats plus différenciés, ainsi que des systèmes de recommandation. Tout l'enjeu est de garder le lecteur à l'intérieur de l'ouvrage.

◆Les tendances clés:

— Plus grande granularité dans l'analyse des contenus permettant peut-être de sortir de la lecture linéaire : possibilité de proposer des modes de lectures à différentes échelles (extraits choisis selon les goûts de l'utilisateur, trames narratives construites pour le lecteur, ...).

— Plus grande hybridation des supports de lecture (en lien avec les tendances actuelles aux expérimentations autour du hardware).

— Pilotage de la production de certains contenus par les data

— Captation des émotions dans la lecture pour favoriser l'engagement : neuromarketing, nudge, ...

— Création de nouvelles pistes de socialisation autour des contenus écrits, par des lieux ou des services.

◆Les données clés:

— Aujourd'hui: Les données techniques sur la lecture (passages préférés, vitesse de lecture, lecture complète ou non, ...).

— Demain ? Les émotions / réactions dans des contextes plus «transmédias».



L'écosystème des contenus vidéo à 5 ans se présente comme :

«Dis moi ce que tu regardes et je te proposerai des nouveaux contenus immersifs et addictifs.»

— Du binge watching pour tous les contenus ?

— Quête par les acteurs d'un équilibre entre l'expérience de découverte (proposer en permanence de nouveaux contenus) et la profondeur du catalogue (permettre d'accéder à beaucoup de contenus, mêmes anciens). Cette industrie est transformée par une recherche «d'hyperfidélisation».

— Ces phénomènes d'addiction engagent dans la durée le spectateur dans un « flow » de consommation culturelle générant une quantité importante de « data ».

◆Les tendances clés :

— Abaissement des barrières à l'entrée de la production ou de la diffusion (de la gopro à youtube).

— Importance croissante (fanstasmée?) des recommandations... pourtant peu «bluffantes» : l'utilisateur a besoin de guides face à un catalogue foisonnant.

— Rapprochement des médias vidéo et vidéo-ludiques (immersion).

— Déclinaison locale de contenus mondiaux.

◆Les données clés:

— Aujourd'hui : Les goûts et préférences, les habitudes de consommation .

— Demain ? Dans des médias immersifs, les émotions et comportements.

EXPLORATION PROSPECTIVE
DE LA PLACE DES DONNÉES
PERSONNELLES DANS LES
MARCHÉS DES CONTENUS
CULTURELS ET LUDIQUES

CNIL · We design services 

ip@cnil.fr



L'écosystème des contenus musicaux à 5 ans se présente comme :

«Dis moi où et quand tu écoutes et je te proposerai des contenus contextualisés.»

- La consommation musicale passe par beaucoup de zappings, des pratiques «boulimiques», des changements de style, des mécanismes de saturation, ... Tout cela est difficile à réduire à un algorithme. Le «parcours utilisateur» évolue tout au long de la journée avec des modes d'accès à la musique par différents devices, medias, ...
- Il y a un intérêt évident de la part des diffuseurs à comprendre les multiples séquences de la consommation de musique au quotidien. (La consommation musicale, moins accaparante, permet d'accompagner l'utilisateur dans des moments clés du quotidien.)

◆ Les tendances clés :

- Changement de rôle des producteurs/diffuseurs : beaucoup plus de self branding par les artistes, les «majors» deviennent agences de prestation et les consommateurs gagnent en pouvoir de prescription.
- Fonction identitaire très forte de la musique, qui va se matérialiser par la revendication de certains contenus (partage, publication, recherche de visibilité sociale) et à l'opposé par des consommations secrètes, cachées.
- Problème persistant de modèles économiques et de «consentement à payer.»
- Contextualisation du contenu musical en fonction de l'individu, du collectif, d'un partage, d'une compilation...

◆ Les données clés :

- Aujourd'hui : Les goûts, les styles, l'intensité des consommations en fonction des moments (trajets domicile - travail, voyages, ...)
- Demain ? Le contexte via des capteurs/objets connectés (la localisation, la taille du foyer, les modes de vie, les modes de transports...)



Ouverture & Enjeux :

Sur les grandes tendances émergentes, plusieurs pistes de recherche sont à envisager:

- Peut-on imaginer une forme de «monétisation des profils-utilisateurs» sur les différentes plateformes des industries ?
- Comment l'externalisation des acteurs clés - signifiant de moins en moins de grandes compagnies et de plus en plus d'intermédiaires, va-t-elle transformer l'industrie ? Par un aplanissement progressif de l'économie des contenus ?
- Comment le transmédia (narration transversale au sein des différents médias) va-t-il agir sur la consommation des contenus culturels ?
- La recommandation algorithmique va-t-elle vraiment transformer la structuration des contenus narratifs vers une plus grande fidélisation des consommateurs ?

QUE NOUS RÉSERVE LE COUPLE DATA / INDUSTRIES CULTURELLES DANS LES CINQ PROCHAINES ANNÉES ?



Qui était là

Julie Knibbe (Product Manager chez Deezer), **Thibaut D'Orso** (fondateur de Spideo, algorithme de recommandation pour les contenus vidéos), **Jana Moser** (Investment Manager, Axel Springer), **Mael Primet** (Chief Technical Officer et co-fondateur de SNIPS, modèles algorithmiques contextuels et services mobiles personnalisés), **Numa Bourragué** (co-fondateur de Escap, Venture Capital dans le secteur digital), **Thomas Beauvisage** (Research Project Manager chez Orange Labs), **Kevin Mellet** (chercheur chez Orange Labs), **Hubert Saint Olive** (administrateur de Trax Magazine, magazine de référence sur la musique électronique), **Sophie Peinado** (chargée d'étude pour Hadopi), **Antoine Durieux** (CEO de Alkemics, big data appliquée à la grande consommation), **Renaud Viage** (Chief Technical Officer de Eventbrite, plateforme de distribution et de gestion d'événements), **Gaetan Gachet** (Head of Business Development chez Algolia, interface de programmation pour améliorer les moteurs de recherche, "Search-As-A-Service"), **Trista Bridges** (partner chez RudeBaquette, premier support digital anglophone sur l'écosystème technologique française).



MUSIQUE

VERS DES CRÉATIONS ASSISTÉES PAR LA DATA ?

Décortiquer un film pour comprendre les ficelles du succès. La promesse a de quoi séduire les producteurs. Le "génomique" d'un contenu, de sa trame narrative à ses personnages, pourra demain être associé à la réaction de son audience, pour percer les mystères des hits ou des contenus originaux et inspirer la création.

Le service **Hyper Movie Generator**

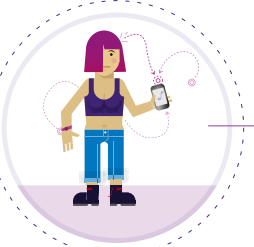
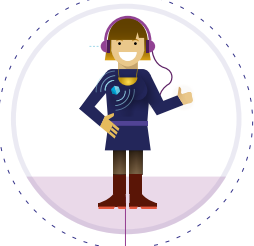
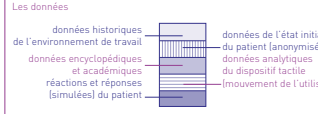
"**Tout le monde peut devenir créateur**" promet l'application. Genre, personnage, sujet ; le système génère des pistes créatives pour créer des films originaux adaptés aux usages et préférences d'un public.



VERS UNE NOUVELLE RÉALITÉ MIXTE.

Le service **Traum Data**

Traum Data est un jeu de réalité augmentée immersif à destination des chirurgiens qui leur propose de tester, d'améliorer et de comparer en continu leur dextérité et leurs aptitudes. Il fonctionne grâce à un dispositif haptique qui capte les données de mouvement des joueurs en temps réel.



VERS UNE NOUVELLE GÉNÉRATION D'ALGORITHMES ?

DES MODELES CONTEXTUELS...

Prévalence des facteurs contextuels... La musique ou le film proposés par les services numériques sera capable de s'adapter à notre humeur du moment et aux données contextuelles qui influencent cette humeur (heure de la journée, météo, etc.).

Le service **The Mood Master**

The Mood Master est un service qui propose des contenus musicaux à l'utilisateur en fonction de son humeur et celle de ses amis, sans effort. Cette technologie d'hypercontextualisation est alimentée par les traces corporelles et physiologiques de l'utilisateur (cardio, pupille, transpiration, etc.), des données géolocalisées enrichies (contexte de l'utilisateur), des données sociales et l'historique de consommation.

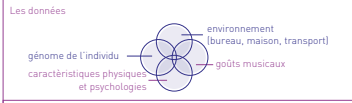
AU PROFIL GÉNÉTIQUE DE L'INDIVIDU

Nous pourrions aller vers un système de recommandation qui délaisse la connaissance a posteriori des comportements sociaux et consommations culturelles au profit d'un **système qui cherche à prédire a priori ces comportements et consommation à partir des critères supposés objectifs de la génétique.**

On passerait donc d'un modèle centré sur la corrélation (si l'utilisateur aime ce produit, il pourrait aimer cette autre produit) à un modèle centré sur la causalité (si le utilisateur aime ce produit, c'est qu'il est programmé pour aimer cette autre offre).

Le service **Life Tracks**

Life Tracks est une application qui associe le profil génétique et l'état émotionnel de l'utilisateur pour lui proposer automatiquement une écoute personnalisée de musique. Ce service se présente sous la forme d'un objet connecté relié à des capteurs ouverts qui parsèment notre environnement. Il permet à l'utilisateur de gagner du temps, réguler son humeur, apaiser des troubles psychologiques, améliorer sa créativité et découvrir constamment de nouveaux morceaux de musique.



DEUX GRANDES CONSTANTES

MODÈLES ÉCONOMIQUES FONDÉS SUR LA PUBLICITÉ - VERS DE NÉCESSAIRES ALTERNATIVES ?

Demain, l'acceptation d'un service et les conditions générales d'utilisation devront aller encore plus loin dans leurs **mécanismes de transparence** pour expliciter concrètement les mécanismes et les données collectées, utilisées et stockées. Plus les données utilisées sont personnelles, plus les modèles ont tendance à refuser la publicité pour ne pas entraver l'acceptation et la qualité du service. L'une des voix qui se dessine privilégie la vente de l'objet connecté, des partenariats avec des services tiers (plateformes) qui pourront ainsi utiliser la technologie (Open API) ainsi que la vente de données sur des thématiques précises, après opt-in et anonymisation (exemple : étude sur la peur dans la ville).

LA QUÊTE DE L'HYPER-PERSONNALISATION, VERS UNE EXPÉRIENCE INVISIBLE ?

L'hyper-personnalisation sera tellement poussée et les technologies tellement bien intégrées à notre réalité (objets connectés, capteurs) que les services numériques se fondront avec l'utilisateur au point de faire disparaître la préoccupation pour l'expérience utilisateur. La fusion parfaite entre le service et l'utilisateur (utilisation des données du génome, la mine de toutes les données, dispositifs haptiques) invite à se poser la question d'un lâcher prise de l'utilisateur. Toute réflexion sur l'usage des données personnelles doit donc identifier un **certain nombre de points de contact** nécessaires pour redonner un minimum d'aspérité au service, pour mieux le comprendre.

A PROPOS

Dans le cadre d'un cycle de recherche collaborative sur le futur des données dans les industries culturelles, l'équipe Innovation et Prospective de la CNIL, et son partenaire Five by Five ont réuni une quinzaine d'acteurs et experts des secteurs de la culture et des technologies le 7 novembre 2014 pour un atelier exploratoire. En ligne de mire, envisager les services des cinq prochaines années et poser les briques de leur modèles économiques. Les quatre services imaginés agissent comme de catalyseurs des questions essentielles.



atelier Prospective #2

