



HAL
open science

Améliorer les contenus de marques grâce aux données émotionnelles de Datakalab : “ promesse communicationnelle ”, discours et imaginaires

Claire Latourre

► To cite this version:

Claire Latourre. Améliorer les contenus de marques grâce aux données émotionnelles de Datakalab : “ promesse communicationnelle ”, discours et imaginaires. Sciences de l’information et de la communication. 2019. dumas-02559455

HAL Id: dumas-02559455

<https://dumas.ccsd.cnrs.fr/dumas-02559455>

Submitted on 30 Apr 2020

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

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

Copyright

Master professionnel

Mention : Information et communication

Spécialité : Communication Marque

Option : Marque et branding

Améliorer les contenus de marques grâce aux données émotionnelles de Datakalab

« Promesse communicationnelle », discours et imaginaires

Responsable de la mention information et communication
Professeure Karine Berthelot-Guiet

Tuteur universitaire : Antoine Bonino

Nom, prénom : LATOURRE Claire

Promotion : 2018-2019

Soutenu le : 12/11/2019

Mention du mémoire : Très bien

REMERCIEMENTS

Ma reconnaissance s'adresse tout d'abord à M. Antoine Bonino, mon rapporteur universitaire du CELSA, qui m'a permis de resserrer le choix de mon sujet, tout en manifestant son intérêt pour ce dernier. J'ai pu grâce à lui passer d'un sujet vaste, le « neuromarketing », à l'étude plus spécifique des données émotionnelles de la start-up Datakalab et à la manière dont elle construit son crédit auprès de ses clients. Il m'a incitée à prendre de la distance sur mon objet de recherche, me donnant suffisamment « d'indices » de la direction à suivre. C'était un plaisir d'annoncer ma sélection à Datakalab après avoir beaucoup travaillé les entretiens.

Mes remerciements s'adressent de manière plus générale au corps enseignant du CELSA.

Un grand merci à François le Corre, mon ancien professeur de « data et marketing relationnel », rapporteur professionnel, et devenu client de Datakalab en juillet 2019. François a immédiatement accepté de m'aider. Il a anticipé les nouveaux enjeux de Datakalab avant même que la structure ne réoriente sa stratégie. Quel heureux hasard de le retrouver sur les différents salons auxquels la start-up participait, comme Big Data 2019 ou le Printemps des Études. J'ai pu également le revoir à Datakalab, mais cette fois-ci en tant que client.

Ce mémoire a été l'occasion pour moi d'aller au-delà des rencontres professionnelles et de tisser des liens encore plus forts avec mes collègues. J'adresse donc mes remerciements les plus sincères à l'ensemble de l'équipe Datakalab.

À Xavier Fischer, CEO, et Lucas Fischer, Chief Data Officer, qui m'ont donné des conseils précieux et appris à être challengée et à challenger durant l'ensemble de mon stage. J'ai pu gagner en confiance grâce à eux. J'admire vraiment la manière dont ils pilotent Datakalab, avec enthousiasme, stratégie, créant ainsi une bonne entente entre les membres de l'équipe. Surtout, ils anticipent et créent constamment les chemins de leur réussite.

Je tiens également à exprimer ma gratitude envers Anne-Marie Gaultier, Ex-Présidente et Soukeina Hassanaly, Ex-Key Account Director, grâce à qui j'ai pu entrer chez Datakalab. C'était aussi un vrai plaisir de travailler avec Frank Tapiro, Chief Emotion Officer et d'échanger avec lui sur de nouvelles idées, notamment dans la rédaction de recommandations pour différentes marques. Soukeina Hassanaly et Aurélien Pecheux, mes deux anciens responsables ont vraiment été bienveillants et m'ont permis de mieux cerner l'environnement

de la start-up. Je remercie aussi Maureen Kahn et Alexandre Amaral pour leur aide dans la sélection de documents pour mon corpus et pour et leur enthousiasme pour ce vaste sujet.

Merci à « l'équipe tech », notamment à Lucas Lugão Guimarães, à Alexandre Macedo et à Éric Daoud qui ont pris le temps de m'expliquer leur métier. Toujours rassurants, j'ai pu compter sur eux. Mes remerciements s'adressent aussi à Kevin Bailly, Arnaud Dapogny et Edouard Yvinec dont les relations ont dépassé l'entente professionnelle. Ils m'ont accordé beaucoup de temps pour m'expliquer certains termes techniques et leurs problématiques de recherches. C'était un plaisir de les retrouver tous les jours et d'échanger sur les différences entre nos domaines d'études.

J'adresse mes pensées à l'ensemble de l'équipe SBT qui a accepté de s'entretenir avec moi pour m'aider à mieux cerner les émotions et leur charge symbolique d'un point de vue psychologique.

Pour finir, je remercie ma camarade du CELSA, Steffi Noël, qui m'a interrogée en tant qu'experte chez Datakalab depuis Shanghai afin de rédiger son article sur les données émotionnelles en Chine. Ensemble, nous avons pu réfléchir à ce sujet ainsi qu'à ses nombreux imaginaires.

ÉVOLUTION ET REPOSITIONNEMENT DE DATAKALAB

Datakalab se présente comme une start-up qui « mesure les émotions des consommateurs à l'aide d'outils issus des neurosciences pour nourrir les stratégies de transformation et d'innovation¹ ». Lors de son lancement, elle opère une véritable révolution en affirmant pouvoir « mesurer les émotions ». La manière dont on construit une « promesse », soit une « proposition sur la communication et ce qu'elle peut apporter à des publics² » et dont on la module en fonction des objectifs retient mon intérêt.

De mars 2019 à août 2019, j'ai effectué mon stage en tant que Project Manager chez Datakalab. Ces six mois ont exactement coïncidé avec la date du repositionnement de la start-up qui a basculé d'une expertise « conseil » à une spécialisation sur sa technologie propriétaire. Dès avril 2019, la direction de Datakalab changeait. Alors qu'Anne-Marie Gaultier, Ex-Directrice générale du Club Med, Ex-Directrice de Communication de Bouygues Telecom et Ex-Directrice Marketing aux Galeries Lafayette quittait la direction pour rejoindre le groupe Intermarché, Xavier Fischer, Centralien, à l'époque Chief Innovation Officer, reprenait les rênes de Datakalab. Mi-avril, lors du changement de direction, la start-up s'est concentrée sur le développement de sa technologie, notamment de sa plateforme et de ses algorithmes propriétaires. Deux nouveaux chercheurs spécialisés en « computer vision³ » avaient été embauchés quelques mois en amont, amorçant cette bascule. Ce changement s'est effectué de manière assez rapide, en l'espace de quelques mois.

Dès lors, j'ai dû repenser l'ensemble de cette étude. Si de par la lecture d'articles et mon embauche, j'avais pu noter plusieurs recommandations, la nouvelle direction les a d'elle-même directement appliquées. J'étais informée de la stratégie de Datakalab, mais j'ai aussi conscience de n'en avoir qu'une vision partielle formée par ce que l'on m'a dit, et par ce que j'ai pu entendre lors d'appels comme de discussions stratégiques.

¹ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, page « ACCUEIL » de Datakalab, « disponible à l'adresse : <https://www.datakalab.com/> »

² Jeanneret, Yves, *Critique de la Trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, Lexique, p.14

³ Padiolleau, Agathe, 26/06/2018, « La vision par ordinateur ou vision artificielle, est une sous-catégorie de l'Intelligence artificielle. Cet ensemble de méthodes et de technologies permet l'automatisation d'une tâche spécifique à partir d'une image »

Suite à la « bascule », l'un des outils de Datakalab, le « bracelet connecté », a cessé d'être utilisé. Jusqu'en mars-avril 2019, il servait à « mesurer les signes physiologiques de l'émotion⁴ » en point de vente. L'outil permettait de combiner des « données brutes » comme la « température extra-cutanée », la « conductance électrodermale », le « battement cardiaque », à dessein de générer un score « d'intensité émotionnelle ». Si le cœur de ma recherche se concentre sur les promesses autour du « facial coding », ce bracelet reste un outil inouï, je tiens donc à lui réserver une place ci-dessous.

⁴ ANNEXE 18, Rubrique « nos outils », *NB* : L'outil « bracelet connecté » a été supprimé du site le 3 juin 2019

SOMMAIRE

REMERCIEMENTS	3
ÉVOLUTION ET REPOSITIONNEMENT DE DATAKALAB	5
SOMMAIRE	7
NOTE INTRODUCTIVE	9
INTRODUCTION	10
1 UNE LÉGITIMITÉ DE LA « FEEL DATA© » CONSTRUITE SUR UN RENOUVEAU TECHNOLOGIQUE ET SUR DES IMAGINAIRES AUTOUR DU NEUROMARKETING	22
1.1 LES DONNEES EMOTIONNELLES SONT ISSUES ET PORTEES PAR PLUSIEURS TENDANCES	22
1.1.1 « <i>Feel data©</i> » fille de deux tendances : « <i>big data</i> » et <i>émotion</i>	24
1.1.1.1 Le « <i>big data</i> » comme système de croyance	24
1.1.1.2 Un nouveau technologique permet l'avènement des données émotionnelles	25
1.1.1.3 Datakalab, un nom chargé d'imaginaires	26
1.1.1.4 L'émotion : un concept flou, évolutif et complexe	27
1.2 DATAKALAB PROMET DE « MESURER LES EMOTIONS », SOIT DE RATIONNALISER L'IRRATIONNEL	28
1.2.1 « <i>Redonner visage humain à la data</i> »	31
1.2.2 <i>Revendication d'un savoir neuroscientifique</i>	32
1.2.2.1 Construction d'un « <i>ethos neuroscientifique</i> »	32
1.2.3 <i>Datakalab se pose en s'opposant aux études déclaratives</i>	34
1.2.3.1 Le neuromarketing comme clé d'accès à l'inconscient	35
1.2.3.2 Un objet complexe, protéiforme et fantasmé	36
1.2.3.3 Ambivalence du neuromarketing : cadre réglementaire et éthique	37
1.2.3.4 Les émotions au cœur du processus décisionnel	38
1.3 TRANSFORMER UNE DONNEE BRUTE EN REALITE INTELLIGIBLE	39
1.3.1 <i>Traductions du réel et simplifications par des formes</i>	39
1.3.1.1 « <i>Scripturer</i> » la donnée pour la rendre péremptoire	40
1.3.1.2 Traductions du langage ingénieur au langage commercial	42
1.3.1.3 Datakalab opère un « <i>acte sémiurgique</i> »	44
1.3.1.4 Une « <i>mesure</i> » complexe des émotions	45
2 ÉVOLUTION DES « DISCOURS D'ESCORTE » DE LA « FEEL DATA© »	47
2.1 MALGRE LA SUPPRESSION DE L'UN DE SES OUTILS, DATAKALAB CONSERVE SA PROMESSE MAÏEUTIQUE	47
2.2 CONSERVATION D'UN TRIPLE POUVOIR HEURISTIQUE PRAGMATIQUE ET NORMATIF	48
2.3 ÉVOLUTION DES « DISCOURS D'ESCORTE » DE LA « FEEL DATA© »	49
2.3.1 <i>Réduction de la compétence humaine d'interprétation</i>	51
2.3.2 <i>Des « algorithmes au-delà du réel » : construction d'une deuxième posture d'expert</i>	52
2.3.3 <i>Datakalab révélée par la création d'algorithmes propriétaires de « facial coding »</i>	54
2.3.4 <i>Valorisation de l'imaginaire scientifique dans la rhétorique commerciale</i>	55
2.4 MAITRISER LE REEL VIA LA « DATA »	56
2.4.1.1 Manque de recul sur la « <i>raison algorithmique</i> »	58
2.4.1.2 Fonction incantatoire des technologies Datakalab	59
3 DATAKALAB APPLIQUE SA NOUVELLE VISION ET S'INTÉRESSE À DE NOUVEAUX MARCHÉS	61
3.1.1.1 Un travail de pédagogie et de vulgarisation autour d'une vérité scientifique	61
3.1.2 <i>Adaptation des « métriques » pour accroître leur opérativité</i>	62
3.2 DATAKALAB, LES INSTITUTS D'ETUDES ET LES ANNONCEURS NOURRISSENT ENSEMBLE LEURS STRATEGIES D'INNOVATION	67
3.2.1 <i>Les appropriations de la « feel data © » témoignent de son statut « d'indice » d'appréhension des « réactions »</i>	69
3.2.1.1 Une approche descriptive et non analytique de la presse liée à une méconnaissance technique	69
3.3 UNE « RATIONALITE OCCIDENTALE » DU NOMBRE, DU CALCUL ET DE LA DONNEE PORTE LES DONNEES EMOTIONNELLES	70

3.3.1	« Gouvernamentalité algorithmique » : la « data » fait loi	71
3.3.2	L'innovation technologique au cœur de la rhétorique commerciale	72
3.3.3	Double instance énonciative ingénieur/communicant : vendre une recherche	73
3.4	RECOMMANDATIONS	76
3.4.1	Datakalab doit continuer d'appliquer sa nouvelle expertise	76
3.4.1.1	Segmentation hyper personnalisée permise par les agences et régies.....	76
3.4.2	Continuer à théâtraliser son expertise tout en rassurant	78
3.4.3	De multiples applications de la donnée émotionnelle témoignent d'un marché en plein « boom »	80
3.4.4	Datakalab doit s'appuyer sur l'aspect « humain »	80
3.4.4.1	Une accélération du développement de Datakalab via un développement à l'échelle internationale	82
	CONCLUSION	84
	BIBLIOGRAPHIE	89
	RÉSUMÉ	97
	MOTS CLÉS	98
	ANNEXE	99

NOTE INTRODUCTIVE

La première difficulté de cette recherche a été de prendre de la distance sur son objet, dans la mesure où celui-ci évoluait et où je contribuais à sa création. En tant que Project Manager, j'étais en charge des études :

- Je les lançais et collectais les données. Je réalisais les présentations et rédigeais les recommandations pour optimiser différents contenus de marques : événements, conférences, affiches, spots puis je les livrais aux clients comme Exterion Media.
- De réflexions à la stratégie de Datakalab et à l'appellation des « métriques », je rédigeais également des études de concurrence et de positionnement.
- En tant qu'administratrice des pages LinkedIn et Facebook de Datakalab, je rédigeais les « posts », tout comme les « newsletters » destinées aux 2 000 abonnés.
- J'organisais et participais aux salons comme Big Data 2019, la Nuit de L'IA, les 20 ans du cabinet de conseil Escal Consulting et le Printemps des Études.

Par ailleurs, mon ancien professeur du CELSA François le Corre, que j'avais expressément choisi comme étant extérieur à Datakalab, travaille également chez DBI, une des agences de Havas en tant que Chief Operating Officer. Dans ce cadre, il est également devenu client de Datakalab dès juillet 2019.

En outre l'une de mes camarades en Marque & Branding, Steffi Noël, m'a interviewée en tant qu'experte chez Datakalab pour rédiger un article sur les « données émotionnelles » en Chine.

Je me suis donc retrouvée dans une situation contradictoire où il fallait à la fois se familiariser avec de nouvelles compétences, s'intégrer à la dynamique de Datakalab et prendre de la distance sur cet objet de recherche. Ce travail va chercher à déconstruire ce que je tenais pour acquis à la fois avant et après cette expérience.

INTRODUCTION

Datakalab promet de « mesurer les émotions⁵ » des consommateurs en les captant directement à leur source, l'inconscient, via des traitements par la machine. En éliminant l'étape de la verbalisation, elle assure à ses clients une mesure « non biaisée⁶ » des émotions face à des contenus de marques. En d'autres termes, elle promet d'aider les producteurs de contenus à les optimiser pour qu'ils puissent attirer l'attention, créer de l'émotion et que cette émotion soit positive. On peut lire sur son site dans la rubrique « Notre promesse⁷ » : « Datakalab mesure les émotions des clients et les transforme en données émotionnelles. Datakalab croise les data transactionnelles, digitales ou relationnelles pour les enrichir avec les feel data© et identifier des corrélations vertueuses et créatrices de valeur. Datakalab visualise les données grâce à sa plateforme et intègre les résultats forte de son expertise marketing et neuroscientifique ». La start-up promet donc d'améliorer les stratégies des créatifs pour augmenter la performance de différents contenus, ce qui institue les « données émotionnelles » comme outils d'aide à la création.

Comme l'indique Google Trends⁸, les recherches « big data », « machine learning », « data science », « data scientist » sont de plus en plus fréquentes. Elles témoignent de l'entrée dans la nouvelle « ère des données ». Ces « data » sont devenues l'or noir du marketing, dont les techniques sont constamment renouvelées. Si le ce dernier analyse les besoins des consommateurs et tente d'influencer leur comportement, nos réactions sont aussi étudiées par les neurosciences, issues de la biologie et de la médecine. Elles se focalisent quant à elles sur les mécanismes chimiques du cerveau qui impactent notre attitude⁹. À la jonction de ces deux domaines, une discipline peu connue en France a été construite au début des années 2000¹⁰, le « neuromarketing ». Idriss Aberkane, neuroscientifique diplômé de l'École Normale Supérieure, le définit comme suit : « Le neuromarketing tente de comprendre comment le cerveau décide s'il achète ou non et comment on associe des émotions au produit ». Les recherches de ce terme

⁵ ANNEXE 20, support de présentation PowerPoint de Datakalab, slide « Qui sommes-nous ? » : « Nous mesurons les émotions à l'aide d'outils issus des neurosciences »

⁶ *Ibid.* Slide « Nos convictions »

⁷ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, page « LE LAB » « disponible à l'adresse : <https://www.datakalab.com/lab> »

⁸ ANNEXE 17, Google Trends, recherches dans le monde de 2004 au 05/09/2019

⁹ Dictionnaire Larousse, « Neurosciences : ensemble des disciplines étudiant le système nerveux », « disponible à l'adresse : <https://www.larousse.fr/dictionnaires/francais/neurosciences/54418> »

¹⁰ Glossaire « Neuromarketing », Emarketing.fr, « disponible à l'adresse : <https://www.emarketing.fr/Definitions-Glossaire/neuromarketing-254948.htm> »

dans Google ont quasiment doublé de 2005 à 2015 selon Google Trends¹¹. Aujourd'hui, les outils issus des neurosciences comme « l'eye tracking » (qui permet de trouver les zones qui attirent le regard), le « facial coding » (qui « mesure l'émotion » en détectant les « expressions faciales ») sont utilisés à des fins marketing pour spécifier la donnée classique, transactionnelle, comme le « panier moyen », le « taux de clics », le « nombre de visites », le « temps passé par page ». Intéressée depuis toute jeune par la psychologie, et plus particulièrement par les liens entre marques et consommateurs, je me suis focalisée sur ce sujet. Néanmoins, pour pouvoir prendre de la distance sur ce phénomène assez récent et complexe, le neuromarketing n'est pas étudié ici *en soi*. C'est bien la manière dont les professionnels de la communication et en particulier la start-up Datakalab, s'appuient sur ses outils pour construire leur crédit qui retient mon intérêt.

Le neuromarketing prétend optimiser les décisions grâce à un savoir scientifique et psychologique. Différents articles de presse professionnelle parus dans LSA¹², Stratégies¹³, l'instaurent comme « avenir du marketing classique ». Ce « marketing classique » regroupe toutes les techniques utilisées par les professionnels de la communication qui ont fait leurs preuves comme les études déclaratives, les indicateurs de suivi, les données transactionnelles évoquées plus haut. « Électroencéphalogramme », « facial coding » « eye tracking » ... en constatant l'étendue de ces termes techniques, innovants et prometteurs, je réalise l'immensité de l'objet. Dans différents reportages, la start-up Datakalab est mentionnée.

Datakalab affirme « mesurer les émotions » sur tous les « touchpoints » de la marque. Par ce terme, elle entend ses points de contact comme des sites internet, des vidéos, des spots, des affiches. La promesse peut fasciner puisqu'elle compte tout bonnement quantifier et rationaliser l'émotionnel. En mars 2016 Frank Tapiro, l'un des membres fondateurs et Chief Emotion Officer¹⁴ de Datakalab dépose le terme de « feel data© » qu'il a inventé. Bertrand Barthelot, professeur agrégé de marketing à HEC (extérieur à la start-up), les définit en décembre 2016 comme « données relatives à un consommateur ou groupe de consommateurs qui peuvent être collectées et traitées de façon à traduire les émotions ressenties par cet individu ou ce groupe ». Les « feel data© » veulent donc capter les émotions humaines, les réactions

¹¹ *Ibid.*

¹² Gavard, Emmanuel, « Le fantasme du neuromarketing » LSA, 17/05/2012, « disponible à l'adresse : <https://www.lsa-conso.fr/le-fantasme-du-neuromarketing,130326> »

¹³ Marchal, Cyril, « Le neuromarketing, arme de séduction massive », 12/05/2015, « disponible à l'adresse : <http://www.strategies.fr/blogs-opinions/idees-tribunes/1014611W/le-neuromarketing-arme-de-seduction-massive.html> »

¹⁴ Jeu sur les initiales « CEO », Frank Tapiro dirige de la communication de Datakalab autour des émotions

face à des stimuli. Jusqu'à début 2019, Datakalab captait ces données grâce à des outils issus des neurosciences comme le « bracelet connecté » et le « facial coding¹⁵ ». La start-up ne vend plus aujourd'hui d'étude avec le « bracelet connecté », leur mise en place s'avérant trop complexe et coûteuse. Datakalab s'est donc spécialisée dans le « facial coding », qui entend « mesurer les émotions » humaines grâce à l'analyse des « expressions faciales¹⁶ ». C'est ainsi que l'inconscient serait le terrain de récolte de nouveaux types de données.

Dans sa rhétorique commerciale, Datakalab affirme que la mesure directe des émotions permet de ne pas déformer le ressenti en supprimant l'étape de la verbalisation, à la différence du déclaratif. Si lors du lancement des « feel data© » en 2016, Datakalab les positionnait comme concurrentes des données déclaratives, les objectifs et la cible de la start-up ont évolué, et ce discours a été modifié. Aujourd'hui, elle affirme vouloir compléter les « données transactionnelles » par les données émotionnelles. La start-up travaille d'ailleurs avec des instituts (comme OpinionWay¹⁷) qui traitent parfois la partie déclarative des études émotionnelles. Ces données déclaratives et émotionnelles promettent, une fois combinées, d'obtenir des vérités clients plus précises pour améliorer les créations publicitaires, les événements liés à la marque. Cependant, il reste difficile pour ces données de trouver leur place face à la notoriété des données déclaratives, majoritairement utilisées par les annonceurs et qui ont déjà fait leurs preuves.

Mon intérêt pour ce sujet dépassant le cadre de cette recherche et m'attirant comme spécialisation professionnelle, je tente alors d'intégrer la start-up. Pour me préparer aux entretiens, j'écoute des conférences. Je vais jusqu'à remonter l'ensemble des articles Datakalab sur les différents réseaux sociaux. Dans ses premières prises de parole en 2016-2017, Datakalab partage les articles des médias¹⁸ qui la mentionnent : Le Figaro, Les Échos, Le Monde... Elle se positionne et convainc de plus en plus d'annonceurs prestigieux¹⁹ comme Disney, SNCF, Novotel, Alfa Roméo. La presse répand rapidement l'information. Lors de son interview à BFM

¹⁵ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, rubrique « Nos outils », NB : L'outil « bracelet connecté » a été supprimé du site le 3 juin 2019

¹⁶ ANNEXE 20, *Op. Cit.*, slide « Facial coding »

¹⁷ ANNEXE 19, Photos de la conférence « OpinionWay x Datakalab », Printemps des Études, 04/2019

¹⁸ ANNEXE 22, Sélection de prise de parole Facebook, articles de 2016 postés dans la rubrique « Ils parlent de nous »

¹⁹ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, Page « LES MISSIONS », « disponible à l'adresse <https://www.datakalab.com/missions> »

TV du 9 octobre 2018²⁰, Anne-Marie Gaultier mentionne même un partenariat avec IBM Tealeaf²¹. Stupéfaite par cette large couverture médiatique, je contacte Datakalab et suis reçue pour un premier entretien. On m’informe que le rôle du Project Manager consiste à réaliser l’entièreté des études, récolte de données, analyse, envoi du livrable au client. Suite à ma rencontre avec Anne-Marie Gaultier, je suis sélectionnée pour rejoindre la start-up. Ce sujet plus ciblé se dessine alors, ayant désormais accès à un terrain très conséquent :

**Améliorer les contenus de marques grâce aux « données émotionnelles » de Datakalab :
« promesse communicationnelle », discours et imaginaires**

Ce sujet fait intervenir une multitude de dynamiques. Imbriquées les unes dans les autres, technique, politique et société le sous-tendent. L’étude veut analyser les « discours d’escorte²² », soit les récits qui accompagnent les données émotionnelles, pour comprendre comment ces objets agissent sur les entreprises. Je m’interroge sur la viabilité des techniques utilisées et sur la réelle possibilité de « mesurer l’émotion ». La construction d’un récit pour convaincre, faire vendre et fidéliser constitue le cœur de ma recherche. Je souhaite identifier, problématiser les promesses des données émotionnelles, leurs discours d’escorte auprès des annonceurs et des instituts d’études. J’étudierai donc la manière dont Datakalab prétend intervenir sur les décisions des acteurs de la communication, son expertise, la circulation de ses outils et de ses résultats.

Il existe de nombreuses sociétés de neuromarketing basées aux États-Unis (Affectiva, Kairos), en Chine (FaceThink, Seeta Tech), en Europe (Immersion Neuroscience, Spark Neuro, IBM Watson, Noldus, Eye Square, Sticky, Crowd Emotions, Real Eyes). Les concurrents de Datakalab utilisent les « feel data© », qu’ils qualifient parfois de « emotional data », soit de données émotionnelles. Les « feel data© » existent à l’étranger, mais le terme n’est déposé en

²⁰ BFM TV, « Business Transformation : place aux data émotionnelles ! », 09/10/2018, « disponible à l'adresse : <https://bfmbusiness.bfmtv.com/mediaplayer/video/business-transformation-place-aux-data-emotionnelles-0910-1107689.html> »

²¹ Solution d’analyse d’IBM pour web et mobile. L’outil permet de capter les interactions des utilisateurs sur l’interface comme le déplacement de la souris

²² Jeanne-Perrier, Valérie, *Communication & langages*, 2015, N° 183, pages 111-131, « Des discours et des signes : pour une évaluation critique de Twitter et de Facebook » : « La notion de "discours d’accompagnement", forgée par le chercheur Philippe Breton à l’occasion de la sortie de son ouvrage *Le culte d’internet* s’arrime à cette idée que les usagers, comme des bambins, ont besoin de bras et de supports chaleureux pour les aider à se lever, se mettre en route, pour avancer ensuite seuls puis revenir vers leurs facilitateurs des « premiers pas ». Ceci afin de les remercier d’avoir joué ce rôle de tuteurs et d’en devenir alors les fils et les filles prodiges, démultipliant la présence des marques impliquées par leurs usages et leurs discours »

France que depuis 2016 par Frank Tapiro. En Europe, le sujet reste relativement tabou puisqu'il soulève des questionnements éthiques majeurs : dans quelle mesure peut-on travailler sur la donnée émotionnelle en tant que service, et dans quel genre de futur s'engage-t-on à vivre avec ce nouveau type de données ? En Chine par exemple, la réglementation n'est pas la même vis-à-vis des données personnelles des citoyens. Elles sont utilisées par le gouvernement pour maintenir l'ordre²³, et les sociétés de neuromarketing sont de plus en plus nombreuses. Comme l'explique Lucas Fischer²⁴, aux États-Unis, les citoyens prennent conscience des dangers relatifs à l'utilisation de leurs données personnelles. Ce terme est chargé d'imaginaires relatifs à la manipulation, il peut même être considéré comme doublement à charge, le préfixe « neuro » mettant l'accent sur le fait qu'on « entre dans le cerveau des gens » et le terme « marketing » étant associé au « faire vendre » en manipulant.

Du côté entreprise, le neuromarketing est considéré comme un investissement risqué, nouveau, inconnu, voire inquiétant. Les professionnels disposent déjà de toute une boîte à outils conséquente pour analyser eux-mêmes les comportements consommateurs. S'ils décident de faire des études, ils y consacrent un certain budget. Leur choix se porte instinctivement sur de grands instituts qui ont déjà fait leurs preuves. Dès que les annonceurs souhaitent faire une étude consommateurs, leurs noms leur viennent spontanément à l'esprit. Avant son repositionnement, Datakalab réalisait des études pour les annonceurs²⁵, et devait d'abord faire face à de la concurrence indirecte, soit aux instituts tels que Nielsen, Ipsos, TNS Sofres, BVA. Néanmoins, suite à sa « bascule », Datakalab a choisi de ne pas chercher à remplacer les instituts d'études, mais au contraire à établir des partenariats avec eux.

Comme en témoignent les articles de Stratégies²⁶, Les Echos²⁷, la large couverture médiatique lui a permis d'être connue des principaux annonceurs. Datakalab doit recruter de

²³ CentraleSupélec « Datakalab : l'analyse des émotions », vidéo postée le 29/03/2018, « disponible à l'adresse : https://youtu.be/FBknC_FvkFc »

²⁴ ANNEXE 7, Entretien avec Lucas Fischer, Chief Data Officer chez Datakalab, 08/07/2019, Paris 17^{ème}

²⁵ *Ibid.* « On s'est rendu compte que les » ²⁶ Soulas-Gesson, Delphine, Stratégies, 15/03/2019 « Disponible à l'adresse : <http://www.strategies.fr/actualites/medias/4026275W/exterion-media-s-allie-avec-datakalab-sur-la-mesure-des-emotions.html> »

²⁶ Soulas-Gesson, Delphine, Stratégies, 15/03/2019 « Disponible à l'adresse : <http://www.strategies.fr/actualites/medias/4026275W/exterion-media-s-allie-avec-datakalab-sur-la-mesure-des-emotions.html> »

²⁷ Loye, Deborah, Les Echos, « CES 2019: Datakalab transforme les émotions en données » « Disponible à l'adresse : <https://www.lesechos.fr/tech-medias/hightech/0600478877092-ces-2019-datakalab-transforme-les-emotions-en-donnees-2234962.php> », 09/01/ 2019

nouveaux clients qui sont prêts à signer des abonnements²⁸, et défendre sa légitimité par un argumentaire qui doit à la fois respecter la confidentialité des études déjà réalisées, se démarquer *via* le positionnement, la technologie et les innovations. Leur irruption est par essence troublante et ne permet pas de rassurer, à la différence des grands groupes précédemment cités. Elle doit donc tenir un discours vulgarisé mais expert, impactant mais bienveillant, rassurant mais innovant.

Malgré les difficultés évoquées, de nombreux dispositifs rendent possible la croyance en ce nouveau type de données. La « feel data© » ne peut être appréhendée sans considérer l'effet de mode autour du « big data ». Si ce dernier traduit la nouvelle « ère des données », il est devenu un « buzzword²⁹ » à partir de 2012, soit un terme à la mode, employé dès que l'occasion s'y prête pour légitimer un discours. « Mesurer les émotions » sur tous les « touchpoints », les points de contact de la marque, constitue donc une « promesse communicationnelle » innovante. Elle séduit et différencie Datakalab des instituts d'études. Promettre une telle mesure revient à affirmer une possibilité de captation par la machine des émotions. Ces dernières sont considérées comme typiquement humaines. Datakalab construit un argumentaire pour légitimer cette promesse surprenante. L'enjeu consiste donc à cerner, définir et problématiser ces « prétentions communicationnelles³⁰ » tout en comprenant comment Datakalab peut leur créer une place dans la boîte à outils déjà conséquente des professionnels de la communication, influencer leurs décisions marketing, voire leurs résultats.

Par quel processus de déconstruction et de reconstruction Datakalab transforme-t-elle l'émotion en donnée, puis la donnée en émotion ? Existe-t-il une dénaturation de l'émotion dans ce processus ? Jusqu'où se développera alors cet intérêt pour la « data » ? Je souhaite alors comprendre si les « feel data© » constituent juste un outil de plus dans le panel du professionnel, ou si, passé les effets de mode et de nouveauté, elles promettent des avancées dans d'autres domaines. Les « promesses communicationnelles » de Datakalab continuent de se définir à l'heure actuelle, cette mobilité devient un autre moteur de travail. Ma problématique formulée

²⁸ ANNEXE 8, Entretien avec Maureen Kahn, Project Manager chez Datakalab, 02/07/ 2019, Paris 17^{ème}, « Il faut vraiment se concentrer sur des clients qui sont prêts à signer des abonnements et perdre moins de temps sur des clients moins importants »

²⁹ ANNEXE 17, Google Trends, recherches dans le monde de 2004 au 05/09/2019

³⁰ Jeanneret, Yves, *Critique de la Trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, Lexique, p.14 : « mode d'intervention sur les processus de communication qui de manière délibérée ou insensible hiérarchise les éléments, détermine des conditions cruciales et légitime une certaines compétences à y intervenir »

plus bas s'appuie alors sur la « bascule » de la start-up pour pouvoir étudier à la fois les discours de l'ancienne expertise « conseil » et de la nouvelle, « technologique ».

« Détecter les émotions » semble être une compétence humaine, qui ne peut pas être rattachée à la machine. Un « algorithme » capable de les mesurer s'avère contre-intuitif, et si innovant que l'on peut douter de sa légitimité. De même si les données émotionnelles se revendiquent comme outil d'aide à la création, le travail du créatif est rattaché à un savoir typiquement humain, irrationnel, indénombrable. Or Datakalab mise ici sur la technologie, sur les données issues d'un algorithme, d'une machine rationnelle pour aider l'homme dans le processus créatif d'optimisation d'un contenu.

La start-up rencontre alors des difficultés pour faire passer ses données émotionnelles comme nouvelles normes. Les tendances comme le neuromarketing peuvent rapidement lui nuire de par les craintes éthiques des acteurs professionnels. Toute l'innovation promise par la start-up joue aussi contre elle, dans la mesure où ce qui est innovant est par essence non établi, non rassurant. L'enjeu consiste alors à établir comme normes des données jusqu'alors inconnues. Datakalab part elle-même « d'indices³¹ » de l'émotion que sont les données brutes comme la « température extra-dermale » captée par le « bracelet connecté ». Une fois captées par leurs outils, ces « traces³² » doivent être mises en forme et être interprétées par les communicants de Datakalab. C'est seulement à terme de ce processus de construction de sens qu'elles sont soumises aux clients en tant que données émotionnelles. Or sur son site³³ et dans ses présentations³⁴, Datakalab instaure ses données comme naturelles. Elles doivent aller de soi pour les clients, alors qu'elles relèvent d'un construit. Plusieurs éléments y sont imbriqués, humainement agencés et le tout est traversé de médiations. Comment alors les instituer à nouveau en tant « qu'indices³⁵ », mais cette fois-ci de la réception, du sens que les consommateurs tirent des contenus auxquels on les expose ? Datakalab doit alors donner les clés d'interprétation de ses données, alors qu'elle est la seule à les détenir. Il existe donc une rupture entre cette lourde construction de sens et l'évidence prétendue de la donnée émotionnelle :

³¹ Yves Jeanneret définit l'indice comme un « simple concept technique » qui relève du monde de la « causalité », 2011, p.76

³² À la lumière du travail d'Yves Jeanneret, je définirai et caractériserai les « traces » captées que Datakalab construit et valorise comme « données émotionnelles »

³³ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, Page « ACCUEIL » de Datakalab « disponible à l'adresse : <https://www.datakalab.com/> »

³⁴ ANNEXE 20, Support de présentation PowerPoint de Datakalab, slide « Qui sommes-nous ? »

³⁵ *Op. Cit.*

Dans quelle mesure la start-up Datakalab s’efforce-t-elle d’instituer ses données émotionnelles comme « indices » capables de « mesurer les émotions » face à des contenus de marques ?

Par le terme « s’efforce », je me penche ici sur les difficultés que Datakalab peut rencontrer. Elle doit notamment recruter un panel, parler à une cible – les instituts d’études – qui doit elle-même en convaincre une autre – les annonceurs... Parmi toutes ces difficultés, il faut garder l’esprit critique et pouvoir prendre de la distance sur ces pratiques nouvelles, innovantes, accompagnées de discours – anciens et nouveaux – et qui font peur.

HYPOTHÈSES

J’ai dégagé ces hypothèses à la lumière des discours de Datakalab et des textes de références dans les sciences de l’information et de la communication.

- 1. Datakalab convainc ses prospects qu’il est possible de ramener la captation et la « mesure de l’émotion » à un traitement computationnel.**
- 2. Datakalab maintient sa « promesse communicationnelle » de « mesure des émotions » malgré son repositionnement.**
- 3. La rhétorique commerciale modulée en fonction des besoins des nouveaux clients accroît l’opérativité sociale et symbolique des « feel data© ».**

MÉTHODOLOGIE & CORPUS

Neuromarketing, données, algorithmes, émotions... autant de concepts entourés d'imaginaires et utilisés par une pluralité d'acteurs, ce qui traduit une véritable fascination à leur égard. Malgré l'ensemble de ces discours, on constate un manque de recul réflexif et un flou définitionnel autour de ces termes. Si les « data » font loi, données et émotions sont-elles compatibles ? Leur combinaison – souvent placée sous l'égide du neuromarketing – peut-elle faire autorité auprès des décideurs ? Ce sujet entremêle une pluralité d'objets, imbriqués avec des liens de cause à effet les uns dans les autres. Cherchant comme Barthes le propose³⁶ : « une réconciliation [...] de la description et de l'explication, de l'objet et du savoir », pour chacun des objets étudiés, j'ai travaillé à partir de la logique des quatre catégories de médiation identifiées par Yves Jeanneret qui ont lieu dans la communication médiatisée : « celle du projet d'intervention dans la relation (promettre quoi à qui ?) » ; celle de « l'énonciation et du discours comme matière travaillée » ; celle de « l'industrialisation de la communication comme un conditionnement » ; celle de « l'action sur les corps et le faire-faire³⁷ ». Tout d'abord, mes recherches ont été en majorité menées pour répondre à la fois à mes trois hypothèses. J'ai ainsi effectué une observation *in situ* de type ethnographique et me suis concentrée sur les éléments qui ont été à ma disposition durant mon stage comme les envois aux annonceurs. J'ai pu rencontrer un grand nombre d'acteurs : experts en marketing et en communication, ingénieurs, psychologues. Pour comprendre les discours d'escorte autour de la donnée émotionnelle, je me suis adressée à des professionnels du marketing et de la communication. J'ai également interrogé des experts en sciences cognitives. Un guide d'entretien a été enrichi suite aux différentes rencontres. Douze entretiens qualitatifs semi-directifs ont ainsi été menés. Ma méthode de recherche s'est basée sur un raisonnement itératif et abductif. Des experts aux profils différents ont ainsi été interrogés entre avril et juillet 2019 à *The Brain Place To Be*³⁸, à Paris dans le 17^{ème} arrondissement. Dans ce lieu de référence où se regroupent les experts en neurosciences, trois sociétés sont présentes : Datakalab, Unyck et SBT. J'ai pu interroger

³⁶ Barthes, Roland, *Mythologies*, 1957

³⁷ Bonaccorsi, Julia, « Compte rendu de Critique de la trivialité. Les médiations de la communication, enjeu de pouvoir », *Communication & langages*, 2015/3, N° 185, pages 31-35 « disponible à l'adresse : <https://www.cairn.info/revue-communication-et-langages1-2015-3-page-31.htm> »

³⁸ The Brain Place To Be, « premier centre dédié aux neurosciences cognitives, un lieu d'échanges pluridisciplinaires, de veille et de prospective pour transformer les usages des organisations, des marques et des dirigeants par l'apport des neurosciences cognitives », « disponible à l'adresse : <https://www.thebrainplacetobe.com/> »

l'équipe de direction marketing de Datakalab composée des membres fondateurs. Nous avons évoqué à la fois les origines de la start-up et les nouvelles applications qu'ils envisageaient pour les données émotionnelles. L'entretien avec « l'équipe tech » de Datakalab m'a permis de prendre de la distance sur les termes employés dans la rhétorique commerciale. J'ai également interrogé deux psychologues de chez SBT sur les façons de « capturer les émotions ». Chaque entretien a fait l'objet d'une prise de notes et a été enregistré, puis retranscrit, avec le consentement des interviewés. J'ai ensuite procédé à l'analyse de ces entretiens³⁹, en faisant ressortir les principaux enjeux évoqués.

Pour rappeler ma première hypothèse, Datakalab affirme qu'il est possible de ramener la « captation » et la « mesure de l'émotion » à un traitement computationnel. Ici je m'appuie sur les éléments identifiés par Yves Jeanneret dans *Critique de la Trivialité* pour illustrer les « prétentions communicationnelles » des données émotionnelles. J'ai d'abord cherché à cerner les prétentions du neuromarketing en lisant des articles professionnels et études « neuroscientifiques ». Cela m'a fait comprendre la diversité des méthodes proposées pour « mesurer l'émotion » et les discours pour convaincre qu'il est possible de traiter l'émotion, irrationnelle, avec des algorithmes propres à la machine, rationnels. Paul Ekman a par exemple théorisé l'universalité des émotions et Antonio Damasio les a définies comme ancres mémoriels, nécessaires à la décision. Lors de discussions avec les professionnels de Datakalab, autour de ces théories en sciences cognitives, j'ai pu prendre de la distance sur ces dernières.

Le nom de la start-up véhicule lui-même des imaginaires autour des neurosciences, de l'innovation et du monde scientifique, technique et technologique. Je mène alors une analyse sémiotique du nom Datakalab. Pour appréhender la construction des données émotionnelles sur un nouveau contexte sociétal et technologique, les « feel data© » sont analysées à partir de la définition de Datakalab à l'origine de cette appellation, mais aussi selon différents articles issus de la presse professionnelle. Les entretiens permettent de connaître le positionnement des professionnels sur leur travail.

Je suppose ensuite que malgré sa bascule, Datakalab maintient sa « promesse communicationnelle » de mesure des émotions. Pour vérifier l'hypothèse, j'analyse le site de Datakalab, dont j'ai recueilli des captures d'écran. De même, à travers une sélection de

³⁹ ANNEXE 3, Résultats de l'analyse terrain

« slides », j'indique que les « discours d'escorte » de la donnée émotionnelle ont évolué. Je me penche sur l'évolution de la communication sur la période de mars à août 2019 et sur des matériaux de différentes formes : emailings, présentations Powerpoint, images, graphes, vidéos, retranscriptions de « pitch ». Les imaginaires jouent sur le chiffre, sur la performance. En ce qui concerne ces discours autour du nombre et du calcul, je m'appuie sur « la logique du Père Noël⁴⁰ » développée par Baudrillard dans *Le Système des objets*. L'auteur explique qu'on choisit de croire ou non aux grandes tendances et représentations sociales. Datakalab s'inscrit alors dans une monstration de ses algorithmes.

Pour finir, je suppose que la rhétorique commerciale modulée en fonction des besoins des nouveaux clients accroît l'opérativité sociale et symbolique⁴¹ des « feel data© », soit la manière dont elles font effet, ici en se rattachant à des imaginaires liés à la société et aux enjeux politiques et économiques. Ces mêmes idéaux « connotent les fonctions référentielles de représentation⁴² », c'est-à-dire qu'en se rattachant à des éléments inscrits dans la culture, ils convainquent du bien-fondé de ce qu'ils affirment. J'étudie les sites et la communication des concurrents principaux, aussi bien dans le cadre de mon stage que par curiosité personnelle. Cette recherche sur différentes sociétés de neuromarketing à l'échelle internationale permet de cerner les promesses des données émotionnelles pour ensuite effectuer des recommandations. J'ai aussi demandé lors des entretiens quelles étaient les possibilités d'applications de la donnée émotionnelle. Ce même procédé a été répété lors des salons auxquels Datakalab participait. Ainsi, j'ai pu enrichir mes idées grâce à des professionnels ou des consommateurs rencontrés par exemple lors de la Nuit de l'IA. Un système d'alertes qui me permettait de recevoir instantanément les dernières parutions sur des mots clés comme « facial coding », « feel data© », « predictive eye tracking ». En ce sens, des articles sur les applications de la donnée émotionnelle ainsi que sur ses dangers étaient lus quotidiennement. Pour traiter ce sujet aussi sensible, il s'avérait aussi crucial de se renseigner sur les lois autour du neuromarketing et sur la loi française dans le cadre de la récolte de données personnelles. La lecture d'articles en SIC m'a également permis de prendre de la distance sur le contexte politique et économique dans

⁴⁰ Baudrillard, Jean, *Le Système des objets*, Gallimard, 1968, analyse de la "foi" publicitaire, mise en rapport avec celle des enfants qui croient au Père Noël

⁴¹ Bigot, Jean-Édouard « Instruments, pratiques et enjeux d'une recherche numériquement équipée en sciences humaines et sociales », « Ce point de vue, opérant dans l'énonciation de la représentation, et les schèmes qui l'instituent, convoquent des *prétentions médiationnelles* issues d'une histoire culturelle de la technologie intellectuelle. Cela définit l'*opérativité connotative*, ou « symbolique » (Davallon, 2007), du dispositif, dans le sens où ces prétentions à la médiation connotent les fonctions référentielles de représentation, p. 254

⁴² *Ibid.*

lequel s'inscrivent tous les idéaux liés au progrès, à l'innovation, et à la confiance accordée aux chiffres dans notre société.

PLAN

Je dégage ici trois axes de réflexion :

- 1. J'analyserai la construction de la légitimité des données émotionnelles de Datakalab sur un renouveau technologique et sur les imaginaires autour du neuromarketing,**
- 2. L'évolution des « discours d'escorte » de la « feel data© » suite au repositionnement de Datakalab,**
- 3. Enfin, Datakalab applique sa nouvelle vision et s'intéresse à de nouveaux marchés.**

1 UNE LÉGITIMITÉ DE LA « FEEL DATA© » CONSTRUITE SUR UN RENOUVEAU TECHNOLOGIQUE ET SUR DES IMAGINAIRES AUTOUR DU NEUROMARKETING

En 2016 lors de leur lancement au salon Big Data, Datakalab présente l'une de ses inventions, les « feel data© ». Ces données qui semblent évidentes relèvent d'un « construit communicationnel⁴³ » au sens de Jeanneret. Révolutionnaires, elles peuvent tout de même être considérées comme issues de tendances telles que le « big data » ou la nécessité de jouer sur l'émotion. Je me concentre alors sur ces données en les considérant en tant que « traces » au sens donné par Yves Jeanneret, c'est-à-dire, un construit qui peut revêtir plusieurs formes⁴⁴. Émotions typiquement humaines (voire animales) et données relatives à la machine constituent deux domaines éloignés. La promesse de « mesure émotionnelle » dispose à la fois d'un aspect intuitif, puisque directement intelligible, mais aussi d'un autre complètement antithétique. Pour qualifier une donnée « d'émotionnelle », Datakalab doit faire passer une donnée brute en donnée intelligible, et tenir les bons discours d'accompagnement. Ces données sont sélectionnées, traitées, et pourtant présentées comme neutres. On peut aussi se demander si elles permettent réellement d'appréhender la charge symbolique des émotions. Comment ramener la « mesure de l'émotion » à un traitement computationnel ?

1.1 Les données émotionnelles sont issues et portées par plusieurs tendances

La « feel data© » a été forgée par Datakalab, mais aussi par le marché lui-même. Celui-ci se développe rapidement tout en s'inscrivant dans d'autres tendances. Développée à partir des années 1950, « l'intelligence artificielle » désigne un ensemble de procédés pour reproduire l'intelligence humaine en imitant son raisonnement⁴⁵. « L'affective computing » soit « informatique affective », l'une de ses branches, s'intéresse aux développements de systèmes capables d'identifier et de modéliser les émotions humaines⁴⁶. L'affective computing peut

⁴³ Jeanneret, Yves, « Faire Trace : Un dispositif de représentation du social », fait partie d'un numéro thématique « De la trace à la connaissance à l'ère du Web » *Intellectica*, 59, p. 42, 2013 « disponible à l'adresse https://www.persee.fr/doc/intel_0769-4113_2013_num_59_1_1085 »

⁴⁴ Péliissier, Daniel, *Présence numérique des organisations*, « "Trace", vous avez vraiment dit " Trace" ? » : « Dans une même perspective, Yves Jeanneret à partir de 2011, a mis en évidence toute la complexité du concept de trace en plaçant pour une catégorisation du concept. En effet, la trace masque des notions différentes comme l'inscription ou le tracé qui portent des enjeux différents et plus polysémique que l'indice »

⁴⁵ Dictionnaire Larousse, « Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine », « Disponible à l'adresse : https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257 »

⁴⁶ ISIR, « Emotion, Affective Computing », « L'informatique affective (Affective Computing) est un domaine

collecter des données non structurées comme des visuels, documents audio, grâce à des caméras ou autres capteurs. En ce sens, l'informatique affective promet d'opérer un enregistrement automatique. Ce dernier permet « l'inscription⁴⁷ » de données brutes qu'évoque Jeanneret. Depuis quelques années déjà⁴⁸, on entend parler de « big data », de « health data » « d'open data » ou de « smart data ». Les recherches dans le monde entier⁴⁹ sur les « data » montrent que certaines typologies existent depuis le début des années 2000. Ces « buzzwords » peuvent être appréhendés comme représentatifs d'une idéologie inspirée des modèles américains⁵⁰ qui recherche l'innovation et le chiffre, dans la mesure où les recherches sont principalement effectuées là-bas, ainsi qu'en Inde et en Australie⁵¹. Le marché de la donnée émotionnelle aujourd'hui estimé à 123 millions de dollars devrait passer à 3,8 milliards de dollars annuels, soit être multiplié par plus de 30 d'ici 2025⁵². Facebook a déjà exploité des milliards de données émotionnelles. En février 2016, soit un mois avant la première prise de parole de Datakalab sur le réseau social, la plateforme a agrémenté son bouton « like » de cinq autres émojis : « j'adore », « haha », « wouah », « triste », « grrr ». Ainsi la plateforme sociale propose-t-elle plus de contenu en affinité avec les envies des internautes, en fonction des données émotionnelles que ces derniers choisissent de communiquer via leurs interactions. La conférence de Hub Institute du 22 avril 2019⁵³, projette que plusieurs secteurs vont se développer dans les prochaines années dans la création de valeur autour de la donnée émotionnelle, notamment dans la distribution, la publicité, les services professionnels, la santé et le gaming. Quelques siècles en arrière, l'Homme ne pouvait appréhender le réel qu'à travers ses cinq sens. Il a lui-même développé des outils pour capter une réalité différente, qui ne s'offrait pas à lui directement. Si pour Datakalab et les clients qu'elle a su convaincre, la donnée émotionnelle semble s'installer

interdisciplinaire récent qui a pour but de donner aux ordinateurs les capacités de reconnaître, d'exprimer, de synthétiser et de modéliser les émotions », « disponible à l'adresse : http://www.isir.upmc.fr/UserFiles/File/chetouani/InteractionVerbale_lecture5_web.pdf »

⁴⁷ Jeanneret, Yves. « Faire trace : un dispositif de représentation du social », *Intellectica*. Revue de l'Association pour la Recherche Cognitive, n°59, 2013/1. De la trace à la connaissance à l'ère du Web, pages. 41-63, « disponible à l'adresse : https://www.persee.fr/docAsPDF/intel_0769-4113_2013_num_59_1_1085.pdf », p. 48 : « L'inscription arrache la trace au monde naturel pour la projeter du côté de l'artefact médiatique, qui rend disponible et lisible une information »

⁴⁸ ANNEXE 17, Google Trends, recherches dans le monde de 2004 au 05/09/2019

⁴⁹ *Ibid.*

⁵⁰ *Ibid.*

⁵¹ *Ibid.* « Innovation » et « performance » par pays

⁵² ANNEXE 16, Tractica, « Emotion Recognition and Sentiment Analysis Market to Reach \$3.8 Billion by 2025 », 07/03/2018, « disponible à l'adresse : <https://www.tractica.com/newsroom/press-releases/emotion-recognition-and-sentiment-analysis-market-to-reach-3-8-billion-by-2025/> »

⁵³ Hub Institute, « Marketing : IBM et Datakalab séparent le grain de l'ivraie avec une IA qui détecte les émotions », 22/04/2019, « disponible à l'adresse : <https://youtu.be/CpxDrvymnq4> »

comme nouvel indicateur pour opérer des décisions stratégiques, ce « tracé⁵⁴ » n'est pas un acquis, mais bien un construit.

1.1.1 « Feel data© » fille de deux tendances : « big data » et émotion

1.1.1.1 Le « big data » comme système de croyance

On peut considérer cette donnée comme fille de deux tendances dont elle hérite du nom, « Big data » et « Émotion ». Comme le démontre Alexandra Ducrot dans sa recherche⁵⁵, le « big data » fait loi, politique, économie, management, entrepreneuriat constituent autant de domaines où il opère. Ses discours d'escorte se multiplient, et le rapport de la société à la donnée évolue. Des plateformes et des logiciels sont créés. De nouveaux postes dédiés apparaissent tels les Data Scientists, Data Analyst, Chief Data Officer. Il est « sur les lèvres de tous les managers et entrepreneurs, fait les gros titres des journaux et anime des conférences entières⁵⁶ » comme le souligne la revue en ligne Maddyness spécialisée dans les start-up et l'innovation. Le contexte socio-économique et professionnel place le « big data » au cœur d'un système de valeurs. Il se transforme progressivement lui-même en domaine d'études, secteur économique et se réclame d'une connaissance illimitée qu'il faut structurer. La donnée suscite alors une fascination, dans la mesure où sa promesse semble infinie. Graal de tout décideur, elle est entourée d'imaginaires, comme si elle détenait une vérité devant être décryptée dans la mesure où son accès n'est pas immédiat. Elle prend même la forme de nombres, de suites non intelligibles pour l'homme. Ainsi sur des sites comme celui d'IBM Watson, on retrouve des formules telles que *unlock hidden value in data to find answers*⁵⁷ dans la rubrique « Watson discovery ». La formule peut être ici comparée à une invitation à un jeu de piste, avec des termes qui relèvent de l'intrigue, voire de la fiction : « unlock », « hidden », « answers ». Ces données sont utilisées par tous les acteurs du marché, mais surtout par les « GAFAM ». Ces investissements et imaginaires autour de la « data » contribuent à faire grossir l'objet « big data ». Grossissant, l'objet se prépare lui-même à accoucher d'un nouveau type de donnée.

⁵⁴ *Op. Cit.* Jeanneret, Yves : « le tracé relève du monde graphique, il donne forme particulière aux images, aux récits et aux symboles », p. 48

⁵⁵ Ducrot, Alexandra, « Discours, objets et imaginaires autour des « technologies big data » : Le cas des consultants en technologie, la construction d'une expertise à travers la « raison algorithmique » Mémoire, promotion 2016-2017, p.7

⁵⁶ Maddyness, « Le Big Data est-il le nouvel eldorado des entreprises européennes ? » disponible à l'adresse : <https://www.maddyness.com/2015/06/05/big-data-eldorado/>, 22/03/2018

⁵⁷ « Je traduis : Libérez le potentiel de la data pour trouver des réponses », « disponible à l'adresse : <https://developer.ibm.com/events/unlock-insights-in-your-text-data-with-watson-discovery-service-zurich-09-12-2019/> »

Comme une étoile devenue trou noir, cet objet absorbe désormais la matière qui contribue à le construire. Cette masse infinie fascine et inquiète à la fois. Selon Bernhard Rieder⁵⁸, ces données, ces « traces » constituent un « dispositif d'observation "total" qui promet de révolutionner les méthodologies classiques en effaçant le clivage entre micro et macro, entre qualitatif et quantitatif ». On redoute qu'elles n'échappent au contrôle de l'homme, voire qu'elles le poussent à contrôler son propre comportement et constituent ainsi la forme moderne du *Panopticon*⁵⁹, la prison au design circulaire étudiée par Foucault. Les gardes peuvent y surveiller les prisonniers depuis leur tour, placée en plein milieu du bâtiment. Ils peuvent ainsi garder un œil sur toutes les cellules. Ce système de surveillance pousse les individus à réguler eux-mêmes leur propre comportement, de peur d'être surveillés, et ce d'autant plus que l'émotion est considérée comme une donnée personnelle, intime, qui risque d'être exploitée par exemple à des fins commerciales.

1.1.1.2 Un renouveau technologique permet l'avènement des données émotionnelles

La transition vers le numérique a fait évoluer plusieurs secteurs. Il y a une vingtaine d'années, le « hardware » soit « l'équipement matériel, mécanique, magnétique, électrique et électronique, qui constitue un ordinateur, ou des machines de traitement de l'information en général⁶⁰ » ne permettait pas de développer de tels algorithmes. L'amélioration de la puissance de calcul de la machine, l'explosion du nombre de données et sa nouvelle facilité d'accès ont permis un renouveau technologique et sociétal. La capacité de stockage des ordinateurs, des serveurs et des logiciels a augmenté. Les mentalités ont évolué au sujet de la donnée. Davantage de chercheurs travaillent aujourd'hui en « machine learning⁶¹ » et en « deep learning⁶² ». De nouvelles théories permettent de développer les « algorithmes ». En mars 2019, les chercheurs Yann Le Cun, Yoshua Bengio et Geoffrey Hinton ont remporté le prix Turing⁶³, pour leurs travaux sur l'intelligence artificielle et le « deep learning ». Le « deep learning » a pris forme

⁵⁸ Rieder, Bernhard, *Études de communication* « Pratiques informationnelles et analyse des traces numériques : de la représentation à l'intervention » p.91-104, 2010, « disponible à l'adresse : <https://journals.openedition.org/edc/2249> »

⁵⁹ Foucault, Michel, *Surveiller et punir*, Éditions Gallimard, Paris, p. 211, 1975

⁶⁰ CNRTL, *Métron*, 1975

⁶¹ Technologie permettant aux ordinateurs d'apprendre sans avoir été programmés à cet effet.

⁶² Yann Le Cun le définit comme le fait « d'utiliser des données pour entraîner une machine à exécuter une tâche particulière au lieu de la programmer explicitement. Par exemple, on peut entraîner une machine à reconnaître des images, en lui montrant des milliers d'exemples d'images. La machine s'ajuste automatiquement pour reconnaître ces objets »

⁶³ Ministère de l'Enseignement supérieur de la Recherche et de l'Innovation, « Attribution du prix Turing à Yann Le Cun », « disponible à l'adresse : <https://www.enseignementsup-recherche.gouv.fr/pid26414-cid140408/attribution-du-prix-turing-a-yann-le-cun.html> », 27/03/2019

dans les années 2010, notamment grâce au développement d'immenses bases de données. La « feel data© » s'est donc appuyée sur un nouveau technologique.

1.1.1.3 Datakalab, un nom chargé d'imaginaires

Le construit que Datakalab donne à voir s'imbrique dans un nouveau politique et économique, qui peut être considéré comme issu des systèmes américains. Lorsque Frank Tapiro m'explique le choix du nom « Datakalab », il évoque certains des imaginaires qui y sont liés⁶⁴ : « J'ai été le premier à inventer le *data capital*, j'avais écrit un papier pour expliquer que plutôt que de se battre contre l'utilisation des données, chacun devrait se les approprier, et les louer, les vendre pour créer un *data capital* ». Ce terme peut renvoyer à une idéologie capitaliste liée à la production et à la vente. Les sciences font loi et le nombre est sacralisé par les décideurs. Le suffixe « Lab » renvoie au laboratoire, lieu de fabrication du savoir. Du latin, *laborare*, travailler, c'est là que les chercheurs expérimentent, font des découvertes, résolvent des problèmes ici autour de la donnée émotionnelle. Dans le nom même de la start-up, on retrouve beaucoup des logiques de médiation qui s'y opèrent. Frank Tapiro ajoute « Datakalab sonnait un peu comme « abracadabra », cela renvoie à la magie. En plus, personne n'arrive à le dire, ce qui me permet de rebondir dessus, d'en faire une blague. Aujourd'hui, en télé, en conférence, tout le monde en parle⁶⁵ ». Différents acteurs-décideurs marketing, créatifs, médiatiques, s'accaparent ce nom qui peut sembler imprononçable ainsi que les discours qui l'accompagnent. Plusieurs imaginaires interviennent dans cette présentation : entrepreneuriat, technologie, innovation, marketing, neurosciences, algorithmes. Cette accumulation permet d'emblée de se forger une idée de ce qu'est Datakalab, et de la manière dont elle peut communiquer. Le nom même de Datakalab invoque donc les imaginaires autour de la donnée, du capitalisme, du laboratoire, en plus de se référer au titre du film *Bienvenue à Gattaca*, qui parle de transhumanisme⁶⁶. La donnée émotionnelle qui se répand partout dans le monde peut aussi être considérée comme issue d'une autre tendance qui souhaite placer l'émotion au cœur des stratégies.

⁶⁴ ANNEXE 4, Entretien avec Frank Tapiro, Co-fondateur et « Chief Emotion Officer », 26/06/2019, Paris 17^{ème}

⁶⁵ *Ibid.*

⁶⁶ *Ibid.*

1.1.1.4 L'émotion : un concept flou, évolutif et complexe

Les marques réalisent aujourd'hui qu'il est nécessaire de jouer sur l'émotion dans leur communication. Dans plusieurs conférences – notamment au Printemps des Études 2019 – les experts soulignent la nécessité de créer de l'affect⁶⁷. Les consommateurs « sont à la recherche d'une publicité qui fait sens et avec laquelle ils peuvent s'identifier » affirme Laetitia Sousa-Tagarro dans l'article de Espub⁶⁸ : « Les feel data, nouveau levier de mémorisation pour les marques ? ». De nouvelles solutions se développent alors pour mieux faire réagir, pallier la lassitude et créer de l'attention dans une société de sursollicitation. Pour créer de l'émotion, les publicitaires utilisent différentes techniques, telles que le marketing « tribal », « expérientiel ». Ici, la marque s'inscrit dans l'imaginaire littéraire, au même titre que les contes, les « mythes » de Barthes. Les campagnes suivent de nouveaux axes créatifs et l'utilisent de plus en plus pour toucher les cibles. *Everyone knows what is an emotion until asked to give a definition*⁶⁹ affirment les psychologues Beverley Fehr et James A. Russell. Si cette définition s'avère si difficile, c'est sans doute parce qu'elle est typiquement enracinée en nous, liée à tant de causes différentes qu'on ne sait la saisir. Aussi sa définition a-t-elle évolué avec le temps. L'émotion a pu être identifiée à une maladie de l'âme qui perturbe la pensée⁷⁰. Elle est utilisée depuis l'Antiquité, lorsque les rhéteurs jouaient sur le *pathos* dans leurs discours. Du latin *e*, « qui vient de, au-delà » et *movere*, « mouvoir », « émouvoir », elle traduit un déplacement qui transcende, dépasse la force d'habitude. L'émotion comme moyen de créer de la différence et donc de la performance constitue un sujet complexe, dans la mesure où celle-ci n'est pas palpable, souvent dissimulée, et surtout, relative au vivant. Si l'émotion est relative à l'homme voire même à l'animal, elle est dans tous les cas rattachée à l'esprit et au conscient. Comment la « mesurer » autrement que par nos sens, notre instinct ? L'humain lui-même ne sait pas toujours quelles émotions ressentent ceux qui l'entourent. C'est là tout l'enjeu de Datakalab : souligner l'évidence de ces données émotionnelles alors qu'elles sont à l'origine des traces brutes captées par une machine.

⁶⁷ ANNEXE 19, Photos de la conférence « OpinionWay x Datakalab », Printemps des études, 04/2019

⁶⁸ Sousa-Tagarro, Laetitia, Espub, « Les feel data, nouveau levier de mémorisation pour les marques ? » 22/01/2018, « disponible à l'adresse : <https://espub.org/feel-data/> »

⁶⁹ Fehr et Russell, « Concept of Emotion Viewed From a Prototype Perspective ». p. 464, 1984, « disponible à l'adresse : <https://pdfs.semanticscholar.org/4c25/97a92ce4de6b58e7c2b4140eed1c1046ffbd.pdf> », « Je traduis : Tout le monde sait ce qu'est une émotion avant que ne demande de la définir »

⁷⁰ Kant, Emmanuel, *Anthropologie du point de vue pragmatique*, 1863, p. 215

1.2 Datakalab promet de « mesurer les émotions », soit de rationaliser l'irrationnel

À son lancement en 2016, Datakalab promet de « mesurer les émotions » grâce aux outils « issus des neurosciences ». Yves Jeanneret définit la « promesse » comme « expression explicite d'une proposition sur la communication et ce qu'elle peut apporter à des publics⁷¹ ». Datakalab opère alors en France une véritable révolution de par cette promesse contre-intuitive et inédite. Grâce à deux outils⁷², le « bracelet connecté » et le « facial coding », Datakalab veut optimiser « l'expérience⁷³ » de marque en analysant des données émotionnelles. Ce que l'on nomme ici « expérience » désigne la manière dont le consommateur appréhende sa relation à une marque ou tout autre producteur de contenu. Datakalab se sert de deux systèmes pour capter et construire des « traces », qui revêtent plusieurs typologies tout au long du processus de création de sens. Au cours de leur traitement, elles sont modifiées, subissant de nombreuses opérations pour synthétiser la réalité captée par la machine et lui donner du sens. Jusqu'à début 2019, le « bracelet connecté », d'origine médicale, promettait de « mesurer » les signes physiologiques de l'émotion comme la « fréquence cardiaque », la « micro-sudation », la « température extra-dermale ». Il « mesurait l'expérience », soit l'ensemble des stimuli qui peuvent nous atteindre dans un point de vente. Plus concrètement, lors de « journées terrain » des hôtes équipaient des volontaires de cet outil. Ensuite, ces personnes effectuaient leur parcours client dans un lieu spécifique comme un magasin ou une gare. Les hôtes les suivaient, et à l'aide d'une tablette, elles « inscrivaient », « labélsaient » l'action en cours (parler avec quelqu'un, acheter quelque chose...). Les « pics émotionnels » s'inscrivaient alors directement sur la tablette de l'hôte à partir des données physiologiques mentionnées. Les hôtes n'avaient plus qu'à montrer au répondant ses pics émotionnels et à l'interroger à chaque fois qu'ils dépassaient certains seuils pour comprendre leurs significations *via* le déclaratif. Les données étaient ensuite agrégées sur une centaine de personnes afin de pouvoir distinguer des tendances et des axes d'amélioration de l'expérience en point de vente.

À son lancement, Datakalab faisait également de la prestation, de la mise en relation entre les annonceurs et les sociétés de neuromarketing qui mettaient à disposition leur outil de

⁷¹ Jeanneret, Yves, *Critique de la Trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris, Editions Non Standard, 2014 lexique, p. 14

⁷² ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019 Rubrique « Nos outils »

⁷³ Nebia, Amelle, CB News, « Datakalab poursuit sa croissance avec de nouveaux clients », 13/12/2017, disponible à l'adresse : « <https://www.cbnews.fr/conseil/image-datakalab-poursuit-sa-croissance-avec-nouveaux-clients-27143> » : « Anne-Marie Gaultier, ex « Présidente de Datakalab », « la mesure des émotions se révèle particulièrement importante pour mieux appréhender l'expérience client »

« reconnaissance des expressions faciales⁷⁴ ». Elle a notamment « optimisé » des affiches publicitaires qui n'attiraient pas assez l'attention, soit le regard des panélistes, et ne déclenchaient pas d'émotion, grâce à des « algorithmes » de « reconnaissance des expressions faciales ». On peut définir l'algorithme comme un « raisonnement » formalisé et automatisé dans le cadre du fonctionnement de certains programmes⁷⁵. Le « facial coding » promet d'identifier les émotions sur les visages des panélistes précisément grâce à ce traitement computationnel. Ici, les études s'effectuent avec des panélistes, soit des personnes qui acceptent de répondre à des questionnaires en ligne. Si elles consentent à ce que la start-up accède à leur webcam, Datakalab leur fait visionner un contenu et capte leurs expressions faciales *via* des « algorithmes » (à l'époque non propriétaires). Comme dans le cas du « bracelet connecté », les données anonymes sont agrégées pour pouvoir tirer des vérités générales sur le contenu testé. L'expression de « mesure » des émotions revêt alors de fortes « prétentions communicationnelles⁷⁶ ». Ce « mode d'intervention sur les processus de communication⁷⁷ » affirme à lui seul évaluer la durée, l'étendue, la quantité, le poids des émotions. Il implique une reconstruction symbolique de toute la charge portée par l'émotion suite à son traitement computationnel. Pour prendre un exemple avec l'ancien outil « bracelet connecté », lors d'une étude⁷⁸ de 2017 pour la SNCF, Datakalab a mesuré les « pics émotionnels » des voyageurs acceptant le test. Des hôtes ont équipé les participants avec le bracelet, et les ont suivis durant leur parcours en gare. Le bracelet enregistrait les « signes physiologiques » choisis comme étant ceux de l'émotion. Ces « métriques » correspondent à ce que plusieurs chercheurs en neurosciences ont pu définir comme tels. Bien sûr, cette sélection n'est pas neutre, elle relève elle-même d'un construit qui a une portée symbolique. Tous ces signes physiologiques captés de l'émotion comme la « fréquence cardiaque » constituent au sens donné par Jeanneret des « indices⁷⁹ » ici de l'émotion. L'auteur affirme : « C'est en effet la transformation de l'indice en inscription, puis de l'inscription en tracé – c'est-à-dire le passage graduel du monde de la causalité à celui de l'expression – qui rend possible une lecture par les uns de ce que font les autres⁸⁰ ». Le bracelet connecté permet l'enregistrement automatique de ces « indices

⁷⁴ Datakalab a ensuite développé son propre outil propriétaire de « facial coding »

⁷⁵ Définition notée lors du cours de Master 2 « La trace et le calcul » dispensé par Antoine Bonino au CELSA, pour les élèves de l'option « Marque & Branding »

⁷⁶ *Op. Cit.* P.14

⁷⁷ *Ibid.*

⁷⁸ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, Page « LES MISSIONS » Étude SNCF Gares et connexions « disponible à l'adresse <https://www.datakalab.com/missions> »

⁷⁹ *Op. Cit.*

⁸⁰ Jeanneret, Yves, « Les harmoniques du Web : espaces d'inscription et mémoire des pratiques », « disponible à l'adresse : http://www.mei-info.com/wp-content/uploads/2014/02/MEI_32_02.pdf »

de l'émotion », les faisant alors passer au statut « d'inscriptions ». Cette collecte mise en forme automatiquement via une courbe permet alors à l'hôtesse de présenter au participant ses « tracés⁸¹ », formes interprétables des « indices » initiaux. Ces « tracés » montrent des « pics d'intensité émotionnelle », construit rendu intelligible, dans un premier temps simplement pour l'hôtesse et le participant. C'est ce même « tracé » qui permet au voyageur d'indiquer à l'hôtesse à quoi correspondent les « pics » rendus visibles. Ce phénomène ne constitue lui-même que le point de départ d'autres traitements de cette « trace » par Datakalab, qui en « agrégeant » tous les résultats des participants forge à nouveau d'autres « tracés », qu'elle soumet ensuite à sa propre analyse, puis à celle des décideurs. Une fois livrés aux clients, ces derniers sont à nouveau ramenés au statut « d'indice », mais cette fois-ci de la réception de l'expérience du consommateur. Ici débute la construction de ce qu'on pourrait nommer une « méta-trace », un « méta-indice ». Ces données émotionnelles sont alors rassemblées par les décideurs, matérialisées en de nouvelles « inscriptions », lorsque les « Chargés d'Études » les intègrent à leurs présentations par exemple. Elles sont à nouveau rendues intelligibles pour être présentées à d'autres acteurs et ainsi de suite. Ces mêmes processus sont traversés par des médiations sociales et symboliques. Dans tous les cas, la start-up promet que ces deux outils, « bracelet connecté » et « facial coding » peuvent « mesurer » des données brutes pour ensuite les traduire en données intelligibles. Ce sont bien ces données rendues compréhensibles et surtout intéressantes pour les clients que la start-up cherche à instituer en tant « qu'indices », alors qu'elles constituent elles-mêmes des formes déjà traitées « d'indices » antérieurs. Elle veut alors placer ces données émotionnelles à la base et au cœur d'un nouveau processus de construction de sens, cette fois-ci acté par les décideurs eux-mêmes, voire par d'autres acteurs comme leurs clients ou les médias. Datakalab doit alors construire une « promesse » suffisamment forte pour imposer cet objet jusqu'alors méconnu comme nouvel essentiel de la boîte à outils des décideurs. À l'époque encore, elle doit faire en sorte de masquer cet artifice pour qu'il apparaisse comme allant de soi *via* des discours autour de cette donnée. Ce travail semble ambitieux, dans la mesure où la construction de données brutes en données intelligibles constitue le quotidien des mêmes acteurs qui veulent la présenter comme naturelle. Datakalab s'appuie alors sur sa vision pour rendre cette promesse péremptoire.

⁸¹ *Ibid.*

1.2.1 « Redonner visage humain à la data⁸² »

La donnée peut sembler indigeste, cette suite de chiffres qu'on voit lorsqu'un algorithme tourne a de quoi étourdir. Datakalab souhaite replacer l'humain au centre, comme le font déjà toutes les stratégies qu'on qualifie de « human centric ». *In fine*, c'est bien lui qui opère les choix communicationnels. Par exemple, ce sont bien les créatifs qui décident des éléments à apposer sur un visuel. La volonté de Datakalab n'est donc pas de réduire la pensée critique et créative à néant, bien au contraire. Les données émotionnelles sont présentées par Datakalab comme « outils d'aide à la création ». En utilisant les « émotions » de la cible, on peut rendre un contenu impactant et mémorable. Datakalab dit chercher à éradiquer les contenus qui laissent indifférents parce qu'ils ne déclenchent pas d'émotion⁸³. Le produit « Datakalab Content » mesure par exemple les « pics d'émotions » pour un contenu de marque. Datakalab a travaillé avec Eric Bompard, qui souhaitait comparer deux versions d'un spot. La version de 30 secondes générait plus d'émotions que la version plus longue de 45. Suite à ce constat, Eric Bompard a réduit d'1/3 ses achats médias. La donnée émotionnelle se revendique donc un outil d'aide à la création. Or, comment convaincre d'une optimisation possible de la création et du processus créatif typiquement humain, grâce à la machine ? Datakalab renverse tout simplement ce traitement computationnel. La donnée émotionnelle qui en résulte est présentée comme *seule* donnée typiquement humaine. Ces données « chaudes⁸⁴ », humaines, récoltées grâce aux « outils issus des neurosciences » sont opposées par Datakalab aux données « froides », qu'elle présente comme « rationnelles, transactionnelles, classiques », telles que le « taux de clics », le « panier moyen », et le « temps passé » sur une page web. Cette distinction opérée par les sociétés de neuromarketing hiérarchise les données. Les données émotionnelles disposent d'une empreinte chaude, sensible, presque vitale, alors que les autres données évoquent la froideur soit le « manque de sensibilité, de passion⁸⁵ », l'absence relative d'émotion qu'on attribue généralement à l'acte marchand.

⁸² ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, Page « LE LAB » « disponible à l'adresse : <https://www.datakalab.com/lab> »

⁸³ ANNEXE 4, Entretien avec Frank Tapiro, Co-fondateur et « Chief Emotion Officer », 26/06/2019, Paris 17^{ème}

⁸⁴ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, Rubrique « notre promesse » : « Datakalab mesure les émotions des clients et les transforme en données émotionnelles et chaudes les feel data© »

⁸⁵ Dictionnaire Larousse

1.2.2 Revendication d'un savoir neuroscientifique

Pour mettre en place cette vision, Datakalab affirme détecter les « micro-expressions », qu'elle présente comme « universelles » en s'appuyant sur un savoir en sciences cognitives. Elle affirme : « Créé à partir des travaux de Paul Ekman en 1970, le *facial coding* permet de détecter les micro expressions du visage et d'identifier les sept émotions universelles : la joie, la surprise, la peur, la tristesse, la colère, le dégoût, le mépris⁸⁶ ». Ici la start-up replace le contexte et introduit les travaux du psychologue. Surtout, elle définit les émotions comme « universelles » en jouant du même appui. À son lancement en 2016, ce sont bien ces sept émotions que Datakalab promet de « mesurer » grâce au « facial coding ». D'après ces recherches, la culture n'aurait donc pas d'influence sur le ressenti et leur expression, elles seraient « universelles⁸⁷ ». En invoquant ce savoir en sciences cognitives, Datakalab s'inscrit dans une posture didactique et pédagogique envers ses clients, ce qui lui octroie du crédit.

1.2.2.1 Construction d'un « ethos neuroscientifique »

Dès 2016, tous les efforts de Datakalab se concentrent alors sur la mise en place d'un « ethos neuroscientifique⁸⁸ ». Emprunté au grec ancien, « l'ethos » désigne l'image que le locuteur donne de lui-même à travers son discours. Cet « ethos » peut être défini comme une « posture, un caractère, une habitude ». Datakalab se présente alors comme une start-up dans la « brain tech », c'est-à-dire dans la « technologie appliquée aux neurosciences ». Celles-ci étudient le système nerveux, notamment le cerveau. En 2016, « brain tech » n'est pas une expression connue en France. Soukeina Hassanaly, ex « Key Account Manager » chez Datakalab m'explique⁸⁹ : « Nous avons décidé d'utiliser ce terme, très courant aux États-Unis. En France il surprenait l'année dernière, maintenant les gens comprennent parce qu'ils

⁸⁶ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019

⁸⁷ Datakalab n'hésite pas à présenter la manière dont Paul Ekman a fait cette découverte. Le psychologue se rend en effet en Nouvelle-Guinée dans le cadre de sa recherche. Il montre des photos de signes exprimés de différentes émotions aux membres d'une tribu aborigène – qui n'ont donc que très peu de contact avec la culture occidentale. La tribu est capable de reconnaître les émotions sur ces photos. Paul Ekman en déduit que certaines expressions faciales émotionnelles sont générées et interprétées de la même façon chez tous les Hommes, indépendamment de la culture. Datakalab revendique ce savoir, mais d'autres anthropologues viennent cependant nuancer ces travaux en considérant que les émotions ne sont pas universelles⁸⁷, qu'elles dépendent de notre caractère, de notre culture et de notre milieu social. Il n'existe donc pas à ce jour de consensus sur ce sujet.

⁸⁸ Alsafar, Ali, « Ethos discursif et construction des rapports intersubjectifs dans les professions de foi des élections présidentielles de 2007 et de 2012 », « disponible à l'adresse : tel.archives-ouvertes.fr/tel-01146590/document » : « l'ethos « désigne l'image de soi que l'orateur construit dans son discours pour contribuer à l'efficacité de son dire ». Dans les sciences du langage, Dominique Maingueneau définit l'ethos comme « ce que l'orateur prétend être, il le donne à entendre et à voir », p. 7

⁸⁹ ANNEXE 5, Entretien avec Soukeina Hassanaly, Ex Key Account Director chez Datakalab, le 05/04/2019, Paris 17^{ème}

s'intéressent davantage à tout ce qui touche aux neurosciences ». Le terme s'inscrit lui-même dans une rhétorique commerciale de l'innovation. Il découle de l'idéologie capitaliste instituée dans les dernières années. D'abord, Datakalab a choisi de s'installer dans *The Brain Place To Be*, lieu de référence pour la recherche en neurosciences. Soukeina Hassanaly, m'explique aussi les raisons de ce choix : « À la base l'objectif était aussi de pouvoir y faire des rencontres, du « networking ». (...) Les trois entreprises qui y travaillent invitent leurs clients, et on parle tous de sujets liés aux neurosciences. (...) l'objectif est de devenir un lieu incontournable ». Sur la page d'accueil de son site, sur ses documents de présentations, elle reprend tout un univers scientifique du « Lab », avec l'utilisation de la table de Mendeleïev pour instituer les émotions comme l'un des éléments constitutifs de notre réalité.



Table de Mendeleïev présente sur la page « ACCUEIL⁹⁰ » du site en juin 2019
« disponible à l'adresse : <https://www.datakalab.com/> »

Datakalab se sert de cette « trace » pour instituer ses données émotionnelles en tant « qu'indice ». En ce qui concerne l'inspiration visuelle, Datakalab utilise le violet bordeaux dans sa charte graphique. La couleur est présente à la fois sur le site internet et dans ses présentations. Évoquant la douceur, l'émotion, le rêve, le violet se rapporte au spirituel et à l'intellect. Ici la nuance bordeaux évoque davantage l'énergie. Datakalab participe aussi à de nombreuses conférences, comme lors du Colloque « Le pouvoir des Emotions⁹¹ » au Centre Pompidou. Anne-Marie Gaultier y prêche l'importance de l'émotion : « Vous avez compris qu'aujourd'hui tout est émotions. On l'a vu, même les discours les plus cartésiens, même les discours de banquiers sont émotions — Sorry Mr Descartes ». Datakalab revendique alors ses connaissances en neuromarketing. Interrogée sur la volonté de suivre cette tendance issue des États-Unis, Soukeina Hassanaly répond : « Oui et non. Xavier a travaillé chez Emotient, donc oui, sauf qu'il est allé plus loin. Il s'est demandé quelles étaient les limites de Emotient. Quelle était la technologie utilisée, et comment améliorer les algorithmes ? (...) Nous ne surfons pas

⁹⁰ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019

⁹¹ ANNEXE 1, Retranscription de la conférence « Datakalab au centre Pompidou ! » du 15/09/2018

sur ces tendances [« l'IA », « les neurosciences », le neuromarketing], nous les avons anticipées. On parlait déjà de neurosciences et d'émotions avant que le marketing ne s'oriente là-dedans. On a été créés avant même qu'on en parle, donc on ne surfe pas sur la tendance. Je pense plutôt qu'on en a été précurseurs⁹² ». Datakalab se présente ici comme moteur qui crée avant même que le besoin d'une nouvelle donnée ne soit ressenti ou exprimé. Ici encore, elle semble anticiper les conditions de sa réussite.

1.2.3 Datakalab se pose en s'opposant aux études déclaratives

À son lancement, Datakalab affirme sa promesse de « mesure émotionnelle » en s'opposant aux études déclaratives. « Le problème des études de marché, c'est que les gens ne pensent pas leurs sentiments, ils ne disent pas ce qu'ils pensent et ne font pas ce qu'ils disent⁹³ » d'après David Ogilvy. Le publicitaire évoque ici la propension de l'individu à ne pas exprimer son vrai ressenti. Les méthodes de recueil d'information – études de marché, questionnaires consommateurs, groupes de discussion notamment – sont soumises à des biais. Ces « déformations, travers⁹⁴ », prennent différentes formes : cognitives, affectives, méthodologiques, sociales et culturelles. L'argumentaire principal des sociétés de neuromarketing défend que les réponses rapides, floues, subjectives, biaisées aux questionnaires déclaratifs, déforment le « vrai » ressenti consommateur. Le déclaratif permet d'analyser les ressentis du public suite à un événement. Les répondants verbalisent, et donc mettent des mots *a posteriori* sur des souvenirs. Yves Jeanneret définit la « prétention communicationnelle » comme un « mode d'intervention sur les processus de communication qui de manière délibérée ou insensible hiérarchise les éléments, détermine des conditions cruciales et légitime une certaine compétence à y intervenir⁹⁵ ». À l'époque, Datakalab « hiérarchise » les études émotionnelles comme supérieures aux études déclaratives. Avec le neuromarketing les ressentis ne sont pas exprimés, ils ne seraient donc pas « déformés » par le *logos*. C'est bien à partir de cet argument que les sociétés de neuromarketing comme Datakalab construisent leurs discours. « Les études déclaratives ne suffisent pas à mesurer objectivement

⁹² ANNEXE 5, Entretien avec Soukeina Hassanaly, Ex Key Account Director chez Datakalab, le 05/04/2019, Paris 17^{ème}

⁹³ Pavier Xavier, Jouanny Corinne, Cathy Daphné, Verez François, *Le design thinking au service de l'innovation responsable*, Edition Maxima Laurent du Mesnil, 2015, p. 33

⁹⁴ CNRLT, « disponible à l'adresse : <https://www.cnrtl.fr/definition/biais> »

⁹⁵ Jeanneret, Yves, *Critique de la trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, p. 15

les ressentis des consommateurs⁹⁶ » explique-t-elle. L'information récoltée serait synthétique, simplifiée parce que formulée. Datakalab promet donc de rationaliser l'irrationnel sans les biais liés à la verbalisation : « Les micro-expressions sont des expressions faciales brèves et inconscientes, donc involontaires, qui apparaissent lorsqu'on ressent une émotion. Elles sont réflexes et ne durent qu'1/15 à 1/25e de seconde⁹⁷ ». Cette présentation joue sur le fait que les micro-expressions, étant réflexes, sont incontrôlables. C'est en ce sens qu'elle revendique une mesure « non biaisée » des émotions des panélistes. Elle donnerait les clés d'accès à un savoir authentique, où l'humain s'abandonne à sa vraie nature, à des « ressentis purs » sans l'intervention du *cogito*.

1.2.3.1 Le neuromarketing comme clé d'accès à l'inconscient

Le neuromarketing prétend analyser les ressorts des décisions consommateurs directement à leur source, « l'inconscient ». Du privatif *in* et de *conscire* en latin, « avoir conscience de », il désigne ce qui échappe à notre conscience, ce qui n'est pas compris dans sa globalité par notre *cogito* et ce qui peut ressurgir et nous échapper de manière réflexe. Un nombre revient constamment dans la rhétorique de Datakalab. Il apparaît dans les discours tenus aux médias, aux annonceurs, et dans les présentations : « 90 % des stimuli qui nous impactent dans une journée sont traités par l'inconscient⁹⁸ ». En d'autres termes, la quasi-totalité de ce qui nous impacte ne passe pas par notre raison. C'est aussi en ce sens que le neuromarketing est fantasmé par les annonceurs. Il est promu par les sociétés de neuromarketing comme « détecteur de vérités ». L'expression est d'ailleurs utilisée par Datakalab à ses débuts lorsque l'on compare ses technologies – « bracelet connecté » et « reconnaissance des expressions faciales » – à un détecteur de mensonge. Sur la page d'accueil de la société de neuromarketing Spark Neuro, fondée en 2016 aux États-Unis, on lit : *tapping into the subconscious to get the truth, to inspire brand love, to drive action*⁹⁹.

⁹⁶ ANNEXE 20, Support de présentation PowerPoint de Datakalab

⁹⁷ *Ibid.*

⁹⁸ ANNEXE 2 Pitch Frank Tapiro – Challenge *Start-me up!* KPMG & Fevad

⁹⁹ « Je traduis : entrer dans le subconscient pour accéder à la vérité, pour faire aimer les marques et mener des actions »



Au 02/11/2019, Page d'accueil de Spark Neuro, « disponible à l'adresse <https://sparkneuro.com/> » et de Immersion Neuroscience au 02/11/2019, « disponible à l'adresse : <https://www.immersionneuro.com/> »

Le neuromarketing est ici présenté comme clé d'accès à la vérité cachée de l'inconscient qui permet d'actionner différents leviers. Une autre société de neuromarketing, Immersion Neuroscience, elle aussi américaine et lancée en 2018, promet « des informations objectives, précises et exploitables¹⁰⁰ ». Ce savoir donnerait les clés d'interprétation des ressentis des consommateurs, graal des décideurs. Comment, par quels outils Datakalab promet-elle alors d'avoir accès à l'inconscient ?

1.2.3.2 Un objet complexe, protéiforme et fantasmé

Le neuromarketing, sous-branche de la « neuroéconomie » se développe au tout début des années 2000. Ces deux domaines s'intéressent aux choix des consommateurs, et à l'étude de leur comportement dans l'économie. Le neuromarketing se concentre sur les réactions du cerveau humain face à des stimuli pour comprendre ce qui influence l'acte d'achat. L'une de ses expériences enseigne l'importance des imaginaires et des émotions dans l'appréciation d'un produit¹⁰¹. C'est en ce sens que Datakalab l'invoque à ses débuts. L'enthousiasme général a construit un mythe autour de ce qui touche aux neurosciences. Selon le chercheur Olivier Oullier, membre du laboratoire de psychologie cognitive à l'université d'Aix-Marseille : « Le neuromarketing a été survendu et mal compris à la fois (...) Certains "neuroopportunistes" ont survalorisé cette pratique en laissant croire que l'on pouvait lire dans le cerveau du consommateur pour le faire acheter¹⁰² ». L'engouement autour du neuromarketing l'a fait

¹⁰⁰ « Je traduis de : *Objective, Accurate and Actionable Information* »

¹⁰¹ Samuel M. McClure, Jian Li Damon Tomlin, Kim S. Cypert, Latane, M. Montague, et P. Read Montague "Neural Correlates of Behavioral Preference for Culturally Familiar Drinks", 2004 « disponible à l'adresse : <http://eugrafal.free.fr/McClu-et-al-2004.pdf> »

¹⁰² Gavard, Emmanuel, « Le Fantôme du neuromarketing », LSA, 17/05/2012 « disponible à l'adresse : https://www.lsa-conso.fr/le-fantasme-du-neuromarketing_130326 »

devenir un véritable « buzzword » comme en témoignent les chiffres de « Google Trends¹⁰³ » depuis 2004. « Neuromanagement », « neuromarketing », « neuroconsommateur », « neuroergonomie », « neuroesthétique », « neuroponcture » ... de nouveaux concepts apparaissent, tous formés à partir de ce préfixe « neuro », terme chargé d'imaginaires autour du cerveau. Il renvoie directement au « neurone », cellule du tissu nerveux.

1.2.3.3 Ambivalence du neuromarketing : cadre réglementaire et éthique

Datakalab doit alors faire face à l'ambivalence de la promesse du neuromarketing à la fois diabolisé, inquiétant, crédible et fascinant. On craint que le neuromarketing ne manipule pour faire consommer plus dans une société de la programmation où toute liberté individuelle a disparu. La France est dans une certaine mesure protégée par le Code civil, selon lequel l'utilisation des « IRM » à des fins marchandes n'est pas autorisée. Si de prime abord, on pourrait croire que les imaginaires autour des « neurosciences » et du neuromarketing portent la croissance de Datakalab, ce nouveau sujet fait d'autant plus peur qu'il est peu connu. Les acteurs de la donnée émotionnelle construisent un système de croyances sur lequel ils s'appuient pour construire leur crédit, à la fois en construisant et en déconstruisant d'autres mythes. Par exemple, Datakalab s'appuie sur les discours neuroscientifiques sur la nécessité de rendre un message impactant pour qu'il soit retenu, mais déconstruit les imaginaires selon lesquels la start-up « manipule les consommateurs ». Elle ne peut pas communiquer comme n'importe quelle entreprise. En 2016, lorsqu'elle fait converger ses efforts pour se faire connaître et rassurer sur ce terme, Les Echos titre « Datakalab, intrusion dans la tête d'un client¹⁰⁴ ». Ce sujet sensible a besoin d'être vulgarisé avec pédagogie. Il s'avère difficile à appréhender et va à l'encontre de notre éthique. Pour des sociétés comme Datakalab, le travail de pédagogie s'avère d'autant plus difficile qu'il faut faire comprendre aux consommateurs que l'on n'entre pas « dans leur tête », tout en convainquant tout de même que l'on a accès à leurs « réactions » inconscientes. Le sujet est délicat dans la mesure où l'émotion est perçue comme une donnée personnelle, sensible et intime. La loi française impose que le consommateur ait donné son accord en amont pour récolter ses données à caractère personnel. Comme on peut le lire sur son

¹⁰³ ANNEXE 17, Google Trends, recherches dans le monde de 2004 au 05/09/2019

¹⁰⁴ Richebois, Véronique, « Datakalab, intrusion dans la tête d'un client », Les Echos, 07/03/2016, « disponible à l'adresse : <https://www.lesechos.fr/2016/03/datakalab-intrusion-dans-la-tete-dun-client-214441> »

site, Datakalab s'oppose à toute utilisation de la donnée qui ne respecterait pas la vie privée des utilisateurs¹⁰⁵.

1.2.3.4 Les émotions au cœur du processus décisionnel

Lorsque Datakalab revendique toujours son expertise « conseil », elle s'appuie sur un savoir issu des neurosciences pour rédiger ses recommandations. Irrationnelle, l'émotion joue aussi pourtant un rôle clé dans nos décisions. Les travaux du neuropsychologue Damasio¹⁰⁶ montrent qu'elle est nécessaire pour prendre des décisions rationnelles. « 95 % de nos prises de décision sont largement influencées par nos émotions¹⁰⁷ » affirme la start-up dans ses premières années. Ici, les recherches en neurosciences se retrouvent « trivialisées¹⁰⁸ » et simplifiées pour vulgariser la « mesure émotionnelle ». Du latin, *trivium*, carrefour, l'expression désigne selon Yves Jeanneret la façon dont « les objets circulent, sont partagés dans la société, dont ils passent d'un espace à un autre, se transforment et prennent des sens nouveaux¹⁰⁹ ». À ses débuts, Datakalab utilise ce savoir pour soutenir sa dimension « conseil ». En juin 2017, Anne-Marie Gaultier, encore « Présidente » déclare : « Aucune décision humaine n'est prise sans émotion. Les professionnels ont intérêt à comprendre le ressenti des consommateurs pour faciliter l'acte d'achat et améliorer le parcours client. L'analyse des sentiments peut guider toutes sortes de décisions stratégiques prises par les entreprises. C'est pour cela que nous vendons à nos clients des données émotionnelles assorties d'une prestation de conseil¹¹⁰ ». Ici la communicante insiste sur l'émotion, le « ressenti » et les « sentiments » dans une isotopie. Ces différents concepts se retrouvent confondus. Elle utilise le champ lexical de « l'expérience client » en se concentrant sur l'affect. Dans la lignée de Damasio, elle place « l'émotion » comme prémisse à la décision – qu'elle soit du côté consommateur pour « faciliter l'acte d'achat et améliorer le parcours client¹¹¹ », ou du côté entreprise pour « guider toutes sortes de décisions stratégiques¹¹² ». L'argumentaire vise à justifier la prestation « conseil ». Les recommandations sont alors rédigées à partir des connaissances issues des neurosciences, par exemple sur la mémorisation

¹⁰⁵ ANNEXE 18, Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019, Rubrique « Notre engagement »

¹⁰⁶ Damasio, Antonio, *L'Erreur de Descartes*, « Cas Eliott », 1995

¹⁰⁷ *Op. Cit.* « disponible à l'adresse : <https://espub.org> »

¹⁰⁸ *Op. Cit.*

¹⁰⁹ Jeanneret, Yves, « La Trivialité » - Lexique des SIC, Effeuilage La Revue, 22/03/2014

¹¹⁰ De Matharel, Lélia, JDN, « Datakalab dissèque les émotions pour aider les pros à booster leur CA », 19/06/17, « disponible à l'adresse : <https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1195466-datakalab-disseque-emotions/> »

¹¹¹ *Ibid.*

¹¹² *Ibid.*

des messages, ou sur leurs conditions d'impact. Citer des chercheurs en neurosciences permet de rassurer sur la possibilité de ramener le traitement de l'émotion à un traitement computationnel.

1.3 Transformer une donnée brute en réalité intelligible

De nombreux procédés s'enclenchent pour convaincre le client, le faire entrer dans des imaginaires relatifs aux tendances et lui donner les « métriques » qui correspondent à ses objectifs. On peut alors s'interroger sur la construction de ce savoir par les traitements computationnels. Ce savoir a-t-il en soi une vraie valeur psychologique qui permet d'appréhender la charge symbolique des émotions ? Ou bien les processus de construction de sens pour lui apporter une information utile délaissent-ils les vérités psychologiques individuelles ?

1.3.1 Traductions du réel et simplifications par des formes

Datakalab capte l'émotion *via* des traitements computationnels et des traductions du réel. Dans le cas du « facial coding », Datakalab s'inspire du modèle de Ekman pour identifier les émotions. Elle part « d'indices¹¹³ » soit de signes exprimés, ici sur le visage, identifiés et choisis comme étant ceux de l'émotion. Ekman est l'un des premiers chercheurs à définir précisément la détection des « micro-expressions ». Il met en évidence les muscles faciaux sollicités pour exprimer des émotions. Le psychologue décompose ces « micro-mouvements » par des « unités d'action¹¹⁴ ». Le *Facial Action Coding System* qu'il développe en recense 46, qui peuvent être combinées pour être associées à des émotions spécifiques. Chaque « unité d'action » dispose d'un numéro, par exemple, le numéro 6 correspond à la « remontée des joues », le numéro 12 à « l'étirement du coin des lèvres ». Combinées, ces deux « unités d'actions » correspondent à la « joie », l'une des « émotions universelles » de Ekman. Ici, on voit qu'on part « d'indices¹¹⁵ », de « traces » des signes exprimés sur le visage de l'émotion. Captées, ces « inscriptions¹¹⁶ » sont ici qualifiées « d'unités d'action ». Combinées, elles deviennent, presque « magiquement », « la joie ». Aujourd'hui, cette méthode est la plus reprise et adaptée pour détecter les expressions faciales et Datakalab utilise un dérivé de ce modèle.

¹¹³ *Op. Cit.*

¹¹⁴ Mouvements des muscles faciaux

¹¹⁵ *Op. Cit.*

¹¹⁶ *Op. Cit.*

Cette mesure permet l'évaluation des contenus audiovisuels tels que les films, séries, jeux vidéo, publicité, conférences, en filmant les spectateurs. Les algorithmes utilisés en fonction des sociétés sont plus ou moins matures et sophistiqués, certains prennent par exemple en compte le mouvement de la tête, son inclinaison, les variations de luminosité, et les différences culturelles. Pour donner du sens à la donnée brute, différentes médiations sont à l'œuvre : traductions mathématiques et littéraires, déplacement de sens, de symboles, de compétences. Dans le cas de Datakalab, on ne devrait pas tant parler de « mesure d'émotions » que de « ressentis » – soit des signes *exprimés* de l'émotion – détectées sur suffisamment de panélistes. Datakalab le sait, ne s'en cache pas, et l'explique même aux clients qui cherchent à comprendre en détail les chiffres fournis. Elle souhaite aussi que les conditions de l'étude ne biaisent pas ses résultats. Concrètement, elle dispose aujourd'hui d'un algorithme qui estime la « géométrie du visage¹¹⁷ » et détecte les « expressions faciales ». C'est là que s'opère « l'inscription ». On sait alors ce qui se passe sur le visage du panéliste, par exemple s'il lève ses sourcils, plisse ses yeux... Cependant, de telles informations sont encore loin d'intéresser les clients de Datakalab. Les ingénieurs traitent alors cette donnée issue de l'algorithme. Ils utilisent un système que l'on peut vulgariser comme étant composé de « seuils¹¹⁸ » et de moyennes pour savoir à partir de quel moment la remontée du coin des lèvres par exemple peut être assimilée à de la « joie ». Ces « seuils » permettraient alors de savoir à partir de quels critères l'algorithme considère qu'il s'agit de « vraies émotions¹¹⁹ ».

1.3.1.1 « Scripturer » la donnée pour la rendre péremptoire

Ces données doivent traduire parfaitement une réalité propre à la situation de consommation pour apporter au client un savoir inouï. Datakalab veut rendre ses données utilisables par les entreprises et reconnues comme outil d'aide à la décision. Pour passer de « données brutes », absconses, à des « données intelligibles », rendre une herméneutique « divine » « lisible » pour reprendre Souchier¹²⁰, il faut « scripturer » les données : « L'homme

¹¹⁷ ANNEXE 21, Sélection de mails envoyés aux 2000 abonnés à la « newsletter » Datakalab, Point 2, « je traduis : *This approach allows to finely estimate the geometry of the face* »

¹¹⁸ Paliers fixés par les ingénieurs de Datakalab à partir desquels les algorithmes considèrent que les données captées indiquent une émotion

¹¹⁹ Lors de plusieurs discussions en interne et de notre entretien, Lucas Fischer s'engage à rendre un travail vrai d'un point de vue scientifique, et avec un gage de qualité dans son rendu : « Je veux [...] donner [aux clients] une information pertinente et vraie. Je ne veux pas leur donner des seuils plus bas mais qui laisseraient passer plus d'erreurs. J'essaye d'être très attentif sur la qualité de la data ».

¹²⁰ *Traces numériques* « De la production à l'interprétation », « Voir le Web et deviner le monde », « La "cartographie" » au risque de l'histoire de l'écriture », Emmanuel Souchier, p. 213-234 « disponible à : <https://books.openedition.org/editions-cnrs/21786?lang=en> »

doit [...] en organiser l'agencement, en définir les limites, en identifier les symboles, en ordonner la lecture et finalement en élaborer la syntaxe¹²¹ ». On donne à voir les réactions des consommateurs via un cadre – celui d'une présentation qui contient des graphes, des scores, des nombres. Datakalab donne sa vision de l'émotion en représentant ces données « prémâchées » sous forme de courbes, d'histogrammes, de tableaux.



« Tracés » données à voir par Datakalab dans ses présentations de cas clients

Les données émotionnelles peuvent être ainsi comparées à des dispositifs performatifs qui font voir la réalité sous un certain prisme, voire rendent visible une réalité intangible. Cette traduction donne une vue synthétique, choisie pour être intelligible. L'ingénieur et le commercial opèrent tous les deux des transformations sémiotiques lorsqu'ils transforment les données brutes en graphes. Ces processus de construction de la donnée et de mise en forme ont un effet d'irréversibilité. Comme l'explique Alexandra Ducrot dans sa recherche¹²² sur les technologies « big data », ils « donnent à voir la réalité sous une certaine forme, ils organisent l'information selon un certain modèle, fruit de plusieurs procédés de traduction, de modification, d'ajustement¹²³ ». Ces représentations s'ancrent dans l'esprit du décideur. Il devient alors difficile de se représenter autrement cette donnée émotionnelle, ce qui la rend péremptoire. Cela permet d'imposer un savoir-lire, une fois que Datakalab donne les clés d'interprétation de cette donnée aux entreprises. La start-up doit alors évoluer de sa conception ingénieure de la donnée à celle d'un communicant.

Pour traiter les « traces », les ingénieurs de Datakalab sélectionnent eux-mêmes les conditions de traitement de la donnée pour ensuite les interpréter : « Un processus commun de

¹²¹ *Ibid.*

¹²² *Op. Cit.*

¹²³ *Ibid.*

sélectivité émerge : tout n'est pas pris en compte et traité par l'informatique. Tout n'est pas perçu par le lecteur et l'interprète des "traces". De chaque côté, des tris s'opèrent qui produisent des effets de visibilité différenciés, allant de l'invisible à l'impression d'évidence existentielle et interprétative¹²⁴ ». Les acteurs choisissent donc la donnée qu'ils donnent finalement à voir et à interpréter. Béatrice Galinon-Méléneq explique ce biais, les résultats sont présentés comme révélateurs d'une « intelligibilité du monde ». On les promeut comme neutres, alors qu'ils sont – malgré toute la bonne volonté des ingénieurs à retranscrire la réalité de la manière la plus exacte – subjectifs et soumis à différents « rouages ». Ces processus de traitement constituent selon Béatrice Galinon-Méléneq des « commodités cognitives et culturelles pour parler le monde ». Or, Souchier se demande¹²⁵ « quels biais instaure la "traduction" de propriétés mathématiques en objets visuels, produits à partir de données elles-mêmes supposées rendre compte du "social" en ce qu'elles sont notamment qualifiées de "traces d'usage" » ?

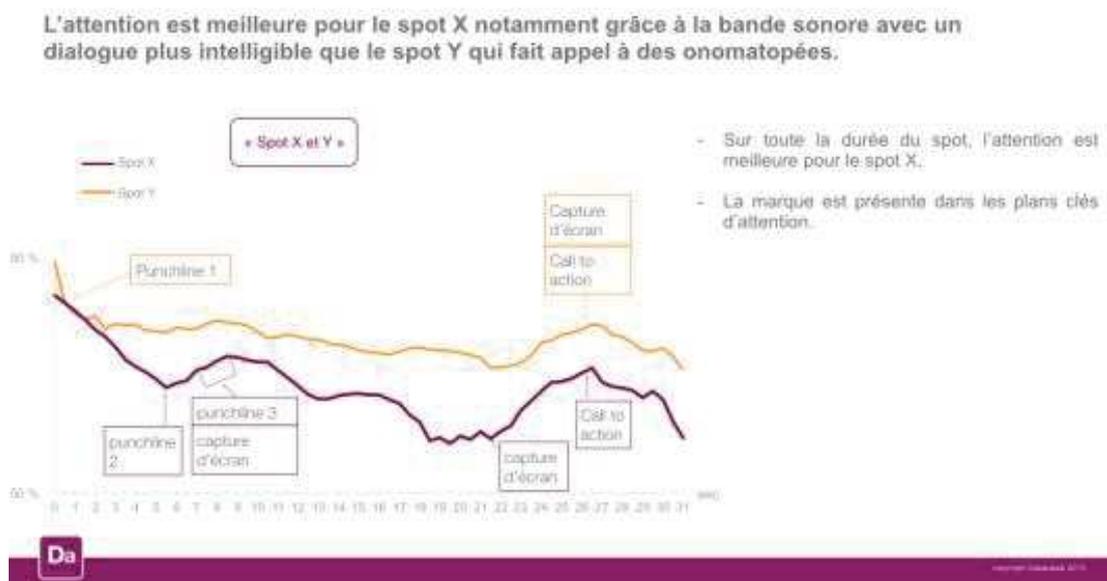
1.3.1.2 Traductions du langage ingénieur au langage commercial

Le passage du langage de l'ingénieur à celui du commercial constitue aussi une traduction, voire une interprétation. Ainsi, les « unités d'action » deviennent « la joie », qui elle-même ensuite combinée avec celle de la « surprise » donne un score « d'adhésion », soit un pourcentage de personnes ayant ressenti « au moins une émotion positive » durant le visionnage du contenu. Ici, des « transmutations sémiotiques » sont à l'œuvre. Jeanneret les définit comme « opérations qui consistent à transformer la nature des signes et leur statut grâce à des conversions et transcriptions en partie automatisées ». Datakalab opère des médiations algorithmiques en passant d'une rationalité ingénieure à une rationalité gestionnaire pour aider les décideurs. Les Project Managers se servent parfois de leurs propres émotions et intuitions pour interpréter cette donnée émotionnelle. Par exemple, des publicités peuvent être améliorées en retirant du texte présent en trop grande quantité, ou en opérant des choix de messages pour ne pas multiplier les points de focalisation. Le chiffre est utilisé pour légitimer la compétence humaine. Datakalab se sert de sa propre construction pour interpréter. Elle est à la fois acteur

¹²⁴ Galinon-Méléneq, Béatrice, Zlitni, Sami, *Traces numériques* : « De la production à l'interprétation », « L'Homme-trace, producteur de traces numériques », p. 7-19

¹²⁵ *Op. Cit.* Souchier, Emmanuël, « Voir le Web et deviner le monde, La "cartographie" au risque de l'histoire de l'écriture », pages 213-234, CNRS éditions, 2013 « disponible à l'adresse : <https://books.openedition.org/editionscnrs/21786?lang=en#ftn3> »

et interprète, réalise « l'écriture et la lecture¹²⁶ ». Les experts de Datakalab prennent ici le rôle du « devin sumérien », le « bârû¹²⁷ » étudié par Souchier. Ce dernier « interroge », « consulte » ou « enquête », ici à partir de données qu'il a lui-même déjà construites. La « data » permet alors de dépasser l'intuition, de la moduler, voire même de l'infirmer. Ce Project Manager lie, compare pour déduire. Au même titre que « les technologues¹²⁸ » de Souchier, ces communicants agissent comme « les devins de l'Antiquité, ils créent un système de cohérence entre des signes¹²⁹ ». Ingénieurs et communicants sont à la fois producteurs et interprètes des signes qu'ils utilisent, et pourtant, ils laissent croire à la neutralité et à l'objectivité de leur propos (sans pour autant *y faire croire*). Parfois des difficultés se posent. La « courbe d'attention » par exemple est le « tracé » choisi pour montrer le pourcentage de personnes qui ont été attentives devant le contenu en fonction des secondes. Des « captures d'écran » sont prises et positionnées en haut des « pics d'attention ».



Support de présentation anonymisé – Courbes d'attention

Cette sélection de visuel doit être représentative de la séquence. Le Project Manager doit aussi y retranscrire l'action. Il insère alors une phrase que Datakalab nomme « punchline ». Dans certains cas, ces choix sont complexes. Le « pic d'attention » peut s'étendre sur plusieurs secondes. Le Project Manager doit utiliser son bon sens pour se demander quel élément a

¹²⁶ *Traces numériques*, « De la production à l'interprétation », « Voir le Web et deviner le monde », « La "cartographie" » au risque de l'histoire de l'écriture », Emmanuel Souchier, p. 213-234, « disponible à : <https://books.openedition.org/editions-cnrs/21786?lang=en> »

¹²⁷ *Ibid.*

¹²⁸ *Ibid.*

¹²⁹ *Ibid.*

véritablement suscité l'émotion du public, et choisir le bon visuel ainsi que la bonne phrase. Cette difficulté a ensuite été résolue par Datakalab qui a trouvé un système dynamique où la vidéo est lue en même temps que la courbe, ce qui permet de retranscrire la complexité contextuelle. Des traductions d'un langage ingénieur en langage gestionnaire et inversement s'opèrent ainsi en boucle et semblent nécessaires pour rendre ces données « intelligibles ».

1.3.1.3 Datakalab opère un « acte sémiurgique¹³⁰ »

Datakalab projette la réalité sous un cadre délimité et construit, celui du test avec des panélistes. Elle est libre de sa création, et rares sont les personnes qui peuvent juger du bien fondé des conditions d'entrée et des « seuils » choisis dans ses paramètres, voire même avant cela, des « indices » définis comme étant ceux de l'émotion. En d'autres termes, seuls les ingénieurs connaissent les limites de la « détection des émotions ». On peut alors affirmer que la création de Datakalab revêt une dimension sémiurgique. En formalisant la confusion émotionnelle, les ingénieurs semblent dépasser les capacités humaines communes et déborder toute rationalité. Souchier souligne ainsi cette puissance : « L'acte fait apparaître ; c'est un acte de pouvoir ». Le sémiurge « donne à voir et régule la signification », il donne forme à ce corps créé dans une dimension qui frôle le démiurgique. Datakalab impose alors son savoir-lire. La start-up doit à la fois s'inscrire dans un système de « signes », de représentations connu, acquis et en sortir de par l'inouï du sujet. Les carrés des « métriques » délimitent l'objet, le score, ce qui traduit des opérations de choix et de synthèses opérés, mais surtout, donne une impression de maîtrise des événements. On conjure l'imprévisible. Dans le cas des courbes et histogrammes que Datakalab donne à voir, on peut même penser à un pouvoir prédictif, transformant une donnée indomptable, irrationnelle et complexe en savoir maîtrisé. Datakalab étend son « acte sémiurgique » en allant jusqu'à inventer de nouveaux postes qui jouent sur l'imaginaire de l'émotion, Frank Tapiro est par exemple Chief Emotion Officer¹³¹. On crée de nouveaux noms de métiers pour s'adapter au nouveau contexte culturel mais surtout, le façonner. Ce phénomène est cristallisé par la reprise du terme « brain tech » et la création des termes « feel data© », « data telling ». Emmanuel Souchier affirme même « ce microcosme est marqué du

¹³⁰ Epstein, Mikhail, "Semiurgy: From Language Analysis to Language Synthesis". Russian Journal of Communication ", 2008, *Semiurgy is the art of creating new signs and sign systems, as opposed to semiotics as the science of signs, and rhetoric as the effective usage of signs. The word 'semiurge' would mean an artisan of signs, the demiurge is the creator of the world.* « Je traduis : La sémiurgie désigne l'art de créer de nouveaux signes et autre systèmes, par opposition à la science des signes, et à la rhétorique en tant qu'utilisation efficace des signes. Le terme "sémiurge" signifierait un artisan des signes, le démiurge est le créateur du monde »

¹³¹ ANNEXE 2, Pitch Frank Tapiro – Challenge *Start-me up*, KPMG & Fevad

sceau de son pouvoir¹³² ». En inventant de nouveaux termes qui viennent enrichir le trou noir suscité par le « big data » Datakalab marquerait du « sceau de son pouvoir » les « artefacts » qu'elle construit. Ces termes disposent donc de la puissance communicationnelle de la technique, de l'innovation, la production et de la performance. Ainsi, toutes les données façonnées par les ingénieurs et ensuite mises en forme par les Project Managers peuvent faire autorité, et être considérées par les décideurs comme naturelles, sans prendre en compte leur « construit communicationnel¹³³ ».

1.3.1.4 Une « mesure » complexe des émotions

Jusqu'où les professionnels peuvent-ils mesurer l'émotion ? À l'heure actuelle, la technologie peut certes appréhender les « signes exprimés de l'émotion » soit les « ressentis », mais ne peut pas les prédire précisément. Leur intensité et leur expression dépendent aussi de tout un ensemble de facteurs relatifs à l'individu. Datakalab ne cache pas cette complexité propre à chacun. Au contraire, les membres de l'équipe en parlent volontiers. Kevin Bailly, « Head of Research » de Datakalab m'explique en entretien¹³⁴ : « Les vraies micro-expressions, l'algorithme ne va pas les capter. Il aurait le temps de les voir, mais il faut ensuite être capable de les différencier du bruit¹³⁵, soit les pics qui n'ont pas de sens. Nos bases de données sont souvent capturées dans de bonnes conditions. Ce n'est pas le cas de ce qu'on va récupérer des personnes qui font ça chez elles avec leur webcam, qui peut être différemment orientée, elles peuvent faire ça dans le noir, surexposées si elles sont devant leur fenêtre ». Si de nombreuses méthodes issues du neuromarketing prétendent « mesurer les émotions », l'adaptation du « facial coding » de Ekman constitue selon les chercheurs la plus simple à mettre en place, et l'une des plus utilisées. Les entretiens m'ont permis de me rendre compte que certains professionnels croyaient une « mesure » totale des émotions, alors que d'autres prenaient plus de distance, évoquant plutôt la « retranscription des signes exprimés de l'émotion ». Ce phénomène peut rappeler la définition du « simulacre¹³⁶ » de Baudrillard, confirmé par Anne

¹³² Souchier, Emmanuël, *Traces numériques* « De la production à l'interprétation », « Voir le Web et deviner le monde », « La "cartographie" » au risque de l'histoire de l'écriture », p. 213-234 (Ghitalla, 2008), « disponible à : <https://books.openedition.org/editions-cnrs/21786?lang=en> »

¹³³ *Op. Cit.* Jeanneret, Yves, p.42

¹³⁴ ANNEXE 10, Entretien avec Kevin Bailly, Head of Research chez Datakalab, 19/07/2019, Paris 17^{ème}

¹³⁵ ANNEXE 12, Entretien avec Éric Daoud, Ex Data Scientist & Machine Learning Engineer chez Datakalab, 04/06/2019, Paris 17^{ème} : « Le bruit correspond aux pics qui n'ont pas de sens. Si par exemple la personne se tourne, l'algorithme pense que ça génère un pic, mais ce pic n'est pas une vraie émotion. C'est juste un bruit, quelque chose auquel l'algorithme ne s'attendait pas, et du coup il croit que ça a un sens et nous dit que c'est un pic. Un signal qui a du bruit, c'est un signal qui a plein de points qui n'ont pas de sens »

¹³⁶ Baudrillard, Jean, *Le Crime parfait*, Éditions Galilée, 1995

Sauvageot dans l'article « Quand Baudrillard nous aide à penser le monde contemporain¹³⁷ » : « L'ordinateur permet d'opérer la substitution du réel par son double hyper-réel, plus réaliste, plus parfait, plus réel que le réel ». On touche ici au « crime parfait¹³⁸ », celui de « la résolution anticipée du monde par clonage de la réalité et l'extermination du réel par son double ». Emmanuel Souchier explique que le processus fonctionnerait « selon une boucle de croyance [...] Je dois croire en mon activité, puisqu'elle me fait vivre ».

Pour conclure cette première partie, la « feel data© » s'inscrit en fait dans toute une imbrication économique, politique, technologique et sociale qu'elle contribue aussi à créer. Cette donnée entend lier deux domaines *a priori* inconciliables, l'humain, naturel et la machine, artificielle. On note donc une antithèse présente à chaque fois dans les différents termes : « mesurer l'émotion », « affective computing », donnée émotionnelle. Cet aspect contre-intuitif découle de la nature typiquement humaine, spirituelle de « l'émotion ». C'est aussi pour cette raison que Datakalab rencontre quelques difficultés pour faire passer cette promesse de rêve à la réalité. Elle convainc bien qu'il est possible de ramener la captation de l'émotion à un traitement computationnel en s'appuyant sur les imaginaires issus des systèmes de collecte qu'elle utilise. La première hypothèse a donc bien été validée et est même dépassée, puisqu'on oublie volontiers la distance avec laquelle on devrait appréhender la « mesure émotionnelle ». Les discours commerciaux s'appuient sur tous ces éléments pour présenter cette donnée comme naturelle, alors qu'elle relève d'un construit. La donnée émotionnelle peut alors être comparée à la « pointe de l'iceberg ». Or ce même construit est aujourd'hui donné à voir dans sa communication.

¹³⁷ Sauvageot Anne, « Quand Baudrillard nous aide à penser le monde contemporain », 24/01/2019

¹³⁸ *Op. Cit.*, Le « crime parfait » de la réalité « c'est celui d'une réalisation inconditionnelle du monde par l'actualisation de toutes les données, par transformation de tous nos actes, de tous les événements en informatique pure : la solution finale, la résolution anticipée du monde par clonage de la réalité et l'extermination du réel par son double »

2 ÉVOLUTION DES « DISCOURS D'ESCORTE » DE LA « FEEL DATA© »

En second lieu, je me demande si Datakalab maintient sa « promesse » de « mesure des émotions » malgré son repositionnement. Suite à la création d'algorithmes propriétaires de « facial coding », Datakalab est passée d'une expertise « conseil », axée sur des recommandations issues d'un savoir neuroscientifique à une expertise technologique. Un outil, le « bracelet connecté » a notamment cessé d'être utilisé. Les « discours d'escorte » tenus autour de la donnée émotionnelle ont évolué, faisant émerger ce qui semble être une nouvelle promesse. J'interroge ici ce renouveau.

2.1 Malgré la suppression de l'un de ses outils, Datakalab conserve sa promesse maïeutique

Les nouveaux chercheurs de Datakalab ont fini par mettre au point un algorithme propriétaire, capable d'estimer la géométrie du visage et dont les communicants affirment qu'il peut « mesurer les réactions ». Jusqu'alors, Datakalab se contentait d'effectuer de la mise en relation entre les annonceurs – ses clients et les sociétés de neuromarketing qui possédaient cette technologie. Or, avec ce nouvel outil propriétaire, la « bascule » de l'expertise « conseil » à l'expertise « technologique » prévue dès le lancement peut être enclenchée. Datakalab cesse alors d'utiliser et de vendre les études conduites avec le « bracelet connecté ». Ce faisant, début 2019 elle semble supprimer la « mesure de l'expérience » en boutique, se concentrant uniquement sur le traitement de données captées en ligne grâce aux panélistes. Datakalab se focalise sur ses algorithmes de « détection des expressions faciales » donc ne s'intéressant plus qu'au visage humain. Dans quelle mesure garde-t-on alors la même promesse de « mesure émotionnelle » tandis que l'outil est différent ? L'emploi du « bracelet » impliquait d'autres « découpes » du réel que les études de « facial coding ». On parlait de données physiologiques comme la « température extra-cutanée » alors que ce sont les « signes exprimés sur le visage de l'émotion » qu'on étudie en « facial coding ». Dans les deux cas, on s'appuie néanmoins sur une donnée brute, absconse en tant que telle, pour lui construire un sens et donner accès à une réalité immesurable autrement. Avant ou après la « bascule », la start-up promet toujours de « mesurer » l'immatériel et de rendre visible l'invisible. Elle « fait accoucher des ressentis » pour donner des « vérités consommateurs » à ses clients et s'inscrit dans une maïeutique. On pourrait même la comparer au philosophe qui vient chercher les

prisonniers dans « L'Allégorie de la caverne » de Platon. Ces signes laissent penser que la promesse n'a pas tellement évolué. Pourtant, une multitude de médiations se trouvent déjà à l'œuvre, ne serait-ce que dans la façon de capter les « traces ».

2.2 Conservation d'un triple pouvoir heuristique pragmatique et normatif¹³⁹

Que ce soit avant ou après son repositionnement, Datakalab dispose de différents types de pouvoirs, grâce à tous les mythes qui l'entourent. Alexandra Ducrot¹⁴⁰ affirme au sujet des consultants en technologie qu'ils disposent d'un « triple pouvoir heuristique, pragmatique et normatif ». « Heuristique » désigne une opération mentale propice à la connaissance, « c'est-à-dire qu'ils donnent une interprétation du réel ». « Pragmatique » prend le sens d'une aide « à la décision et à l'action concrète ». « Normatif » signifie enfin la capacité « d'induire une norme des pratiques professionnelles ». D'abord, Datakalab peut revendiquer un pouvoir heuristique puisque ces données émotionnelles permettent bien d'aller au-delà de l'intuition et de donner des clés d'accès à la performance. Elles guident les décideurs dans leur réflexion. Concernant leur pouvoir pragmatique, les « feel data© » constituent selon Datakalab des « outils d'aide à la décision ». Suite à leur interprétation, les directions prennent bien des décisions : par exemple enlever du texte sur une création publicitaire, agrandir le logo et autres modifications visuelles. Finalement le pouvoir normatif s'exerce dans la mesure où Datakalab s'efforce d'instituer sa donnée émotionnelle en tant « qu'indice » de la réception des publicités, nouvel outil dans la boîte à outil des décideurs. Datakalab cherche notamment à établir des normes pour réguler l'efficacité d'une publicité. Elle s'inscrit ici dans un long débat qui cherche à savoir si ce genre de règles existe. Le pouvoir normatif existe aussi *via* la création des « métriques » qui tentent de reproduire une réalité, de faire exister une construction du réel. Elles doivent être compréhensibles de tous les secteurs, et par tous les corps de métiers. En se multipliant, ces « traces » s'auto-légitiment petit à petit. Datakalab montre dans ses présentations qu'elle a pu collecter de nombreuses données pour créer des normes : « 80 annonces presse et affiches, 55 vidéos et films¹⁴¹ ». Elle les construit en faisant des moyennes à partir des scores observés. La start-up veut que ce construit soit lui-même interprété, croisé avec les autres « traces » captées par les instituts d'études. Dans ses normes, Datakalab combine les données de tous les

¹³⁹ *Op.Cit.*

¹⁴⁰ *Op. Cit.*

¹⁴¹ ANNEXE 20, Support de présentation PowerPoint de Datakalab, slide « Les normes »

secteurs¹⁴². Des articles circulent également sur l'éventuelle possibilité d'établir des normes en publicité. Certaines associations comme l'*Advertising Research Foundation*¹⁴³ sont même spécialistes du sujet. Datakalab exerce donc ainsi ce triple pouvoir heuristique pragmatique et normatif. Ces pouvoirs constituent alors des notions suffisamment conceptuelles pour ne pas avoir été impactées par le repositionnement de la start-up. Ils prennent suffisamment de hauteur théorique sur ces pratiques pour continuer d'être véridiques malgré un contexte remanié. En ce sens, on peut maintenir qu'il existe un cœur de promesse inébranlable malgré les séismes de son repositionnement.

2.3 Évolution des « discours d'escorte » de la « feel data© »

La start-up s'est repositionnée, délaissant sa « prestation conseil » en 2019. Cette expression renvoie aux recommandations fournies dans les présentations pour aider les annonceurs à optimiser leurs contenus de marques. Ces recommandations étaient rédigées à partir du savoir neuroscientifique analysé en première partie. La start-up ne revendique plus ces connaissances. On peut même aller jusqu'à dire qu'elle a effectué un changement assez conséquent dans l'ensemble de sa communication. Anne-Marie Gaultier, Ex Présidente, ainsi que la Key Account Director, Soukeina Hassanaly ont quitté la start-up début 2019. Ces deux communicantes étaient spécialisées dans la partie marketing/commerciale et rédigeaient jusqu'alors les recommandations. Elles laissent alors les rennes de Datakalab à deux ingénieurs centraliens, Xavier et Lucas Fischer. S'en suit un changement de communication : une valorisation de la technologie et une suppression rapide de la dimension « conseil ». Désormais, la majeure partie de la communication de Datakalab s'effectue en anglais : « newsletter », prises de paroles Facebook, LinkedIn... Cela lui permet de toucher une cible internationale. Si à son lancement, Datakalab se définissait en tant que « laboratoire conseil en neuromarketing », aujourd'hui, ces qualificatifs ont été repensés. Elle reste bien sûr un laboratoire, dans la mesure où des chercheurs y travaillent, font des découvertes et tentent de faire progresser la recherche en « computer vision », mais elle n'est plus spécialisée dans le « conseil », et n'emploie plus le terme de « neuromarketing » dans sa communication. Interrogé à ce sujet, Xavier Fischer, CEO de Datakalab, explique : « Il ne faut pas employer des termes qui font peur. Notre objectif, c'est

¹⁴² Comparer des « scores d'attention » du secteur automobile à celui du secteur du luxe peut paraître déroutant. Datakalab affirme au contraire que le consommateur fait face à une nuée de stimuli commerciaux, et qu'une affiche d'un secteur n'est pas uniquement en concurrence avec un message concurrent, mais avec l'ensemble des autres stimuli d'autres secteurs

¹⁴³ Association américaine dans la publicité et les médias fondée en 1936 qui cherche à améliorer les pratiques publicitaires

de dédramatiser¹⁴⁴ ». Datakalab semble ne plus autant jouer sur son « ethos neuroscientifique » qu'à ses débuts. Pourtant, il est difficile de construire un discours qui s'appuie sur les découvertes en neurosciences sans parler de neuromarketing. Datakalab a donc passé ses premières années à essayer de rassurer sur le fait qu'elle ne cherchait pas à manipuler les gens en ayant accès à leurs pensées intimes¹⁴⁵. Désormais, elle ne communique plus du tout sur ce terme à cause de sa portée symbolique qui peut être rattachée au courant behavioriste, qui considère le comportement de l'individu comme étant conditionné et réagissant à une suite de « stimuli ». De même, à ses débuts, Datakalab dit « mesurer » les différentes émotions de base de Ekman. C'est bien pour cette raison qu'elle prêche leur importance dans l'acte d'achat dans toutes les conférences. Désormais, l'émotion ne semble plus être le cœur de Datakalab, qui évoque le visage comme étant sa « nouvelle clé ». Toute la communication est donc réactualisée, et Datakalab modernise ses présentations. Elle change ses visuels, supprimant par exemple un nuage de logos pour le remplacer par un visuel des rues de Times Square inondées de publicités, illustrant les idéologies auxquelles elle semble se rattacher.



Ancien (à gauche) et nouveau (à droite) supports de présentation de Datakalab

¹⁴⁴ ANNEXE 6, Entretien avec Xavier Fischer, Co-fondateur et Chief Executive Officer de Datakalab, 26/07/2019, Paris 17^{ème}

¹⁴⁵ Datakalab précise systématiquement : « les résultats sont agrégés et on demande son accord au panéliste »

2.3.1 Réduction de la compétence humaine d'interprétation

De la suppression de cette expertise découle une réduction de la compétence humaine d'interprétation. Dans la « deuxième version » de Datakalab, le Project Manager ne fait plus autant de recommandations : il ne conseille plus les marques sur la manière d'améliorer leurs contenus. Désormais, la start-up fournit essentiellement des scores, des « tracés ». Avant la bascule, les communicants de Datakalab interprétaient la donnée. En réduisant la part de compétence humaine d'interprétation du Project Manager, on s'écarte de la prestation « conseil » pour se rapprocher d'un *faire loi* de la technologie. On conserve un savoir inné, absolu détecté par la machine. Emmanüel Souchier parle de « l'effacement des sources d'énonciation¹⁴⁶ » et de « négation de la médiation¹⁴⁷ ». Cet effacement participe à la construction du mythe de la neutralité et de l'objectivité des médiateurs : « Si je ne suis pas l'énonciateur, si les cadres d'énonciation me sont "naturels", alors la parole que je profère est une parole des dieux ou de l'univers¹⁴⁸ ». Derrière les scores donnés par Datakalab qui ne fait plus de recommandations s'opère alors une délégation d'une compétence humaine à la machine. Dans la « première version » de Datakalab, l'interprétation du « savoir » de la machine¹⁴⁹ était confrontée aux intuitions du communicant. Aujourd'hui, la start-up semble de plus en plus fournir les scores et « tracés » simplement accompagnés d'une prise de note descriptive, ce qui renforce la « boucle de croyance¹⁵⁰ » d'Emmanüel Souchier. Datakalab se place en « porte-parole » du savoir « divin » de l'algorithme, lui octroyant une certaine responsabilité du « savoir » produit. Ceux qui vulgarisent la donnée (les communicants) ne sont pas ceux qui la produisent (les ingénieurs). Sans « produire » le sens, le communicant le « subit » et se contente de le transmettre. Comme le prophète, ce sont « les signes ou les dieux qui parlent à travers lui¹⁵¹ ». Les données émotionnelles ne sont plus accompagnées d'un discours de « conseil » qui les enrobe. Elles n'en ont plus besoin, puisqu'elles existent en tant que telles. En ce sens les conditions d'énonciation de la donnée émotionnelle sont modulées. Selon Souchier et Jeanneret : « les discours ne se contentent pas d'accompagner, de l'extérieur, les objets

¹⁴⁶ Souchier, Emmanüel. *Voir le Web et deviner le monde : La « cartographie » au risque de l'histoire de l'écriture, Traces numériques : De la production à l'interprétation* Paris : CNRS Éditions, 2013, p.224, « disponible à l'adresse : « <file:///Users/clairerlatourre/Desktop/editions-cnrs-21786.pdf> »

¹⁴⁷ *Ibid.* p.225

¹⁴⁸ *Ibid.* p.224

¹⁴⁹ *L'expression semble antithétique.*

¹⁵⁰ *Ibid.*

¹⁵¹ *Ibid.* p.223

concernés. Ils les configurent. Ils opèrent en eux, les qualifient¹⁵² ». Comme ces données ne sont plus enrobées d'un discours aux prétentions neuroscientifiques, mais livrées en tant que telles, les données et les présentations ne sont plus appréhendées de la même manière. Pour conserver la valeur symbolique de ses rendus, Datakalab doit alors miser sur la construction d'une posture d'ingénieur expert, qui crée plus qu'il ne guide. Or comment et jusqu'où peut-on se montrer en train de créer, dès lors qu'on vient d'affirmer que Datakalab présentait sa donnée comme naturelle, pure et intacte ? Là repose le cœur du sujet. Les discours d'escorte ne revêtent pas simplement une fonction d'accompagnement de la promesse, ils semblent la forger. Mais si promesse et discours forment un tout indissociable, ne perd-on pas une part de la promesse en supprimant les discours qui enrobent les données émotionnelles, les recommandations ? Comment remplacer ce vide laissé dans la promesse elle-même ? Les réécritures vont alors sans doute la modifier, dans la mesure où « signifiant » et « signifié » semblent ici confondus.

2.3.2 Des « algorithmes au-delà du réel¹⁵³ » : construction d'une deuxième posture d'expert

Datakalab déplace son expertise. Elle passe d'une maîtrise en neurosciences à une expertise en « computer vision ». Le statut d'expert vient s'ajouter à celui d'ingénieur, déjà très valorisé dans la société. La strate permet une légitimité immédiate, péremptoire. Cette citation de Yves Jeanneret est souvent reprise : « Un métier, une compétence ne peuvent s'imposer sans communication ; leur valeur et leur image sont construites par un récit¹⁵⁴ ». Les prétentions de Datakalab ne sont appréhendées que par ses « discours d'escorte » qui l'instaurent en tant que « qu'experte en computer vision ». Dans les salons, on assiste à une « préfiguration¹⁵⁵ ». Datakalab place ses éléments visuels : kakémonos, bannières, de manière à attirer le public en nombre. Elle fait la démonstration des algorithmes visible en annexe¹⁵⁶. Les experts organisent « les conditions de la communication » pour que le public les voie directement.

¹⁵² Jeanneret, Yves, Souchier, Emmanuël, « Présentation ». Communication et langages, n°128, 2ème trimestre 2001. Dossier : Les T.I.C., objets de valeur. p. 33, « disponible à l'adresse : https://www.persee.fr/doc/AsPDF/colan_0336-1500_2001_num_128_1_3072.pdf »

¹⁵³ Expression de Frank Tapiro, « Pitch Datakalab » tenu dans les salons et devant les annonceurs, soit la présentation succincte de la start-up qui donne envie aux prospects de se renseigner

¹⁵⁴ Tétu, Jean-François, « Yves Jeanneret, *Penser la trivialité*. Volume 1 : *La vie triviale des êtres culturels* », Questions de communication, 2009, mis en ligne le 17/01/2012, « disponible à l'adresse : <https://effeuillage-la-revue.fr/portfolio-item/la-trivialite-pensee-par-yves-jeanneret/> »

¹⁵⁵ Jeanneret, Yves, *Critique de la trivialité*. « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, p. 14 « L'attitude dont certains acteurs disposent pour organiser les conditions de la communication de manière plus ou moins stricte »

¹⁵⁶ ANNEXE 22, Sélection de prise de parole Facebook, Prise de parole du 05/06/2019 sur les attributs que le nouvel « algorithme » sait détecter : présence de lunettes, de maquillage, bouche ouverte...

Maingueneau qualifie la scénographie de « scène de parole¹⁵⁷ » c'est-à-dire « le processus par lequel une scénographie légitime un énoncé qui, en retour doit la légitimer ». La start-up met aussi en avant le parcours prestigieux de ses ingénieurs pour attirer la confiance avec les expressions « issu de Centrale », « diplômé de Polytechnique » dans ses présentations.



Présentation de l'équipe Datakalab en 2019

Datakalab présente leurs réalisations : « Les travaux de Kevin Bailly ont reçu le Premier Prix à deux reprises au concours *Facial Expression Recognition and Analysis* contre le M.I.T., Cambridge et l'université de Californie à San Diego ». Elle met en scène la participation de ses chercheurs aux événements internationaux¹⁵⁸. Le caractère inouï des algorithmes est présenté *via* des superlatifs. Frank Tapiro explique dans les salons : « Nos chercheurs ont gagné tous les concours internationaux, tous ! Ce sont les meilleurs chercheurs au monde. [...] Vous avez vu notre démo ? L'algorithme qu'ils ont réussi à créer est au-delà du réel¹⁵⁹ ». Cet « émerveillement » des acteurs pour reprendre Souchier porte la croissance de la start-up, dans la mesure où il permet lui-même de créer de « l'attention », et de susciter des « réactions » positives. Datakalab applique ici ses propres enseignements. On éveille l'intérêt en jouant sur le suspense, la tension, la nouveauté, le jamais-vu. Ce discours de présentation des algorithmes

¹⁵⁷ *Op. Cit.* Alsafar, Ali, « Ethos discursif et construction des rapports intersubjectifs dans les professions de foi des élections présidentielles de 2007 et de 2012 », Linguistique, Université Paul Valéry - Montpellier III, 2014, p. 18 « ethos et scène d'énonciation » selon Maingueneau

¹⁵⁸ ANNEXE 21, Sélection de mails envoyés aux 2000 abonnés à la « newsletter » Datakalab, Mail du 23 juillet 2019 sur la sélection à ICCV 2019

¹⁵⁹ *Op. Cit.*

revêt plusieurs fonctions. Datakalab joue sur sa posture didactique, qui forme et informe ses clients, les invitant à l'action de « tester la démonstration ». Ces champs lexicaux accroissent la nouveauté du spectacle, construisent la figure d'expert en « computer vision » et permettent à Datakalab de légitimer les « discours d'escorte » de ses études émotionnelles. Datakalab construit aussi son crédit grâce aux mails qu'elle envoie tous les mois à ses 2 000 abonnés¹⁶⁰. Elle y intègre une présentation en bas de page : « Datakalab a développé ses propres algorithmes et plateformes et analysé plus de 10 millions d'images d'expressions faciales collectées lors de mesures d'affiches, d'annonces presse, de vidéos, de films et de conférences¹⁶¹ ». La start-up met en avant son expertise grâce au développement de ses outils propriétaires. « 10 millions d'images » est un nombre difficile à appréhender tant il semble immense. Elle rappelle aussi la diversité de son expertise avec l'accumulation « mesures d'affiches, annonces presse, vidéos, films et conférences ». Elle cite ses principaux clients comme « TF1, LVMH, SNCF, Disney, Médiamétrie, Exterion Media ». Le choix de ces sociétés, stratégique, mêle annonceurs, afficheurs et instituts d'études, rendant le travail avec Datakalab envisageable pour ces interlocuteurs. Elle assoit ainsi sa légitimité et pose les bases de la construction d'un « ethos algorithmique » qui porte sa promesse.

2.3.3 Datakalab révélée par la création d'algorithmes propriétaires de « facial coding »

Toute la communication de Datakalab s'axe alors sur la monstration de sa technologie pour se construire un « ethos algorithmique ». Elle montre son expertise *via* la plateforme propriétaire de récolte de vidéos des panélistes. Certains annonceurs peuvent même se rendre sur cette plateforme et accéder directement aux résultats de leur affiche ou de leur spot. Datakalab communique sur la démonstration des algorithmes, elle présente la technologie à l'œuvre, en train d'agir.

¹⁶⁰ ANNEXE 21, Sélection de mails envoyés aux 2000 abonnés à la « newsletter » Datakalab

¹⁶¹ *Ibid.*



Présentation de l'algorithme et de la plateforme propriétaires en 2019

La start-up montre son apport théorique avec les papiers de recherche¹⁶², son algorithme d'estimation de la géométrie du visage, la plateforme de traitement des vidéos. Avant, tout ce travail de traitement par les algorithmes était masqué dans la communication, parce que non propriétaire. On tenait des discours autour du savoir neuroscientifique, mais on traitait tout de même la donnée. Mieux, on réunissait les conditions de la naissance de la technologie propriétaire en embauchant de nouveaux chercheurs. Alors que ce traitement et le processus de construction de sens étaient « couvés », cette construction est rendue visible, révélée. En ce sens Datalab est allée elle-même jusqu'à jouir de sa propre maïeutique. Modélisé par des flèches¹⁶³, on montre les facilités permises par cette technologie : récolte de données, traitement, extraction. Datalab accomplit une « prouesse technique¹⁶⁴ » : « Cet algorithme est plus puissant que celui de Facebook ou SenseTime et permet même de prendre en compte des aléas comme une main devant la bouche, un contre-jour ou une tête qui se tourne¹⁶⁵ » explique Frank Tapiro.

2.3.4 Valorisation de l'imaginaire scientifique dans la rhétorique commerciale

Datalab doit précisément se prévaloir d'une vérité scientifique alors qu'elle a toujours traité des émotions. Elle persuade et suscite l'adhésion grâce aux valeurs rationnelles qu'elle

¹⁶² *Ibid.* ANNEXE 21

¹⁶³ ANNEXE 20, Présentation de la démarche de respect de la loi de la protection de la donnée personnelle

¹⁶⁴ Candel, Etienne, dans son HDR « Textualiser les interfaces. Opérativité et épistémologie d'une requalification », souligne la vision utopique et technocentrée à l'égard des interfaces techniques : « se déploie, de façon à la fois très explicite et discrète, l'idée générale que recouvrent ces interfaces dans cet usage spécifique : une prouesse technique, un exploit, placés au profit et au service des utilisateurs, et porteurs de promesses inattendues et inespérées pour lui », p. 23

¹⁶⁵ Pitch Frank Tapiro – Challenge Start-me up! KPMG & Fevad

incarne en s'appuyant sur les mathématiques. La science fait loi et constitue à elle seule un « être culturel¹⁶⁶ » au sens donné par Jeanneret, soit un « ensemble d'idées et de valeurs qui incarne un objet de la culture dans une société tout en se transformant constamment à partir de la circulation des textes, des objets et des signes ». Le nombre et le calcul disposent de pouvoirs symboliques qui font autorité, comme dans la « logique du père Noël¹⁶⁷ » de Jean Baudrillard, soit une représentation sociale que les individus décident de croire ou non. Alors que le calcul relève du rationnel, il porte aussi en lui la capacité à nous faire tomber dans une « régression infantile¹⁶⁸ ». La réflexion que Baudrillard applique à la publicité peut être transposée sur la « gouvernance algorithmique ». La marque et ses discours font exister ce qui est énoncé. On adhère d'autant plus à ces technologies qu'elles disposent d'une origine rationnelle, scientifique, et donc d'un caractère infaillible fantasmé. Datakalab s'appuie alors sur la « rationalité occidentale¹⁶⁹ » et sur la foi en la base de données. Yves Jeanneret évoque la capacité montrée par l'informatique « de se muer, de simple procédure de calcul qu'était à l'origine la "computer science", en une discipline de l'archive¹⁷⁰ ». Tous ces imaginaires s'acheminent vers le progrès, l'impératif d'innovation de notre société. Datakalab deviendrait alors une instance énonciative à la « rationalité algorithmique¹⁷¹ ». Il s'avère difficile de soutenir que la promesse est restée en tout point la même, dans la mesure où ces discours sur les chercheurs, sur la plateforme et les algorithmes propriétaires n'étaient pas présents dans ce que Datakalab nomme sa « première version » étudiée en première partie. Datakalab s'appuie ainsi sur le crédit de ses débuts pour moduler sa promesse initiale et augmenter sa rentabilité.

2.4 Maîtriser le réel *via* la « data »

Datakalab appréhende désormais les « réactions » des consommateurs dans leur dimension complexe. En passant du terme « émotions » à celui de « réactions », on perd la dimension flottante qui *transporte au-delà* pour se concentrer sur un terme qui peut sembler

¹⁶⁶ *Op. Cit.* Lexique

¹⁶⁷ Baudrillard, Jean, *Le Système des objets*, Gallimard, 1968, l'analyse de la « foi » publicitaire, mis en rapport avec celle des enfants qui « croient » au Père Noël.

¹⁶⁸ Aubrun, Frédéric, Bihay, Thomas « Publicité en série : lorsque la marque se raconte sur le Web », *Communication & langages* 2015/3, N° 185, p. 137, Disponible à l'adresse : <https://hal.archives-ouvertes.fr/hal-01569933/document>

¹⁶⁹ Kuçuradi, Ioanna, revue *Diogène* : « on peut comprendre cette notion » comme une forme particulière d'expliquer et justifier les choses », lorsque l'Homme explique les phénomènes naturels non pas « théologiquement » ou « métaphysiquement », mais par l'observation et l'expérimentation, ou par la « science positive » (« Rationalité' » et « rationalités » dans le cadre du débat modernisme – post-modernisme », 2003, n° 202, pages. 12-20)

¹⁷⁰ *Ibid.* Pages 32-33

¹⁷¹ SADIN Éric, *La Vie algorithmique* « Critique de la raison numérique »

plus froid, mécanique. Le terme de « réactions » peut même être évoqué au sujet d'une machine. Néanmoins, il insiste davantage sur le fait que la personne réagit devant un élément, ce qui rend mieux compte de la réalité du travail de Datakalab. Ses algorithmes sont désormais capables d'estimer la géométrie du visage, « l'orientation de la tête », le « clignement des yeux », ainsi que de détecter des attributs comme le sexe, la présence d'accessoires, de maquillage. Lucas Fischer, co-fondateur de Datakalab m'éclaire durant notre entretien¹⁷² : « Dans notre promesse, on a décidé de faire évoluer l'offre, et de ne pas se cantonner aux émotions. On explique maintenant que les émotions ne sont qu'une partie de l'information. On va essayer de tout détecter sur le visage ». Datakalab a désormais une vision plus large et relaye l'émotion au second plan, alors qu'elle a toujours joué le premier rôle dans ses discours. Mais c'est aussi en ce sens qu'elle reste crédible et légitime. La start-up ne « bascule » pas du tout au tout, elle module simplement son expertise. Ce nouveau discours vient même appuyer la vision initiale de « redonner visage humain à la data », dans la mesure où elle affirme désormais vouloir comprendre tous les « signes » sur le visage. Xavier Fischer explique en présentation : « Si je vous demande de vous souvenir de toutes les émotions que vous avez ressenties seconde par seconde, vous n'y arriverez pas car nos cerveaux n'ont pas la capacité de stockage ». Les ordinateurs ont la capacité de stockage de nos émotions et « réactions » supérieures à celles de l'humain. Datakalab s'inscrit dans la logique de délégation de pouvoir à la machine. L'interprétation se veut pluridimensionnelle, elle tente de prendre en compte le contexte de l'étude, le contenu présenté, les tendances dans lesquelles tout s'imbrique. La rhétorique mentionne une captation absolue des « réactions », et donne à voir le monde dans sa globalité seconde par seconde. Cette réalité s'avère encore plus précise que ce dont le cerveau peut se souvenir. Les algorithmes captent des données qui semblent infinies, et c'est bien pour cette raison qu'elles apparaissent légitimes pour cerner l'expérience dans sa globalité. Or, le livrable – des scores, « tracés » comme les histogrammes, courbes – réduit au cadre d'interprétation ces premières « traces » et simplifie avec des opérations comme des moyennes la réalité presque infinie captée initialement *via* les données brutes. Le traitement computationnel fait donc d'abord exploser le nombre « d'inscriptions¹⁷³ » captées : inclinaison de la tête, remontée du coin des lèvres, des sourcils, clignement des yeux... pour ensuite agréger ces scores et donner une version intelligible en réduisant ces objets à une représentation, simple « trace ». Celle-ci cherche l'adhésion du plus grand nombre sans garder mémoire de ces processus de création. Les chiffres et méthodes utilisés ne constituent que des outils d'appréhension du réel. On peut

¹⁷² *Op. Cit.*

¹⁷³ *Op. Cit.*

établir une comparaison avec le philosophe qui tente de cerner le réel avec ses propres outils, les mots. Dans *La vie algorithmique* « Critique de la raison numérique », Éric Sadin souligne ce désir d'une connaissance infinie et constamment enrichie : « Des technologies informationnelles [...] imposent un mode de rationalité fondé sur la définition chiffrée de toute situation et sur une maîtrise indéfiniment accrue du cours des choses¹⁷⁴ ». Il évoque une « raison algorithmique¹⁷⁵ », soit un recours systématique à la data, qui vient sans cesse impacter les décisions dans les sphères managériales, politiques, sociales et économiques. Comme lors de l'apparition de chaque nouvelle technologie, des discours révolutionnaires sont tenus. Grâce au renouveau technologique, on accède à un nouveau savoir, jusque-là inaccessible, inatteignable et captivant.

2.4.1.1 Manque de recul sur la « raison algorithmique¹⁷⁶ »

Néanmoins, Antoinette Rouvroy souligne le manque de recul sur la notion de « raison algorithmique ». Les acteurs font confiance à la donnée plutôt qu'à leurs intuitions. Lorsqu'on traite l'émotion, il peut paraître étonnant de se fier aux chiffres plutôt qu'à son instinct. D'après elle « l'opérationnalité sans épreuve du comportementalisme numérique¹⁷⁷ » n'est pas assez interrogée. Or c'est aussi parce que la « data » semble quasiment infinie qu'on la pense qualifiée pour cerner le réel. Cette prétention résonne avec le fantasme du nombre et du calcul et avec les imaginaires du numérique. L'utopie des modèles scientifiques et du nombre se réalise du fait même d'être évoquée dans des discours qui grandissent. Comme l'explique Alexandra Ducrot dans sa recherche¹⁷⁸, nous devrions être alertés par les prétentions de la donnée à cerner le réel dans sa totalité. Cette prise de recul s'avère aussi cruciale dans la mesure où les discours tenus dans ce contexte opèrent de manière performative. C'est bien en ce sens que la nouvelle promesse se trouve forgée et modulée par les mêmes discours qu'elle a créés et véhiculés. On peut alors interroger le langage lui-même.

¹⁷⁴ Sadin, Éric, *La vie algorithmique* « Critique de la raison numérique », quatrième de couverture

¹⁷⁵ *Ibid.*

¹⁷⁶ *Op. Cit.*, Sadin

¹⁷⁷ Rouvroy, Antoinette, « Pour une défense de l'éprouvante inopérationnalité du droit face à l'opérationnalité sans épreuve du comportementalisme numérique », *Dissensus, Revue de philosophie politique de l'ULg*, N°4, 04/2011, p. 127

¹⁷⁸ *Op. cit.*

2.4.1.2 Fonction incantatoire des technologies Datakalab

L'homme croit au langage mais une prise de distance est parfois nécessaire pour appréhender la réalité de manière plus juste. Dans *Le Gorgias* de Platon, on trouve déjà des marqueurs de méfiance envers le logos et la rhétorique¹⁷⁹. Le lexique de Datakalab tourne autour de la « data », en témoignent le nom de la start-up et les termes qu'elle a forgés « data telling » « feel data© » « data crossing ». Ces termes sont constitués à partir de la base « data », épice de la rhétorique. Cette révolution sémantique amène à souligner les conditions de pouvoir des mots : « facial coding », « machine learning » « deep learning », « computer vision », « predictive eye tracking » ... autant de termes utilisés à la fois en interne et dans la communication de Datakalab. S'ils renvoient évidemment à de vrais domaines, ces appellations peuvent surprendre de prime abord. Innovantes, dans une langue étrangère, tendance et absconses pour le néophyte, ces formules presque magiques accordent du crédit à celui qui les invoque. Ce magicien ou rhéteur peut être un employé de Datakalab, un décideur d'entreprise, un journaliste, un consommateur. Dans tous les cas, les imaginaires cristallisés en ces termes valorisent le discours tenu. Leur utilisation devient un levier rhétorique, ces « algorithmes faiblement supervisés », « points caractéristiques qui estiment la géométrie du visage » crédibilisent Datakalab. Le langage soulève des enjeux de pouvoir. La start-up se sert de ses données émotionnelles pour établir un rapport de force avec les décideurs, qu'ils soient instituts d'études, régies, agence de communication, média, annonceurs ou consommateurs. Une fois les clés d'accès à ces données « offertes », Datakalab conserve son statut d'expert en opposition aux profanes qu'elle a formés. Les termes anglais jouent sur l'imaginaire du milieu de la start-up. Les termes comme « scalable » « from scratch » « data », « one to one » « B to B » « brainstormer » « process » « deadline » « insights » « baseline » « stand up » « proof of concept » renvoient à la fois à la réalité de l'entrepreneuriat et à l'univers de la communication. Utilisés dans la rhétorique commerciale, ils revêtent une valeur incantatoire. Par ailleurs certains, propres à la communication sont également employés par l'équipe « tech » ce qui montre que cet idéal de production massive et de performance touche tous les domaines. Datakalab psalmodie cette incantation lors de rituels : lorsqu'il présente les objectifs de la start-up, Frank Tapiro accumule ainsi les infinitifs « quantifier, structurer, analyser, stocker les émotions¹⁸⁰ ». Ces idéaux techno-centrés produisent donc des discours performatifs et

¹⁷⁹ IJsseling, Samuel, « Rhétorique et philosophie. Platon et les Sophistes, ou la tradition métaphysique et la tradition rhétorique », *Revue Philosophique de Louvain*. Quatrième série, tome 74, n°22, 1976. pages 193-210, « disponible à l'adresse : https://www.persee.fr/doc/phlou_0035-3841_1976_num_74_22_5884 »

¹⁸⁰ ANNEXE 2, Pitch Frank Tapiro – Challenge Start-me up! KPMG & Fevad

constitutifs de la nouvelle promesse. Ils doivent être expliqués, ce qui enclenche encore de nouvelles réécritures des « discours d'escorte » de la donnée émotionnelle.

En ce début de deuxième partie, je me suis penchée sur les indices d'une promesse inchangée. En effet, malgré la création de nouveaux algorithmes propriétaires et un repositionnement, Datakalab promet toujours de cerner une vérité sur les consommateurs *via* une maïeutique couplée d'un triple pouvoir heuristique, pragmatique et normatif, tout en redonnant visage humain à la donnée. Cependant, de nouveaux « discours d'escorte » sont tenus autour des « feel data© » et modulent l'ancienne promesse de « mesure émotionnelle ». Ils l'altèrent du fait même de changer ses conditions d'énonciation. Datakalab ne promet plus une « mesure » des différentes émotions identifiée par Ekman (joie, peur, colère, tristesse, surprise, mépris, dégoût) mais une mesure des « réactions ». Datakalab se spécialise alors sur les signes exprimés de l'émotion visibles sur le visage : « orientation de la tête », « clignement des yeux »... C'est bien en ce sens que l'expression « mesure des réactions » semble plus appropriée. Tous ces discours se confondent et se cristallisent dans ce nouvel engagement. L'hypothèse d'une promesse intacte suite au repositionnement n'est donc qu'en partie validée : si la vision de « redonner visage humain à la data » perdure, promesse et « discours d'escorte » évoluent ensemble. Portée par une « rationalité occidentale » qui sacralise le chiffre et la donnée, cette promesse contribue elle-même à la construction d'un système techno-centré. La nouvelle ambition de Datakalab de « mesure des réactions » tout comme le fait qu'elle soit passée d'une expertise « conseil » à « technologique » traduisent finalement une délégation du savoir humain à la machine. Jusqu'où cette nouvelle promesse est-elle alors modulée pour accroître l'opérativité sociale et symbolique des données émotionnelles en fonction des nouvelles cibles de Datakalab ?

3 DATAKALAB APPLIQUE SA NOUVELLE VISION ET S'INTÉRESSE À DE NOUVEAUX MARCHÉS

« L'adaptation », telle que l'entend Jeanneret, consiste à faire varier un message selon le « destinataire anticipé en faisant le pari que son appropriation par le public et sa circulation en seront favorisées¹⁸¹ ». Ici, je me demande si le discours de vente n'opère pas d'autant plus qu'il est adapté en fonction des besoins des clients de Datakalab. Suite à sa « bascule », la start-up se détache de sa première cible, les annonceurs. Ce faisant, elle semble moduler ses « discours d'escorte » pour répondre aux besoins spécifiques de ses nouvelles cibles, qui ne sont pas détaillées ici par souci de confidentialité¹⁸². Elle modifie son modèle économique en cherchant à passer des partenariats avec cette fameuse cible. Datakalab a été amenée à ajouter de nouveaux indicateurs dans ses rendus, et ses clients s'improvisent désormais co-constructeurs des études. Ces grands groupes ont l'habitude de travailler avec des prestataires et connaissent bien les besoins des annonceurs. Ils peuvent eux-mêmes orienter Datakalab dans ses choix de « métriques » et autres formes qu'elle donne à voir. Datakalab, cette nouvelle cible et les annonceurs semblent nourrir ensemble leurs stratégies d'innovation. Les clés du savoir, une fois données par Datakalab, permettent de nouvelles médiations qui tendent vers l'infini. Jusqu'où les discours peuvent-ils suffire pour convaincre et amorcer un tel cercle vertueux ? Enfin, quelles stratégies adopter pour parachever la transformation de la donnée émotionnelle en tant qu'indice des réactions des consommateurs face aux contenus ?

3.1.1.1 Un travail de pédagogie et de vulgarisation autour d'une vérité scientifique

Dans la mesure où Datakalab ne fournit plus de recommandations mais plutôt des scores et des « tracés », il est d'autant plus important de vulgariser, de faire de la pédagogie pour donner aux décideurs les clés d'interprétation de ce savoir. Il faut faire comprendre aux clients le sens des chiffres. Cette vulgarisation prend elle-même plusieurs formes. D'abord, elle est explicitée lors de la présentation de la technologie, des algorithmes, du « facial coding ». Aussi,

¹⁸¹ *Op. Cit.* Jeanneret, p.14

¹⁸² Ainsi, certains passages de cette dernière partie ont été masqués

¹⁸³ « L'avenir du Big data sera émotionnel », *Influencia*, 07/03/2016, « disponible à l'adresse :

<http://www.influencia.net/fr/actualites/tendance,tendances,avenir-big-data-sera-emotionnel,6132.html> »

¹⁸⁴ Jeanneret Yves, Souchier, Emmanuël. Présentation, *Communication et langages*, n°128, 2ème trimestre 2001, Dossier : Les T.I.C., objets de valeur. p. 33, « disponible à l'adresse : https://www.persee.fr/doc/AsPDF/colan_0336-1500_2001_num_128_1_3072.pdf »

¹⁸⁵ *Ibid.*

¹⁸⁶ *Ibid.*

¹⁸⁷ En mai 2019

certaines termes techniques ou qui se rapportent à des domaines particuliers comme « computer vision » doivent-ils être traduits. La vulgarisation s’effectue aussi de manière implicite, grâce aux éléments visuels. Les images, couleurs utilisées renvoient à des imaginaires de la performance et des émotions. Datakalab fait le « pari¹⁸⁸ » d’une meilleure « appropriation par le public¹⁸⁹ » de cette nouvelle promesse, liée à la manière dont les décideurs vont ensuite s’emparer de ce savoir pour faire des choix. Datakalab doit aussi effectuer un travail de pédagogie pour donner les clés d’interprétation aux décideurs. Le sujet s’avère complexe, l’univers de l’ingénieur et de ses algorithmes peut effrayer le néophyte. Datakalab fournit alors tout un travail de pédagogie lié à la nouveauté du sujet. Les scores des « métriques » s’avèrent parfois contre-intuitifs, difficiles à faire circuler. Les « métriques » sont systématiquement présentées dans le livrable¹⁹⁰. Grâce à cette dimension pédagogique, les décideurs peuvent s’approprier les résultats, les faire circuler. Elle contribue elle-aussi à asseoir la posture experte de Datakalab. En donnant les clés de lecture de son travail, la start-up confère une partie de son pouvoir et de son « ethos algorithmique » à ses clients.

Datakalab constate des freins de la part de certaines de ses cibles. Si l’émotion suscite de plus en plus d’intérêt, utiliser sa mesure, ou celle de « réactions » pour valider des stratégies requiert une certaine ouverture d’esprit de la part des décideurs. Celle-ci est difficile à corréler au retour sur investissement. La rationalité de ces démarches reste contestée, et si les décideurs ne prennent pas le temps de s’intéresser à celle de Datakalab, ils peuvent être freinés par les mêmes imaginaires qui portent la start-up. Même si les stratégies axées sur le client et sur l’humain se développent de plus en plus, Datakalab est la première start-up dans la « brain tech » en France qui analyse les réactions humaines exprimées sur le visage. Si ce sujet nouveau ne suscitait pas autant de fascination, elle pourrait ne pas connaître un tel succès.

3.1.2 Adaptation des « métriques » pour accroître leur opérativité

Depuis son repositionnement, les « discours d’accompagnement » de la donnée émotionnelle ont été réécrits, adaptés aux besoins des nouvelles cibles. « L’attention » par exemple est l’un des éléments qui a été revalorisé suite au repositionnement. Soukeina

¹⁸⁸ *Op. Cit.* Jeanneret, p.14

¹⁸⁹ *Ibid.*

¹⁹⁰ ANNEXE 20, Support de présentation PowerPoint de Datakalab, slide « Les métriques »

Hassanaly, ex « *Key Account Director* » chez Datakalab, explique¹⁹¹ : « L'objectif [...] aujourd'hui, c'est de faire la différence. Il faut créer une publicité qui ait suffisamment d'impact pour déjà attirer mon œil. Si oui, ça veut dire que j'ai rendu la personne attentive ». Datakalab cherche donc à répondre aux attentes de sa nouvelle cible. Yves Jeanneret définit « l'attente¹⁹² » comme « préconstruit social et historique susceptible d'être mobilisé dans une situation de communication particulière, explicitement ou insensiblement ». En choisissant les appellations de cette métrique, Datakalab répond à une « attente » issue d'un « préconstruit social », celle de la bataille pour l'attention. Elle combat la sursollicitation des consommateurs et rebondit directement sur les nouveaux besoins clients. Aujourd'hui, les formats publicitaires se multiplient, la lecture des sites web est polluée par des bannières et des vidéos qui se lancent automatiquement. Il y a quelques années, ces derniers étaient moins envahissants et les bloqueurs de publicités n'existaient pas. EMarketer¹⁹³ prévoit qu'en 2019 plus de 30 % des Français utiliseront un « adblock », contre 28,7 % en 2018. Ce pourcentage traduit l'agacement, « le ras-le-bol » croissant à l'encontre de la publicité. Une large part de la stratégie de Datakalab réside dans ses choix de « métriques ». Celles-ci doivent rendre compte de cette nouvelle réalité, elles doivent avoir un sens à la fois d'un point de vue de la réalité scientifique, mais aussi du nouveau contexte socio-économique. Datakalab a donc créé trois « métriques¹⁹⁴ » : « l'attention », « l'émergence » et « l'adhésion ».

« L'attention » constitue la première et désormais principale « métrique » de Datakalab. Proche du « temps de cerveau humain disponible » de Patrick Le Lay, elle est définie par Datakalab comme « le pourcentage de personnes ayant été attentives durant tout le visionnage d'un contenu¹⁹⁵ ». Datakalab explique que l'individu est confronté tous les jours à environ 15 000 stimuli commerciaux¹⁹⁶. Frank Tapiro explique¹⁹⁷ : « Le nerf de la guerre aujourd'hui c'est l'attention. S'il n'y a pas d'attention, tout le reste ne sert à rien. On a tendance à oublier que c'est l'attention qui fait tout, c'est le point de départ. S'il y a de l'attention, on peut « mesurer l'émotion ». Sans cette « métrique », les autres n'auraient pas lieu d'être. « L'économie de l'attention » s'intéresse à ce concept comme une ressource rare. Dans

¹⁹¹ ANNEXE 5, Entretien avec Soukeina Hassanaly, Ex Key Account Director chez Datakalab, le 05/04/2019, Paris 17^{ème}

¹⁹² Jeanneret, Yves, *Critique de la trivialité*. « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, p. 72

¹⁹³ Ropars, Fabian, BDM/Media, 07/09/2018, « 30 % des internautes Français utiliseront un ad blocker en 2019 »,

¹⁹⁴ ANNEXE 20, Support de présentation PowerPoint de Datakalab, slide « Les métriques »

¹⁹⁵ *Ibid.*

¹⁹⁶ ANNEXE 20, Support de présentation PowerPoint de Datakalab, slide « La réalité d'aujourd'hui »

¹⁹⁷ ANNEXE 4, Entretien avec Frank Tapiro, Co-fondateur et Chief Emotion Officer, 26/06/2019, Paris 17^{ème}

l'introduction de *L'Économie de l'attention* « Nouvel horizon du capitalisme¹⁹⁸ ? » Yves Citton explique que depuis les années 1990, nous basculons d'une économie « traditionnelle » à une « économie de l'attention » et que « le principe de rareté paraît s'être radicalement déplacé, depuis le pôle de la production vers le pôle de la réception¹⁹⁹ ». Avec cet intérêt pour « l'attention » Datakalab s'inscrit donc dans le nouveau contexte économique et politique d'incitation à la performance.

Datakalab nomme « émergence²⁰⁰ » sa seconde « métrique », soit la présence d'émotions en réaction à un contenu. Ce score indique la « non neutralité ». Datakalab explique que si l'on suscite des émotions, on garantit une meilleure mémorisation, revendiquant encore son ancien « ethos neuroscientifique ». À ses débuts, Datakalab parlait « d'engagement émotionnel », ensuite, « d'émergence émotionnelle » puis simplement « d'émergence ». Le passage « d'engagement émotionnel » à « émergence émotionnelle » a été établi pour souligner la nécessité pour une création publicitaire d'émerger, soit de ressortir dans la masse formée par d'autres publicités. Ensuite, la direction a estimé que « émergence émotionnelle » s'apparentait à un pléonasme et a simplifié en ne gardant que « émergence ». Certains clients choisissent eux-mêmes de renommer cette « métrique » en « présence d'émotions ». L'expression peut être considérée comme moins vendeuse, mais elle s'avère directement intelligible. Celle-ci est complexe à traduire²⁰¹, et la start-up cherche un terme qui puisse être utilisé en anglais comme en Français. Frank Tapiro propose alors de renommer « l'émergence » en « impact ». Il explique²⁰² : « l'émergence est encore un mot qui sonne "conseil". L'impact, tout le monde comprend, parce que dans tous les domaines : la communication, le cinéma, s'il n'y a pas d'impact, il n'y a pas d'émotion ». Datakalab cherche donc à créer des « métriques » qui fonctionneraient « dans tous les domaines », cherchant encore à maximiser sa production. Selon le publicitaire, la dialectique pour comprendre les « réactions » face à un contenu consiste d'abord à se demander s'il a capté « l'attention » de son public, généré des « ressentis », et si ces « ressentis » étaient positifs ou négatifs.²⁰³ En psychologie les « ressentis » désignent

¹⁹⁸ *L'Économie de l'attention*, « Nouvel horizon du capitalisme ? » sous la direction de Yves Citton, p. 07, « disponible à l'adresse : <http://www.yvescitton.net/wp-content/uploads/2017/09/CITTON-EconomieAttention-Intro-2014.pdf> »

¹⁹⁹ *Ibid.*

²⁰⁰ Dans les présentations, « l'émergence » est définie comme « pourcentage de personnes ayant ressenti au moins une émotion »

²⁰¹ En anglais, la traduction s'en est tenue à « emotional engagement », soit « engagement émotionnel » la première version de la « métrique »

²⁰² ANNEXE 4, Entretien avec Frank Tapiro, Co-fondateur et Chief Emotion Officer, 26/06/2019, Paris 17^{ème}

²⁰³ *Ibid.*

justement les signes exprimés de l'émotion. C'est en ce sens que Datakalab pourrait s'en servir dans sa communication. Frank Tapiro propose même d'ajouter le « rejet » à la liste des « métriques » officielles²⁰⁴. Cette dernière a pu être utilisée dans certaines études. Pour la construire, Datakalab effectue une moyenne de l'ensemble des émotions négatives. Plus le pourcentage est élevé, plus le contenu fait réagir les personnes de manière négative. Sans attention ou émotion, le message ne peut être retenu, et le contenu ne doit pas laisser indifférent. Frank Tapiro ajoute²⁰⁵ : « Je préfère une campagne détestée, conspuée, censurée à une campagne sans saveur, sans odeur. Au moins, on a brisé la barrière de l'indifférence. Aujourd'hui, on ne peut plus être en-deçà de la barrière de l'indifférence, c'est le nerf de la guerre dans tout ». L'expression « barrière de l'indifférence » sonne et fait prendre conscience du manque d'intérêt des cibles pour les messages qui leur sont adressés. Selon lui, tant que les contenus ne se démarquent pas, ils restent dans une forme de médiocrité, s'avèrent fades, banals, dépourvus d'identité.

Enfin, « l'adhésion », troisième métrique, est définie comme « le pourcentage de personnes ayant ressenti au moins une émotion positive ». Cette dernière permet de s'assurer que cette émotion soit au service de la marque. Elle combine la joie et la surprise pour indiquer si les contenus ont fait réagir positivement les panélistes. L'intensité émotionnelle ne fait pas tout, il faut encore que celle-ci soit en accord avec les objectifs de la marque pour laisser un souvenir favorable.

Datakalab va jusqu'à personnaliser ses métriques en fonction du secteur de ses clients. Elle a travaillé avec Do The Right films, une société qui produit des études spectateurs et effectue du conseil auprès des distributeurs et producteurs de films grâce *via* des tests. Dans le cas du cinéma, les métriques classiques de Datakalab, « attention », « émergence », « adhésion », ne correspondaient pas totalement aux enjeux de l'industrie. Ce qui lui importait, ce n'était pas tant de savoir si le film captait « l'attention » du spectateur par exemple, mais plutôt s'il laissait des « ressentis positifs » ou « négatifs », voire s'il laissait indifférent. Pour ces études, Datakalab a dû inventer de nouvelles formes, en l'occurrence des émojis qui traduisent ces différents ressentis. Tout comme « les icônes des écrans ne sont pas choisies pour

²⁰⁴ ANNEXE 20, Support de présentation PowerPoint de Datakalab, slide « Les métriques »

²⁰⁵ *Ibid.*

leur cohérence sémiotique mais pour leur familiarité dans la vie quotidienne²⁰⁶ » selon Jeanneret, ces émoticônes qui font respectivement un sourire et une moue triste ont été choisis pour leur pouvoir symbolique. En ce sens, Datakalab cherche à s'inscrire dans les formes connues, tout en inventant ses propres « métriques ». Elle doit aussi marquer « une rupture avec les attentes établies du genre », soit les « attentes » relatives aux instituts d'études dont elle doit se démarquer. « Ressentis positifs » soit la moyenne des émotions dites positives et « ressentis négatifs » se sont vues donner ces « corps ». Les émoticônes relèvent de l'infra-ordinaire au sens donné par Gorges Perec. Yves Jeanneret et Emmanüel Souchier développent la pensée de l'auteur : l'objet « disparaît à force d'évidence » et « on ne (lui) prête que peu d'attention²⁰⁷ ». À travers cette forme choisie, apparemment banale, se cristallisent en fait des enjeux de pouvoir, ici celui de se faire reconnaître par ses clients. Ces nouvelles métriques de Datakalab « charrient une riche mémoire sociale de formes antérieures²⁰⁸ ». Ils ne sont utilisés que parce qu'ils sont connus, porteurs de sens. Une promesse et une implication « ne créent du sens et du lien que si les formes sont reconnues des sujets sociaux - reconnues, c'est-à-dire reliées à un déjà là²⁰⁹ ». À travers l'emploi de cette forme, Datakalab a donc voulu adapter sa promesse, personnaliser l'étude pour fidéliser Do The Right Films. L'opérativité de ces données émotionnelles s'en est trouvée accrue, et a pu circuler au sein de la direction de la société, qui s'est accaparé le savoir de Datakalab pour le transmettre à ses propres clients.

²⁰⁶ Jeanneret, Yves, p.35, *Critique de la Trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014

²⁰⁷ Jeanneret, Yves et Souchier, Emmanüel, « L'énonciation dans les écrits d'écrans », *Communication et langages*, l'emprunte de la technique dans le livre, n°145, 3^{ème} trimestre, pages 3-15, « disponible à l'adresse : https://www.persee.fr/doc/colan_0336-1500_2005_num_145_1_335 »

²⁰⁸ Jeanneret, Yves, *Critique de la trivialité*. « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014 : « les modèles, si nouveaux soient-ils, charrient une riche mémoire sociale de formes antérieures ; les formes de la communication sont déterminées de manière peu visible mais bien réelle par les outils qu'elles mobilisent, qui ont leur propre histoire et leur propre logique », p. 40

²⁰⁹ *Ibid.* p.73

3.2 Datakalab, les instituts d'études et les annonceurs nourrissent ensemble leurs stratégies d'innovation

La technologie de Datakalab constitue un outil d'aide à la décision. Pour l'étudier de manière critique, on peut se référer aux travaux d'Emmanuel Souchier qui affirme au sujet de « l'écriture du web²¹⁰ » : « les instruments dont elle se saisit sont aussi ceux par lesquels elle est elle-même saisie, [...] les dispositifs matériels par lesquels elle se fabrique sont aussi des dispositifs qui la fabriquent, auxquels elle se plie ou avec lesquels elle se débat²¹¹ ». Dans la mesure où les décideurs se servent de cette donnée comme outil de gouvernement, il convient d'étudier la manière dont ces acteurs se saisissent de cette technologie. Datakalab doit tenir un discours à deux destinataires différents aux annonceurs et aux instituts d'études. La majorité des arguments fonctionnent pour l'un comme pour l'autre, et des médiations s'opèrent lorsque les instituts d'études s'approprient le discours de Datakalab pour convaincre les annonceurs de passer des partenariats avec eux. En travaillant avec elle, les instituts d'études se légitiment à la fois en interne en montrant aux employés une volonté de croissance, mais aussi en externe, face aux annonceurs et à la presse. Ils revendiquent les connaissances de type neuroscientifique encore vulgarisées par Datakalab. Le communiqué de presse de Médiamétrie²¹² : « L'émotion des spectateurs au cœur de la mesure cinéma de Médiamétrie » du 11 juillet 2019 illustre ce phénomène. Médiamétrie y explique : « Pour 76 % des spectateurs, le cinéma est avant tout un vecteur d'émotions ». Celles-ci sont utilisées pour établir un lien entre Datakalab et le cinéma. Le leader des études médias s'adresse à la fois aux annonceurs avec lesquels elle cherche à passer des partenariats, et à la presse qui construit son image. « Pour la première fois en France, Médiamétrie va plus loin dans son analyse du ressenti des spectateurs en utilisant une technologie développée par la société Datakalab ». Les expressions « pour la première fois en France », « va plus loin » « technologie » renvoient à la recherche de l'innovation et de la performance. Dans son argumentaire, Marine Boulanger, « Directrice du pôle Cinéma et Entertainment » de Médiamétrie, insiste sur la valeur ajoutée de ce nouveau type de données : « Ces données émotionnelles viennent appuyer et détailler les données des spectateurs que nous recueillons dans nos études, pour apporter une réponse encore plus concrète sur l'attractivité et

²¹⁰ *Op. Cit.* « La "cartographie" au risque de l'histoire de l'écriture », CNRS Editions, 2013, p. 231, « disponible à : <https://books.openedition.org/editions-cnrs/21786?lang=en> »

²¹¹ Chatelain et Jacob, 2004

²¹² « L'émotion des spectateurs au cœur de la mesure cinéma de Médiamétrie », Communiqué de presse Médiamétrie du 11/07/2019, « disponible à l'adresse : <https://www.mediametrie.fr/fr/lemotion-des-spectateurs-au-coeur-de-la-mesure-cinema-de-mediametrie> »

le potentiel d'un film, ou l'efficacité d'une bande-annonce²¹³ ». L'objectif de Datakalab est atteint. Le leader des études média affirme lui-même le cœur de sa nouvelle rhétorique, à savoir, le fait que ces données émotionnelles complètent les « données des spectateurs » que Médiamétrie collecte dans ses études. Ici, les « feel data© » sont introduites non pas par Datakalab, mais par un autre protagoniste. À la manière de l'acolyte du soldat fanfaron dans la littérature, la start-up se laisse présenter par un autre acteur de crédit. Médiamétrie narre ses exploits alors qu'elle est étrangère à ces domaines technologiques : « Le deep learning, une branche de l'intelligence artificielle, permet à Datakalab de mesurer l'attention et l'émotion des spectateurs grâce à des caméras placées de part et d'autre de l'écran de cinéma²¹⁴ ». La limite de cette comparaison se trouve dans le fait que le travail de Datakalab ne revêt en aucun cas le fantasme ou le caractère affabulateur du *miles gloriosus*²¹⁵. Si ensuite les annonceurs acceptent de passer des contrats avec les instituts d'études et Datakalab, ils peuvent à leur tour se montrer en train d'innover par ce type de communiqué. Ce qui se joue ici, c'est bien une appropriation du pouvoir, une course à l'accapuration pour un règne de la performance. L'appel – même indirect – à des sociétés telles que Datakalab témoigne d'une certaine ouverture d'esprit, d'une volonté de travailler avec des technologies modernes et d'une quête du chiffre. Cela contribue à définir l'image de la marque, aux yeux des collaborateurs et employés, consommateurs, et médias. Ces scores leur permettent de justifier des choix stratégiques au sein même des sphères directionnelles. Les entreprises s'appuient alors sur le savoir fourni par Datakalab, restitué par les instituts d'études pour construire à leur tour de nouveaux « discours d'escorte » de la donnée émotionnelle. Ces paroles se retrouvent bien « trivialisées²¹⁶ » au sens donné par Jeanneret, dans la mesure où elles revêtent un « caractère formateur et créatif [...] à travers différents espaces sociaux²¹⁷ ». L'utilisation de ce savoir inouï est elle-même à nouveau soumise à d'autres autorités au sein de l'entreprise. La « raison algorithmique²¹⁸ » confère une posture rationnelle, donne du pouvoir et justifie des décisions de différents types : organisationnels, créatifs, stratégiques. On assiste donc ici à un triple processus légitimant la donnée émotionnelle, Datakalab, les instituts d'études, et les annonceurs s'adressant les uns aux autres, parfois de manière directe, parfois sous l'égide de la presse.

²¹³ *Ibid.*

²¹⁴ *Ibid.*

²¹⁵ Latin pour « soldat fanfaron »

²¹⁶ Jeanneret, Yves, *Critique de la trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, p. 15

²¹⁷ *Ibid.*

²¹⁸ *Op. Cit.* Sadin

3.2.1 Les appropriations de la « feel data © » témoignent de son statut « d'indice » d'appréhension des « réactions »

Au sein de l'entreprise, le savoir relatif à la donnée émotionnelle circule. Il est à nouveau traduit par différents acteurs. Les commerciaux l'expliquent à leur direction, et la direction s'en saisit à son tour pour énoncer sa stratégie. La circulation de cette « raison numérique²¹⁹ » peut être appréhendée comme la manifestation du nouveau système de croyance, non plus seulement en la donnée, mais en l'une de ses formes, celle dite émotionnelle. Le commercial qui a lui-même été convaincu des discours tenus par Datakalab doit à son tour revêtir l'habit du « messie²²⁰ » qui « prêche » la donnée émotionnelle auprès de sa propre entreprise et doit la vulgariser. Il passe donc d'une posture d'élève, simple initié, à une posture de maître. S'il s'appuie sur le savoir revendiqué par Datakalab, il doit aussi adapter ce discours aux enjeux de sa propre entreprise. Ce « messie » promet alors de meilleures performances, un meilleur retour sur investissement. Il est arrivé qu'Exterior Media contacte Datakalab pour demander des arguments qui convainquent un client de l'efficacité des données émotionnelles. Datakalab a dû alors façonner un argumentaire à partir de cas existants. Ici, le savoir a été « trivialisé » à plusieurs échelles. D'abord la donnée brute a été construite et mise en forme pour intéresser le client par les ingénieurs, interprétée par l'un des Project Managers de Datakalab, puis livrée à l'afficheur. L'afficheur a dû lui-même convaincre son propre client à partir de ce savoir. Le client a dû à son tour en informer sa direction pour ensuite revenir vers Exterior Media, à nouveau revenu vers Datakalab. Ces échanges forment une boucle autour du savoir dont les « acteurs sociaux s'emparent [...] pour produire du pouvoir et de la valeur²²¹ ». Plus ce dernier circule, plus il prend du poids.

3.2.1.1 Une approche descriptive et non analytique de la presse liée à une méconnaissance technique

Une légitimation de ce nouveau type de données est opérée par le « discours d'escorte » des professionnels et de la presse. La presse légitime la donnée émotionnelle en lui produisant elle-même ses propres « discours d'accompagnement ». Pourtant les médias non spécialisés s'en tiennent pour l'instant à une simple description. Les titres des articles prennent la forme

²¹⁹ Sadin, Éric, *La vie algorithmique*. « Critique de la raison numérique », Paris : Éditions l'Échappée. DL 2015, (quatrième de couverture)

²²⁰ *Op. Cit.*, adjectif utilisé par Alexandra Ducrot au sujet des acteurs qui se saisissent des technologies big data, p. 73

²²¹ Jeanneret, Yves, *Critique de la trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, p. 20

interrogative, ainsi Méta Media²²² du 15 mars 2016 : « L'avenir de la data sera-t-il émotionnel ? ». L'absence de nette prise de position est liée au fait que les journalistes connaissent mal le sujet dont la nouveauté suscite l'engouement. Cet enthousiasme est renforcé par l'objet traité, celui des nouvelles technologies, du « facial coding », des univers qui inspirent autant une certaine fascination qu'ils suscitent la crainte. Dans tous les cas, ils engendrent des « réactions ». Autant de critères qui font que les discours autour de ces tendances se multiplient. Les différents acteurs tels que les entreprises et les médias s'intéressent à la fois de plus en plus à la donnée et aux émotions. Appliquer la compréhension des émotions humaines au monde numérique captive, et plus les discours sont tenus, plus ils sont « trivialisés ».

3.3 Une « rationalité occidentale » du nombre, du calcul et de la donnée porte les données émotionnelles

Tout ce qui concerne la science et la technologie est aujourd'hui couplé d'imaginaires rarement questionnés ou remis en perspective. Ces discours d'escorte font autorité puisqu'il s'agit de mythes opérants qui éveillent notre « rationalité occidentale²²³ ». Beaucoup de sociétés de neuromarketing comme Affectiva jouent sur l'effet de fascination autour de la donnée et des algorithmes. On peut lire dans la rubrique « how it works » : *using computer vision speech analytics, deep learning and a lot of data, we analyse human states in context*²²⁴. Ce qu'on vend ici, c'est bien la technologie, une « prouesse technique » mathématico-scientifique. D'après Ioanna Kuçuradi dans « "Rationalité" et "rationalités" dans le cadre du débat modernisme – post-modernisme²²⁵ », notre « rationalité occidentale » est marquée par le fait qu'on souhaite expliquer le monde par « l'observation et l'expérimentation », ou par la « science positive ». Notre « rationalité occidentale²²⁶ » fait valoir la démonstration rationnelle, le scientisme qui affirme que : « l'homme doit se borner à ce qu'il peut savoir de manière certaine, grâce à la science ». Sur son site, Kairos, l'un des concurrents de Datakalab en reconnaissance faciale fondé en 2012 aux États-Unis, évoque un « contrôle ultime de la donnée²²⁷ » et donne à voir du code directement sur sa page d'accueil.

²²² Balla, Émilie, France Télévisions, Direction de la Prospective, 15/03/2016, « disponible à l'adresse : <https://www.meta-media.fr/2016/03/15/lavenir-de-la-data-sera-t-il-emotionnel.html> »

²²³ *Op. Cit.*

²²⁴ Je traduis « à partir des analyses de discours issus de la vision par ordinateur, du deep learning et de nombreuses données, nous analysons les états humains en regard d'un contexte ».

²²⁵ *Op. Cit.*

²²⁶ *Op. Cit.*

²²⁷ Kairos.com, « je traduis de : *Integrate Face Recognition via our cloud API, or host Kairos on your own servers for ultimate control of data* ».



Visuels présents sur la page d'accueil de Kairos, « disponible à l'adresse : <https://www.kairos.com/> »

Pour le néophyte, cette représentation impressionne et elle est liée à l'imaginaire de la base de données. On adhère alors à un système de représentations dont on ne connaît rien uniquement parce qu'il a été créé par des experts du domaine, des ingénieurs. On leur cède notre confiance, puisque le contexte politique et social nous apprend à adhérer aussi bien aux sciences exactes qu'aux affirmations des professionnels. Plus ce type de discours se répand, plus le pouvoir même des « traces », des données s'accroît. La donnée a donc été amenée par l'avènement du numérique au cœur des discours contemporains.

3.3.1 « Gouvernamentalité algorithmique » : la « data » fait loi

Aujourd'hui, nous vivons donc dans cette « société du calcul » que dépeint Dominique Cardon. Dans *À quoi rêvent les algorithmes*²²⁸, le sociologue précise les domaines d'intrusion de la donnée dans notre quotidien : « Nos vies à l'heure des "big data" utilisent l'expression de "société du calcul" dans le sens où tout est chiffre : "Classement, personnalisation publicitaire, recommandation de produits, mesures corporelles, orientation et déplacements" ». Toute situation est systématiquement chiffrée en vue d'une meilleure maîtrise du réel. C'est dans cette même optique qu'on peut penser que les acteurs de Datakalab exercent une « raison algorithmique ». On assiste à la fois à la multiplication des « traces » et des discours qu'on leur porte. Éric Sadin mentionne une « puissance toujours plus totalisante détenue par les systèmes

²²⁸ Cardon, Dominique, *A quoi rêvent les algorithmes ? : Nos vies à l'heure des big data*, Paris, Seuil, La République des idées, 2015, Quatrième de couverture

computationnels²²⁹ ». Les émotions sont elles-mêmes mises en vente considérées *as a service*²³⁰ par certaines sociétés de neuromarketing comme Affectiva. En quantifiant l'émotion grâce à la machine, c'est peut-être la vie dans son ensemble qu'on met à vendre. Dominique Cardon redoute expressément cet aspect de la numérisation de la société : « La constitution de grandes bases de données confère une place de plus en plus centrale aux algorithmes qui gouvernent les comportements de chacun ». Cette constitution de base de données pour appréhender le monde est elle-même dénoncée par Antoinette Rouvroy, *via* le concept de « gouvernementalité algorithmique²³¹ ». Par le terme de « critique de la gouvernementalité algorithmique » Antoinette Rouvroy réalise que les acteurs de la donnée souhaitent « que se produise cette substitution d'une réalité numérique au monde tel qu'il est²³² ».

3.3.2 L'innovation technologique au cœur de la rhétorique commerciale

Pour séduire, la nouvelle promesse de Datakalab doit donc s'inscrire dans cette quête perpétuelle de croissance issue d'idéaux américains. Son nouveau discours y répond parfaitement, dans la mesure où il est « techno-centré », promettant ainsi une réponse quasi totale aux besoins des nouvelles cibles de Datakalab. Ces discours peuvent opérer puisqu'on assiste à un idéal de performance, d'innovation axé sur le chiffre. Antoine Ribou, fondateur et président de Danone, définit l'innovation comme « une alliance entre recherche, marketing, instinct, imagination, produit et courage industriel ». Datakalab mêle tous ces éléments, en particulier celui de « courage industriel » dans la mesure où faire le pari de se repositionner uniquement sur l'expertise « technologique » et délaisser subitement le « conseil » peut sembler rapide, voire même brutal (la bascule s'est faite en quelques mois seulement) et risqué (sans expertise « conseil », les données émotionnelles convaincront-elles à elles-seules ?). Il en va de même pour le choix de s'autofinancer à ses débuts, soit les trois premières années, sans levée de fonds. Dans le cas de Datakalab, « innovation » semble même un euphémisme, puisqu'à son lancement sa promesse de « mesure des émotions » était souvent qualifiée de « révolution ». Selon l'encyclopédie Universalis²³³ : « Étymologiquement, la « révolution » s'entend comme

²²⁹ Sadin, Éric, *La vie algorithmique* « Critique de la raison numérique », Quatrième de couverture

²³⁰ « Je traduis : en tant que service » : *Emotion as a Service enables you to analyze human expressions of emotion*, Affectiva, « disponible à l'adresse : <https://www.affectiva.com/product/emotion-as-a-service/> »

²³¹ Séminaire « Big data et anticipation : vers une gouvernementalité algorithmique ? », organisé le 16/04/2015 à l'Espace éthique/IDF (disponible sur YouTube, vidéo publiée le 22 mai 2015 par Espace Éthique, « disponible à l'adresse : www.youtube.com/watch?v=RUEkEsB6FJA »

²³² *Ibid.*

²³³ « Disponible à l'adresse : <https://www.universalis.fr/encyclopedie/revolution/> »

"retour sur soi", comme réitération de ce qui a été et comme prévision de ce qui sera ». Mais historiquement, elle prend le sens de rupture. On parle souvent de « révolution » lorsqu'une nouvelle technologie est inventée. Le discours de « révolution » autour de l'innovation technologique inscrit les start-up comme Datakalab dans la recherche permanente de la modernité, elle a d'ailleurs elle-même opéré sa propre « révolution » en créant sa technologie suite à un « retour sur soi ». Le mythe du progrès apparu au XVII^{ème} siècle est ici réactivé. Le progrès, défini par Frédéric Rouvillois dans *L'Invention du progrès* comme « mouvement ascendant et nécessaire²³⁴ » de la société pousse les hommes à se dépasser. Une fois que de nouveaux formats médiatiques se sont développés, les cibles en prennent l'habitude. L'innovation se retrouve aussi prise au piège de ses propres idées. Dans l'avenir, les cibles se laisseront sans doute également des publicités personnalisées. De nouveaux formats médiatiques resteront sans doute toujours à inventer laissant à Datakalab de nouvelles possibilités d'innovation. Ces idées rejoignent la nécessité d'innover théorisée par Hartmut Rosa, qui va jusqu'à évoquer la « pathologie de la modernité » : « Comme la société moderne et capitaliste ne reproduit sa structure qu'à travers l'innovation et l'accélération – ce que je qualifie de "dynamisation du temps" – on a besoin de croître et d'innover pour préserver les structures économiques, sociales, politiques ». Pour séduire, la nouvelle promesse de Datakalab doit donc s'inscrire dans cette quête perpétuelle de croissance. Son nouveau discours y répond parfaitement, et plus encore, il même les besoins de ses cibles.

3.3.3 Double instance énonciative ingénieur/communicant : vendre une recherche

Au sein même de la direction de Datakalab on assiste à ces discours hybrides. Parmi les quatre membres fondateurs, les Centraliens Xavier et Lucas Fischer disposent d'une expertise technologique et en intelligence artificielle. De l'autre côté deux publicitaires, Frank Tapiro et Anne-Marie Gaultier (jusqu'à début 2019), disposent d'une expertise « conseil/marketing ». On constate alors une « polyphonie énonciative », définie comme suit par Bakhtine : « L'hybridité se produit lorsque deux ou plusieurs discours se disputent l'autorité de l'énonciation²³⁵ ». Les

²³⁴ Rouvillois Frédéric, *L'Invention du progrès*, CNRS Editions, « "L'idée de Progrès" apparue lors de la révolution scientifique mécaniste du XVII^e siècle et conçue comme « mouvement ascendant et nécessaire » de la société, marque la philosophie de l'histoire de la modernité », « disponible à l'adresse : www.lefigaro.fr/lefigaromagazine/2011/02/19/01006-20110219ARTFIG00581-lectures-le-mythe-du-progres.php »

²³⁵ Le comptoir des presses d'universités, Résumé, « disponible à l'adresse www.lcdpu.fr/livre/?GCOI=27000100318130 » sur *Le texte hybride*, Presses Sorbonne Nouvelle

logiques des communicants et des ingénieurs ne sont pas toujours similaires²³⁶. Datakalab se construit *via* un échange entre ces deux pôles. Sans les ingénieurs, le pôle commercial/stratégique ne pourrait pas construire son discours, tout comme les ingénieurs seraient moins guidés dans leur travail sans les communicants. On assiste ainsi à une hybridation des compétences, les uns ayant identifié les besoins et contraintes des autres. En ce sens, il arrive aux commerciaux de parler de code, « d'API » et « de points caractéristiques du visage » et de beaucoup d'autres termes issus de « l'ethos algorithmique ». Les ingénieurs prennent quant à eux part aux discussions avec la direction stratégique et expliquent leurs avancées, ce qui permet aussi de savoir sur quoi communiquer. À la fois polyvalente et ambivalente, cette double instance énonciative s'auto-légitime. La polyphonie garantit l'expertise technico-scientifique sous-jacente. Cette expertise est enrobée par des discours commerciaux et vulgarisée. Corrélés et indissociables, ces discours caractérisent la posture d'énonciation des entreprises expertes. On amoindrit presque l'intention commerciale par un discours sur l'innovation, la recherche, la science. C'est cette même instance qui régule les besoins de la cible, constituant elle-même un outil d'aide à la décision pour Datakalab. On retrouve ce mouvement à une plus large échelle au sein de l'équipe. Ces postures énonciatives correspondent aussi aux différents corps de métiers présents au sein de la start-up. L'ingénieur dispose d'un « ethos algorithmique » lié aux enjeux de son métier. Il utilise couramment des termes techniques tels que la « descente de gradient ». Le commercial produit quant à lui un discours qui doit parfois s'éloigner de la réalité technique pour convaincre. Lorsque les commerciaux présentent l'outil de « facial coding » de Datakalab, ils le rattachent à de l'acquis, à du « connu » afin de mieux le faire comprendre. Ainsi évoquent-ils à son propos la série télévisée américaine *Lie to Me*, et le travail de son inspirateur et conseiller Paul Ekman spécialiste des « micro-expressions ». Ce faisant, ils créent un nouvel imaginaire, celui de Datakalab, à partir d'une base elle-même imaginaire. Cette allégorie, imaginaire, fictive, passe ensuite du stade de représentation mentale à celui de « réalité ». Interrogé sur les discours commerciaux tenus autour de son travail scientifique, Arnaud Dapogny, Machine Learning Scientist chez Datakalab et Chercheur à Sorbonne Université répond²³⁷ : « Nous devons essayer de créer des systèmes performants. Du côté marketing/communication, vous devez rendre ces données intéressantes pour le client ». Cette imbrication de discours se produit à chaque fois

²³⁶ Par exemple Frank Tapiro souhaite parfois communiquer sur plusieurs sujets sur les réseaux sociaux, prétextant que « abondance de biens ne nuit pas », tandis que Xavier Fischer préfère prendre la parole sur un sujet technique par semaine

²³⁷ ANNEXE 11, Entretien avec Arnaud Dapogny, Machine Learning Scientist chez Datakalab, 27/06/19, Paris 17^{ème}

qu'il est nécessaire de vulgariser les connaissances techniques d'un domaine, de les « rendre intéressantes » pour les vendre. Cette tension entre véracité scientifique et nécessité de vulgarisation est soulignée par Yves Jeanneret : « L'écrit de médiation est marqué par une tension permanente entre cet idéal d'optimisation communicationnelle des textes et la réalité d'une écriture savante marquée par une auctorialité scientifique²³⁸ ». Dans le cas de « l'ethos algorithmique », cette logique s'applique, les commerciaux faisant appel à des ressorts rhétoriques, des superlatifs. Les chercheurs de Datakalab sont par exemple présentés comme « meilleurs chercheurs au monde » sur les salons. Leurs algorithmes constituent l'une des fiertés de la start-up. Cette nécessité de se vendre pour vendre se retrouve même dans la recherche scientifique. Arnaud Dapogny m'explique sa conception des termes issus de ces modèles : « "L'affective computing" se rapproche d'un terme "bullshit" politique pour vendre. C'est un nouveau terme comme celui de "brain tech". À mon sens, tous ces termes veulent dire la même chose. Je considère que la recherche devrait être débarrassée de ces problématiques de traitement de l'information²³⁹ ». Ces discours peuvent autant porter la promesse que la réduire à une dimension uniquement liée à la production de masse au détriment d'un gage sincère et humain.

En début de troisième partie, je me demandais si les discours modulés en fonction des besoins des nouveaux clients de Datakalab permettaient d'accroître l'opérativité des données émotionnelles. Datakalab module et adapte son discours en fonction des différents secteurs et enjeux pour convaincre. Lorsqu'elle travaille avec des instituts d'études par exemple, on assiste à une triple légitimation, dans un triangle qui implique également l'annonceur. Datakalab partage ses clés d'interprétation de la donnée, accordant ainsi aux clients une part du pouvoir issu de l'algorithme. Les instituts d'études s'approprient alors cette donnée, et créent eux-mêmes de nouveau « discours d'escorte ». On cherche en fait à atteindre un idéal du nombre, porté par la « rationalité occidentale » pour le calcul. Ainsi, ces discours d'escorte sont donc bien modulés en vue d'une opérativité, ce qui valide ma dernière hypothèse.

²³⁸ Jeanneret, Yves, *Critique de la trivialité*, « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014, p.100

²³⁹ ANNEXE 11, Entretien avec Arnaud Dapogny, Machine Learning Scientist chez Datakalab, 27/06/19, Paris 17^{ème}

3.4 RECOMMANDATIONS

3.4.1 Datakalab doit continuer d'appliquer sa nouvelle expertise

Datakalab dispose d'une nouvelle expertise, qu'elle doit désormais appliquer. Elle initie sa nouvelle cible pour la convaincre d'avoir recours à son nouvel outil. Comme l'explique Xavier Fischer, il faut signer des abonnements avec elle. Les contrats signés au début par Datakalab comme avec la SNCF engageaient des sommes bien plus importantes²⁴⁰. En revanche, ces contrats de type « one shot », soit en une seule fois, étaient si onéreux pour les annonceurs que Datakalab n'avait aucune garantie qu'ils soient renouvelés. Les annonceurs demandaient aussi de nombreuses adaptations dans les présentations en fonction de leurs enjeux. Elle est désormais spécialisée dans sa technologie propriétaire de « facial coding ». En ce sens, elle a récemment finalisé sa première levée de fonds et peut envisager de se développer plus rapidement.

L'hybridation est ici un concept clé : la donnée déclarative passe du statut d'ennemie à celui d'alliée, avec laquelle se construire. La donnée émotionnelle se révèle un levier pour la nouvelle cible qui lui permet de se démarquer par une « brique » originale et innovante. Cela permet de spécifier la donnée livrée, à l'image des résultats déjà envoyés à Exterior Media, segmentés en fonction du sexe des panélistes²⁴¹. Croiser les données déclaratives et émotionnelles ne suffit pas. Une stratégie idéale prendrait en compte d'autres indicateurs clés de l'entreprise comme son chiffre d'affaires, ses investissements publicitaires pour inscrire les études dans une réalité business. La start-up ne doit pas abandonner pour autant la cible des annonceurs dans la mesure ils lui font mieux comprendre les besoins du marché, et parfois signer des contrats plus importants. Ainsi Datakalab pourrait-elle continuer de se rapprocher d'annonceurs dont elle a déjà testé le contenu.

3.4.1.1 Segmentation hyper personnalisée permise par les agences et régies

Datakalab pourra se concentrer sur d'autres sociétés qui collectent des données des consommateurs, comme les agences médias et les régies, et réfléchir à la façon dont combiner

²⁴⁰ ANNEXE 4, Entretien avec Frank Tapiro, Co-fondateur et Chief Emotion Officer, 26/06/2019, Paris 17^{ème} : « Datakalab était à l'époque principalement portée sur le conseil. On vendait ce produit cher, mais la marge était moins forte qu'aujourd'hui avec le facial coding »

²⁴¹ Datakalab donnait par exemple un score d'attention global, puis le score hommes et enfin celui des femmes

leurs expertises. Il semble pertinent de se rapprocher d'agences créatives²⁴², mais aussi d'agences média²⁴³. Avec les régies, elles peuvent présenter à l'annonceur ses gains, à savoir une réduction des investissements médias, comme dans le cas Bompard en première partie. Le partenariat avec Exterior Media en 2019 a été un succès. Si celui-ci le permet, Datakalab pourrait se rapprocher d'autres régies²⁴⁴. Les instituts d'études et afficheurs peuvent recroiser les résultats de « mesure émotionnelle » avec le média sur lequel est diffusé le contenu afin d'avoir un contenu plus riche à livrer à l'annonceur. Par exemple, dans l'article de La Réclame, « La publicité digitale capterait moins l'attention que les autres médias²⁴⁵ » publié le 14 janvier 2019, on voit que « l'attention » accordée aux publicités varie en fonction des médias. Datakalab cherche aussi à trouver de nouveaux marchés où il est possible de collecter des données de manière continue comme l'expliquait déjà Xavier Fischer lors de notre entretien²⁴⁶ : « Les *market research*²⁴⁷ collectent des données en discontinu, étude par étude, tandis que les marchés où tu utilises les algorithmes tous les jours, toutes les minutes, sont encore plus intéressants pour nous ». Ayant désormais quitté Datakalab, je demande fin octobre 2019 à voir une présentation plus récente et comprend qu'une nouvelle offre est en train d'être développée. En ce sens Datakalab s'intéresse à des solutions pour capter les données des utilisateurs qui passent devant des affiches publicitaires dans la rue. Ces données ne seraient pas conservées, ce qui permettrait à Datakalab de respecter la réglementation sur la donnée personnelle.

²⁴² Comme Publicis, Havas Paris, Buzzman, TBWA, Ogilvy, La Chose, Dragon Rouge

²⁴³ Comme Havas Media, OMD, Dentsu Aegis Network, Phd, Starcom

²⁴⁴ Comme Lagardère, JCDecaux, Mediatransports, Clear Channel, France TV publicité (avec qui elle a déjà travaillé), Adexo, (qui travaille déjà avec Ittention, l'un des concurrents de Datakalab), et de régies presse comme La Tribune

²⁴⁵ Carcolse Élodie, La Réclame, « La publicité digitale capterait moins l'attention que les autres médias », 14/01/2019, « disponible à l'adresse : <https://lareclame.fr/mymedia-alpha-attention-digital-212390> »

²⁴⁶ ANNEXE 6, Entretien avec Xavier Fischer, Chief Executive Officer de Datakalab, 26/07/2019, Paris 17^{ème}

²⁴⁷ Les instituts d'études et tous les autres organismes qui collectent des données sur les consommateurs



Solution

State of the art deep learning algorithms to quantify anything on people faces.



© Copyright Datakalab 2019

Solution Datakalab : *state of the art deep learning algorithms to quantify anything on people faces*²⁴⁸ - support de présentation de octobre 2019

On peut pousser encore plus loin la personnalisation. Les nouveaux algorithmes de Datakalab sont capables de détecter de nouvelles informations, et Datakalab réfléchit à leurs utilisations. En quoi le fait que le panéliste — voire même le citoyen si on se concentre sur les caméras placées dans les rues — porte un chapeau, peut-il aider les annonceurs, instituts d'études, régies, agence de communication ou média ? Peut-on implémenter un système qui détecte et enregistre ces informations, lorsque l'on passe devant une boutique ou bien devant une affiche ? C'est ce que Datakalab propose dès octobre 2019.

3.4.2 Continuer à théâtraliser son expertise tout en rassurant

Datakalab doit poursuivre ses efforts en communication. Elle semble partir dans la bonne direction en employant l'anglais sur les réseaux sociaux : LinkedIn lui permet de toucher une cible experte, et Facebook de se vulgariser. Les newsletters contiennent un certain nombre d'éléments et incitent à se renseigner encore davantage. Une refonte intégrale du site internet s'avère néanmoins nécessaire pour être en accord avec la nouvelle focalisation sur le visage. Sa qualité particulière doit apparaître sitôt ouverte la page d'accueil. La manière de présenter appelle d'emblée l'attention et peut prédisposer favorablement le visiteur à s'intéresser à ce qui va être montré. Cette refonte du site est en cours. J'ai été amenée à faire des propositions. Certaines des idées ci-dessous ont été validées par la direction. La start-up compte par exemple

²⁴⁸ « Je traduis : des algorithmes issus de l'état de l'art du *deep learning* pour tout quantifier sur le visage humain ».

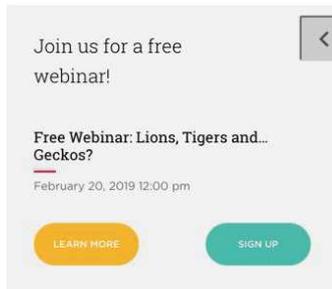
redynamiser son site avec un fond interactif, qui évoque les émotions. Des citations clients sur le site, le nombre de projets accomplis, et des cas clients permettraient de convaincre. Mettre plus en avant les marques qui font confiance s'avère crucial. Jusqu'alors, il faut encore les faire défiler lentement de gauche à droite. On souhaiterait une vision directe des clients prestigieux : annonceurs, instituts d'études, régies et agences.



**Captures d'écran du site internet de Datakalab au 01/06/2019, page
« ACCUEIL » de Datakalab « disponible à l'adresse : <https://www.datakalab.com/> »**

Le site pourrait alors proposer des démonstrations qui permettraient de tester un contenu déjà en ligne, par exemple sur YouTube. Ainsi les décideurs se feraient-ils eux-mêmes une idée concrète des promesses de la technologie. La plateforme propriétaire de collecte des vidéos constitue l'une des forces principales de Datakalab. Montrer également une « démo » sur le site, comme le fait son concurrent Feng-GUI²⁴⁹ qui propose de tester du contenu en ligne depuis 2009 pourrait s'avérer pertinent. Par ailleurs il est rare que les articles sur Datakalab évoquent ses « métriques ». Il importe de les mettre en évidence sur le site, comme particulièrement intéressantes pour les instituts d'études. Seraient également bienvenus des éléments de type pédagogique. Le sujet étant encore peu connu en France, on peut faire plus de contenu pédagogique pour convaincre les décideurs. Ainsi l'on s'accorde à la fois à la nécessité de rassurer et à celle d'asseoir son crédit pour intervenir. La start-up pourrait offrir du contenu pédagogique en échange de l'adresse mail du visiteur pour enrichir la base de données de prospects. L'ajout d'événements type Webinars sur le site, comme la société de neuromarketing Sparkneuro, pourrait éventuellement inciter à renouveler les visites, mais nécessiterait un certain temps de production et de lancement.

²⁴⁹ « Disponible à l'adresse : « <https://feng-gui.com/> »



Pop-up sur la page internet de Sparkneuro, Webinar « disponible à l'adresse : <https://sparkneuro.com/webinars/> »

3.4.3 De multiples applications de la donnée émotionnelle témoignent d'un marché en plein « boom »

Datakalab a choisi la « reconnaissance des expressions faciales, l'une des solutions existantes pour reconnaître les « réactions ». Sur ce point, je me retrouve dans l'incapacité de faire davantage de recommandations dans la mesure où m'interroger sur la meilleure façon de capter les « ressentis » via ces choix technologiques ne relève pas de mon domaine d'études. Toutefois, on peut penser à de nouvelles applications. Maintenant que la start-up a défini clairement son expertise, elle peut continuer de trouver de nouveaux marchés. Ainsi, lors de discussions et d'échanges informels, j'ai pu interroger différents acteurs sur des salons sur les possibilités pour le produit Datakalab. Je me suis aussi focalisée sur les réponses données lors des entretiens présents en annexe.

3.4.4 Datakalab doit s'appuyer sur l'aspect « humain »

Datakalab traite d'un sujet sensible, et d'informations personnelles. En France, les technologies de « reconnaissance des expressions faciales » commencent à se développer, mais cette avancée intervient de manière assez tardive par rapport à d'autres pays comme la Chine ou les États-Unis. Premier laboratoire au monde dans ce domaine, la Chine se sert de la reconnaissance faciale pour signaler les piétons qui traversent au rouge, les automobilistes ivres par exemple. Les taxis Japonais vont proposer à leurs clients des spots ciblés suite à la reconnaissance de leur âge et de leur sexe. Datakalab entend continuer à respecter l'éthique, enjeu central pour un tel sujet. Cette volonté est partagée par l'ensemble des membres fondateurs. Dans les messages de présentation des « études émotionnelles » adressés aux panélistes, les objectifs sont clairement énoncés. Datakalab construit alors une relation de confiance en étant la moins intrusive possible. Parler uniquement de « réactions » permet certes de s'éloigner des imaginaires diaboliques autour du neuromarketing, mais casse l'aspect

« féérique » de la promesse de mesure émotionnelle, au profit d'un cadre qui semble plus rationnel. Néanmoins, pouvoir capter des données aussi précises déborde une fois encore toute rationalité.

La révolution de la donnée émotionnelle peut être conçue comme un véritable progrès sociétal notamment pour le marché de la santé. Ces données promettent notamment d'améliorer la santé mentale de personnes vulnérables ou malades. C'est tout le travail de Datakalab avec l'institut Rafaël, la Maison de l'après cancer²⁵⁰. Datakalab prévoit de filmer les réactions des volontaires pour aider les médecins à faire passer plus d'émotions dans leurs messages pour encourager les patients à prendre leur traitement. Dans le milieu hospitalier, les médecins sont souvent débordés et n'ont que peu de temps à accorder aux patients. Replacer l'humain au centre de leur discours s'avère crucial. On peut également penser à d'autres applications comme la télémédecine, la détection de la douleur. Datakalab pourrait passer des partenariats avec des hôpitaux, tester les affiches préventives qui y figurent. Des sociétés comme Affectiva utilisent les données émotionnelles pour détecter et prévenir les risques comme la somnolence en voiture, ainsi que les états émotionnels du conducteur grâce à des caméras et des micros.

En ce qui concerne l'éducation, la reconnaissance des expressions faciales peut être utilisée sur bien des plans comme l'affirme Anne-Marie Gaultier²⁵¹. En milieu scolaire par exemple, pour identifier ce qui n'a pas été compris, ce qui a ennuyé, voire irrité. Toutes les réactions suscitées chez les élèves durant les années d'études deviennent alors elles-mêmes objet d'étude, dans le but d'améliorer les conditions de la vie scolaire et de pallier l'échec. La manière dont les professeurs doivent présenter un contenu pour faire retenir un message pourrait être explorée d'une façon neuve. Tout jugement porté sur la pensée neuroscientifique pourrait se faire alors en meilleure connaissance de cause. Par conséquent, la peur que suscitent encore souvent ses applications s'en trouverait éloignée. Informer de ses enjeux, c'est en effet éloigner la peur que cette pensée neuroscientifique engendre encore chez de nombreux consommateurs. Les discours autour de ses outils promettent un progrès sociétal. L'extension de ses applications à des domaines divers ne peut que contribuer à la rendre familière.

²⁵⁰ ANNEXE 5, Entretien avec Soukeina Hassanaly, Ex Key Account Director chez Datakalab, 05/04/2019, Paris 17^{ème}

²⁵¹ ANNEXE 1, Retranscription de la conférence « Datakalab au centre Pompidou ! » du 15/09/2018

Finalement, on peut penser à de nombreuses applications de la donnée émotionnelle dans le secteur des affaires. Datakalab pourrait « mesurer » des applications, des logiciels, en plus des sites internet. D'autres pistes peuvent être creusées comme celle du jeu vidéo. La technologie Datakalab permettrait par exemple de tester des cinématiques pour les lancements, voire pour le jeu vidéo lui-même. Leur très grand succès faciliterait probablement le recrutement des panélistes. Cette donnée est si prometteuse qu'on peut finalement lui fantasmer tous types d'utilisations : artistique (évaluer le niveau technique des danseurs), rhétorique (elle permettrait par exemple à tout type de décideur de s'entraîner à être impactant, dans un discours politique, un « pitch », dans le domaine de la communication). En ce qui concerne le sport, certaines stratégies de match sont pilotées par des intelligences artificielles, il existe des modélisations de comportement. On peut penser à ces nouvelles technologies pour optimiser de l'efficacité des équipes, sortir des joueurs en difficulté physique, voire progresser dans l'éducation physique et sportive de manière plus générale. Sur les salons, de nombreuses personnes sont aussi intéressées par l'utilisation de cette technologie pour des problématiques relatives aux ressources humaines. Mais l'accord des personnes est nécessaire pour récolter leurs données personnelles, et Datakalab se soucie réellement de l'utilisation de sa technologie. C'est l'une des raisons pour lesquelles ce marché, bien que sans doute porteur, n'est pas exploré par la start-up.

3.4.4.1 Une accélération du développement de Datakalab *via* un développement à l'échelle internationale

Réfléchir à l'échelle internationale impliquerait une certaine refonte du « business model » de Datakalab. De nombreuses sociétés de neuromarketing existent à l'étranger, mais en France, Datakalab est la seule société de « facial coding ». Elle doit prendre le meilleur de ses concurrents, et s'éloigner du pire. Parmi les principaux concurrents de la start-up Affectiva, Kairos, Realeyes, certains comme Affectiva permettent à n'importe quel consommateur de tester ses propres contenus en ligne. Datakalab pourrait faire de même et continuer de réfléchir à l'ouverture de sa plateforme à tous les utilisateurs au lieu d'être réservée aux utilisations en interne. Ce point a été évoqué lors des entretiens avec Alexandre Macedo²⁵², « Software

²⁵² ANNEXE 13, Entretien avec Alexandre Macedo, Software Engineer chez Datakalab, 03/06/2019, Paris 17^{ème} : « Je pense que pour les utilisateurs, c'est plus intéressant de pouvoir analyser directement leurs contenus, mais il y a beaucoup de travail à faire pour qu'on puisse y parvenir »

Engineer », et Maureen Kahn²⁵³ « Project Manager » chez Datakalab. Ce faisant, on donnerait les clés d'accès de ces données « aux profanes ». Cela impliquerait aussi un changement de cible, soit de s'adresser aux consommateurs qui postent régulièrement du contenu, et donc de passer en « B to C ». Datakalab peut adresser de nouveaux marchés où un flux massif d'images et de vidéos défilent constamment. C'est la définition des réseaux sociaux, des applications de rencontre. Datakalab pourrait tester de tels contenus. En vogue sur les réseaux sociaux, « gifs », « memes » pourraient aussi être analysés. En revanche, Datakalab prendrait le risque de devenir une énième application d'optimisation visuelle. Il faut encore et surtout, réfléchir à la difficulté de contrôler l'utilisation de cette technologie si elle est rendue publique.

Pour synthétiser l'ensemble de ces recommandations, je retiendrai d'abord la nécessité de se concentrer sur les nouvelles cibles et de réfléchir à plusieurs combinaisons d'expertises qui tendent vers une segmentation hyper personnalisée. Pour convaincre l'ensemble de ces acteurs, Datakalab doit continuer de théâtraliser son expertise tout en rassurant sur ses utilisations. Pour ce faire, une refonte intégrale de son site actuel doit être envisagée. Elle peut s'appuyer sur l'aspect humain dont elle se réclame *via* sa vision : « redonner visage humain à la donnée ». Dans tous les cas, ce respect de l'éthique s'avère primordial, et ce d'autant plus que le marché de la donnée émotionnelle se développe très rapidement. On peut alors appliquer cette donnée à de multiples marchés comme la santé et l'éducation. Les possibilités d'applications semblent infinies et le marché florissant. C'est bien en ce sens que la start-up doit rapidement s'intéresser aux marchés étrangers et se développer à l'international. Envisager « d'ouvrir sa plateforme » à tous les utilisateurs comme le font déjà certains de ses concurrents semble aussi pertinent pour gagner du temps sur les analyses. Parmi toutes ces possibilités, au regard des rachats de plusieurs start-up en neuromarketing par les « GAFAM », on peut supposer qu'une proposition sera faite à Datakalab.

²⁵³ ANNEXE 8, Entretien avec Maureen Kahn, Project Manager chez Datakalab, 02/07/ 2019, Paris 17^{ème}, « Le fait d'ouvrir la plateforme permettrait de passer moins de temps sur les études au bureau »

CONCLUSION

On ne peut que souligner l'exploit de Datakalab : parvenir à faire passer une donnée inconnue en 2016, la « feel data© » à un essentiel dans la boîte à outil des décideurs. Amorce d'une véritable révolution, cette donnée entend dépasser la frontière communément admise entre homme et machine. La donnée émotionnelle peut être appréhendée comme étant issue de plusieurs tendances. « Feel data© » est ainsi née dans la nuit du 7 au 8 mars 2016 au Salon Big Data à Paris. On peut considérer que ses parents, « big data » et « émotion », lui ont légué leurs noms. En effet, on ne compte plus les discours tenus autour de ces tendances et tous ceux qu'elle produit. Le contexte d'exposition permanente des individus aux contenus de marques oblige les publicitaires à se renouveler. Jouer sur l'émotion devient la nouvelle nécessité. C'est bien à partir de tous ces imaginaires combinés que Datakalab promet une « mesure émotionnelle ». Grâce à des discours séduisants, la start-up explique pouvoir mesurer l'immesurable, rationaliser et raisonner l'irrationnel, voire rendre visible l'invisible. Montrer l'évidence et le naturel de ces données est d'autant plus complexe pour Datakalab que tout son cœur de métier repose précisément sur le traitement de ces données brutes. Afin de convaincre d'y faire appel, elle s'est renseignée sur les imaginaires que véhiculent les mêmes outils dont elle se servait. Comme les études déclaratives, le neuromarketing cherche à identifier les motivations clients, à mieux connaître les profils des consommateurs, leur impact sur la perception des stimuli, à déclencher des réactions qui le pousseront à l'achat, et à prédire le comportement. À la différence des instituts d'études, le neuromarketing prétend à une nouvelle forme de « vérité » *via* une fonction maïeutique. Datakalab revendiquait ainsi un « ethos neuroscientifique » et décidait de s'installer dans *The Brain Place To Be*. Elle s'appuyait sur des théories comme celles d'Ekman ou de Damasio pour convaincre de leur opérativité.

Au-delà de tout ce discours de vente, restait à savoir si la mesure de Datakalab pouvait réellement appréhender la charge symbolique de nos émotions. Une « prétention sémiurgique » voire « démiurgique » se joue aussi en filigrane de la construction de la donnée. L'ingénieur ordonne le chaos, il est libre maître de sa création, choisit ses conditions d'existence, donne corps à l'immatériel tout en rendant visible l'invisible. Datakalab invente aussi des noms et crée des normes. Or Datakalab ne mesure pas tant les émotions que leur expression qui devrait plutôt être qualifiée de « ressenti » selon les définitions de ces termes en psychologie. Ma première intuition a donc été validée et dépassée puisqu'au temps de son expertise « conseil », certains acteurs de Datakalab se laissaient parfois eux-mêmes séduire par leur propre discours de vente.

En second lieu, je me demandais si Datakalab maintenait sa « promesse communicationnelle » de « mesure des émotions » malgré son repositionnement. Entre création d'algorithmes propriétaires de « facial coding », nouvelle expertise « technologique », la suppression de l'expertise « conseil » et de l'outil « bracelet connecté », on pouvait croire à une toute nouvelle promesse, différente de celle initiale. Or la nouvelle promesse conserve son cœur, sa fonction maïeutique et sa volonté de « donner visage humain à la donnée ». En ce sens elle conserve également ses pouvoirs heuristiques, pragmatiques et normatifs.

En passant de l'expertise « conseil » à l'expertise « technologique », Datakalab opère un changement complet de communication. Elle cesse d'y employer le terme de « neuromarketing » dans sa communication, et fait passer les émotions du premier rôle dans l'acte d'achat à l'un des rôles secondaires, l'une des multiples informations que ses algorithmes sont désormais capables de capter sur le visage. Datakalab cesse aussi de fournir des recommandations, délaissant la compétence humaine du Project Manager toujours dans une optique de spécialisation « technologique ». Datakalab s'inscrit ici dans les idéaux techno-centrés d'aujourd'hui. En cessant d'enrober ses résultats d'une dimension « conseil », elle les fait exister comme « suffisants » en tant que tels, soit comme de nouveaux « indices » de la réception des consommateurs aux différents contenus. Ici, les discours d'escorte tenus autour de la donnée émotionnelle continuent à forger la promesse. Datakalab ne revendique plus un savoir issu des neurosciences dans ses résultats. Cette même communication se fait désormais en anglais et tente d'atteindre un public étranger dans une même volonté d'élargissement et de développement. Elle promet non plus une « mesure des émotions » mais bien une « mesure des réactions ». Le passage du terme « émotion » à celui de « réaction » lui fait perdre sa dimension symbolique fantasmée de rationalisation de l'irrationnel ou de mesure de l'intangible. Cette nouvelle promesse semble plus ancrée dans la réalité. Promettre de « mesurer les réactions » devant un contenu implique ainsi une perspective plus large. Datakalab prend désormais en compte plusieurs critères que ses algorithmes propriétaires parviennent à identifier. Cette technologie peut détecter de nouveaux attributs, tels que l'âge, le sexe, l'orientation de la tête, le clignement des yeux, la présence de maquillage. En ce sens, la promesse semble désormais plus large.

Datakalab communique sur son nouveau point fort, à savoir la « prouesse technique » de ses chercheurs. Dépassant « l'ethos neuroscientifique » identifié en première partie, Datakalab revendique maintenant un « ethos algorithmique » qui lui permet de se crédibiliser et de faire croire à cette nouvelle promesse. On montre donc le processus de création qui restait naguère dans l'ombre. Avec la création de nouveaux algorithmes propriétaires

de reconnaissance des expressions faciales et le repositionnement d'une expertise « conseil » à une expertise « technologique », on constate une variation des conditions d'énonciation. De nouveaux discours d'escorte sont alors tenus autour des données émotionnelles et modulent donc l'ancienne promesse de mesure émotionnelle en changeant ses conditions d'énonciation.

Dans un dernier temps, je me demande si les discours modulés en fonction des besoins des nouveaux clients permettent d'accroître l'opérativité sociale des données émotionnelles. Il existe un double bénéfice pour Datakalab et ses nouvelles cibles, qui se montrent en train d'innover à la fois aux yeux de leurs collaborateurs et de la presse. Tenant des discours autour de la performance *via* la réduction des investissements média par exemple, Datakalab parle la même langue que les décideurs qu'elle cherche à séduire. En retour, ils se prêtent au jeu de Datakalab et introduisent des données émotionnelles à leurs présentations. Ils s'approprient et adaptent alors le discours de la start-up aux enjeux de leur entreprise. Dans ces mêmes processus s'opèrent alors de nouvelles réécritures des discours d'escorte de la « feel data© ». Donnant les clés d'accès à son savoir, Datakalab aide ces acteurs à comprendre cette donnée. En contre-don²⁵⁴, Datakalab reçoit des informations plus spécifiques sur les cibles et sur le marché. Toutes ces médiations contribuent alors à en créer de nouvelles, chacun de ces acteurs produisant toujours plus de discours autour de cette donnée, ce qui favorise encore la posture communicationnelle de Datakalab. On assiste donc à un cercle vertueux. En fonction des domaines d'études de ses clients, Datakalab adapte sa méthodologie, choisit d'intégrer des messages différents dans ses présentations, et va parfois même jusqu'à adapter ses « métriques » au secteur de son client comme elle a pu le faire avec la société Do The Right Films. Cet exemple illustre que la start-up peut personnaliser ces indicateurs initialement créés pour répondre de manière générale aux besoins d'une « société de sursollicitation ». Les « discours d'escorte » sont donc bien modulés pour être performants, convenir aussi bien aux clients qu'au nouveau contexte politique, économique, technologique et sociétal dans lequel ils s'insèrent. Ma dernière hypothèse se trouve donc validée.

Datakalab avait élaboré sa stratégie de « repositionnement » dès sa création. Tant que l'on dispose d'un point de vue extérieur à Datakalab, on ne peut pas appréhender une telle réflexion. En six mois, je considère aussi n'en avoir eu qu'un aperçu. Toutes les modifications demandées s'avéraient réfléchies. Ainsi, le choix de se repositionner sur la technologie semble tout à fait pertinent au regard des exigences de production, de performance, d'innovation, de

²⁵⁴ Notion développée par Marcel Mauss dans *Essai sur le don*, « Forme et raison de l'échange dans les sociétés archaïques », PUF, « Quadrige Grands textes », 2007

chiffre et de la confiance accordée aux algorithmes dans notre société. Pour « prêcher » la donnée émotionnelle, Datakalab doit alors continuer à montrer son expertise, ce qui passe notamment par la refonte de son site internet. Tirer le fil rouge initial et « redonner visage humain à la donnée émotionnelle » semble aujourd'hui crucial. Cette vision répond particulièrement bien à la promesse de « mesure des réactions », qui implique elle-même d'envisager de nouvelles utilisations de ce que l'algorithme propriétaire détecte désormais. En ce sens elle devrait se concentrer sur des marchés comme la santé et l'éducation. Tout en continuant à rassurer, Datakalab doit réfléchir à de nouvelles utilisations de cette donnée *via* des marchés qui permettent de récolter des données en continu, comme par exemple les réseaux sociaux, les caméras dans les rues, notamment à l'échelle internationale et garder en tête que certaines propositions peuvent lui être faites de la part des « GAFAM ».

Le sujet s'avère si nouveau et innovant qu'il mériterait donc un traitement bien plus en profondeur. On pourrait ainsi produire des études entières sur chacun des objets apparus en étant parti des « feel data© ». Le « bracelet connecté » et le « facial coding » portent de nombreuses promesses, qui opèrent *via* les imaginaires issus d'un nouveau contexte politique, économique, sociétal et technologique. Ils mériteraient eux-mêmes un traitement toujours plus poussé. Aussi, ce travail pourrait être prolongé en s'attachant aux promesses des autres produits Datakalab. Une étude serait ainsi portée sur le produit que Datakalab nomme « Datakalab Event », à savoir la mesure de conférences, débats, spectacles. Il conviendrait aussi d'analyser dans une nouvelle recherche uniquement le produit « Datakalab Website » qui promet de « mesurer les réactions » des panélistes face à un site internet. Ces deux autres produits semblent fascinants. L'analyse de site internet est un produit de Datakalab toujours en développement qui n'a pour l'instant été implémenté qu'auprès d'un seul client, Éram. Il est compliqué de collecter un nombre suffisant de panélistes qui acceptent de participer à de telles études, qui impliquent déjà d'accepter de donner accès à sa caméra et de visiter le site de Éram en cliquant sur plusieurs onglets. Le traitement de ces données semble particulièrement délicat et complexe : Comment savoir quel stimulus a engagé quelle réaction ? Or le sujet « promesses communicationnelles des données émotionnelles de Datakalab » est déjà relativement précis. Avec ce repositionnement et toutes les médiations qui en découlent, l'occasion se présentait de s'intéresser aux réécritures sur l'ensemble de ces outils. Mes réflexions peuvent donc leur être appliquées.

Les difficultés de Datakalab pour instituer ses données émotionnelles en tant « qu'indice » peuvent grâce à ce traitement général être déplacées sur ces autres objets d'études.

Datakalab utilise encore d'autres techniques, à savoir le « predictive eye tracking », qui promet d'estimer où le regard va se porter sur un visuel. Elle développe aussi sa propre technologie « d'eye tracking » comme l'explique Edouard Yvinec lors de notre entretien²⁵⁵, afin que ce savoir ne soit plus le résultat d'un traitement de bases de données, mais issu d'une expérience avec des panélistes. Or les promesses de ces mêmes outils mériteraient aussi d'être explorées. L'objet de recherche continue d'évoluer très rapidement. Ce travail constitue alors une définition momentanée des « promesses communicationnelles » autour de la « feel data© » de Datakalab. Ce travail peut être prolongé en continuant de se renseigner sur l'actualité de la start-up. Il semble qu'on ne peut que suivre Datakalab, qui anticipe et choisit constamment les chemins de sa réussite.

²⁵⁵ ANNEXE 9, Entretien avec Edouard Yvinec, Machine Learning Scientist Junior chez Datakalab, 14/06/2019, Paris 17^{ème}

BIBLIOGRAPHIE

OUVRAGES

Baudrillard, Jean, *Pour une critique de l'économie politique du signe*, Paris : Gallimard, Coll. « Les Essais » (n° 168), 1972

Baudrillard, Jean, *Le Système des objets*, Paris, Gallimard, 1968

Damasio, Antonio, *L'erreur de Descartes : la raison des émotions*, Paris, Odile Jacob, 1995

Jeanneret, Yves, *Critique de la trivialité*. « Les médiations de la communication, enjeu de pouvoir », Paris : Editions Non Standard, 2014

Perec, Georges, *L'Infra-ordinaire*, Paris, Seuil, 1989

ARTICLE PRESSE PROFESSIONNELLE

Neuromarketing

Binninger, Anne-Sophie, Goudey, Alain, « RMS, Le neuromarketing : nouvel eldorado de la recherche ou dérive manipulatoire ? » *Monde des grandes Ecoles et Universités, Marketing Grandes Ecoles* n°11,04/2012 « disponible à l'adresse : <http://www.mondedesgrandesecoles.fr/rms-le-neuro-marketing-nouvel-eldorado-ou-derive-manipulatoire/> » « date de dernière consultation : 03/08/19 »

Chamak, Brigitte, « Neuromarketing : une fusion suspecte », *Presses Universitaires de Paris Nanterre*, pp.113-123, 2017 « disponible à l'adresse : « <https://halshs.archives-ouvertes.fr/halshs-01498133/document> » « date de dernière consultation : 20/09/19 »

Gavard, Emmanuel, « Le Fantasma du neuromarketing », *LSA Commerce et Consommation*, 17/05/2012 « disponible à l'adresse : <https://www.lsa-conso.fr/le-fantasme-du-neuromarketing,130326> » « date de dernière consultation : 07/05/19 »

« La start-up française Moodstock rachetée par Google », *Les Échos*, AFP, 6/07/2016

« disponible à l'adresse : <https://www.lesechos.fr/2016/07/la-start-up-francaise-moodstocks-rachete-par-google-212584> » « date de dernière consultation : 10/07/19 »

Nebia Amelle, CB News, « Datakalab poursuit sa croissance avec de nouveaux clients », 13/12/2017, disponible à l'adresse : « <https://www.cbnews.fr/conseil/image-datakalab-poursuit-sa-croissance-avec-nouveaux-clients-27143> »

Reuters, « Facebook rachète Face.com le spécialiste de la reconnaissance faciale », Journal 20 minutes, 18/06/2012

« disponible à l'adresse : <https://www.20minutes.fr/web/955883-20120618-facebook-rachete-facecom-specialiste-reconnaissance-faciale> » « date de dernière consultation : 10/07/19 »

Données émotionnelles

Benalal, William, Gaultier, Anne-Marie, « IBM et Datakalab séparent le bon grain de l'ivraie avec une IA qui détecte les émotions », HUB Institute, 22/04/2019 « disponible à l'adresse : <https://hubinstitute.com/2019/hubday/data/replay-IBM-WilliamBenalal-IA-emotions-experience-client> » « date dernière consultation : 25/04/19 »

Destribats, Juliette, « L'émotion des spectateurs au cœur de la mesure cinéma Médiamétrie », Médiamétrie, communiqué de presse, 11/07/2019, « disponible à l'adresse : <https://www.mediametrie.fr/fr/lemotion-des-spectateurs-au-coeur-de-la-mesure-cinema-de-mediаметrie> » « date de la dernière consultation : 23/07/19 »

De Matharel, Lélia, JDN « Datakalab dissèque les émotions pour aider les pros à booster leur CA » « JDN » 19/06/2017 « disponible à l'adresse : <https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1195466-datakalab-disseque-emotions/> » « date de dernière consultation : 25/04/19 »

« L'avenir du Big Data sera émotionnel » « Influencia » « revue digitale » 7/03/2016 « disponible à l'adresse : <http://www.influencia.net/fr/actualites/tendance,tendances,avenir-big-data-sera-emotionnel,6132.html?> » « date de dernière consultation : 25/04/19 »

Loye, Deborah, Les Echos, « CES 2019: Datakalab transforme les émotions en données »
« Disponible à l'adresse : <https://www.lesechos.fr/tech-medias/hightech/0600478877092-ces-2019-datakalab-transforme-les-emotions-en-donnees-2234962.php> », 09/01/ 2019, « date de dernière consultation : 01/11/19 »

Lagarde, Émilie, « L'IA friande de nos émotions » CBNews marques, 11/11/2017 « disponible à l'adresse : <http://www.cbnews.fr/cb-info/l-ia-friande-de-nos-emotions-a10386> » « date de dernière consultation : 23/05/19 »

Lunel, Bénédicte, « Mesurer l'engagement émotionnel en seulement trois mots », Influencia média numérique, 25/03/2016 « disponible à l'adresse : <http://www.influencia.net/fr/actualites/tendance,observatoire-influencia,mesurer-engagement-emotionnel-seulement-3-mots,6201.html> » « date de dernière consultation : 05/07/19 »

L'Effet Papillon, Revue digitale « Les Feel data, rêve de la neuroscience ou réalité du marketing ? » 27/02/2018
« disponible à : <http://www.effetpapillon.fr/feel-data-reve-de-neuroscience-realite-marketing/> » « date de dernière consultation : 23/05/19 »

Marchalant, Charlotte, « La feel data© décrypte les secrets des cœurs (des consommateurs) », emarketinf.fr, 8/03/2016 « disponible à l'adresse : <http://www.e-marketing.fr/Thematique/data-1091/Breves/Big-Data-2016-Feel-Data-decrypte-secret-coeurs-consommateurs-302697.htm> »

Nebia, Amelle, CBNews« Datakalab prend le pouls des émotions sur les sites e-commerce », 08/01/2019 « Disponible à l'adresse : http://www.cbnews.fr/conseils/datakalab-prend-le-pouls-des-emotions-sur-les-sites-e-commerce-a1047227?utm_source=newsletter&utm_medium=email » « date de dernière consultation : 01/11/19 »

Noël, Steffi, « Emotional Data in China : when emotions become measurable », daxue consulting, 17/07/2019, « disponible à l'adresse : <https://daxueconsulting.com/emotional-data-china/> » « date de dernière consultation : 01/09/19 »

Richebois, Véronique, « Datakalab : Intrusion dans la tête d'un client », Les Échos, 07/03/16, « disponible à l'adresse : <https://www.lesechos.fr/2016/03/datakalab-intrusion-dans-la-tete-dun-client-214441> » « date de la dernière consultation : 25/04/19 »

Robillard, Olivier, « Les IA ne comprennent rien à nos grimaces », L'ADN, 15/05/2018, « disponible à l'adresse : https://www.ladn.eu/entreprises-innovantes/transparence/reconnaissance-faciale-les-ia-ne-comprennent-rien-a-nos-grimaces/?utm_source=newsletter_ladn&utm_medium=email&utm_campaign=news_mai_2018&utm_content=20180515 » « date de la dernière consultation : 01/09/19 »

Soulas-Gesson, Delphine, Stratégies, 15/03/2019 « Disponible à l'adresse : <http://www.strategies.fr/actualites/medias/4026275W/exterion-media-s-allie-avec-datakalab-sur-la-mesure-des-emotions.html> », « date de dernière consultation : 01/11/19 »

Sousa-Tagaroso, Laetitia, « Les feel data, nouveau levier de mémorisation pour les marques ? » ESP Ecole Supérieure de Publicité, 22/01/2018, « disponible à l'adresse : <https://espub.org/feel-data/> » « date de dernière consultation : 23/05/19 »

ARTICLES NEUROSCIENTIFIQUES

Eustache, Francis, Desgranges, Béatrice, Giffard, Bénédicte, Guillery-Girard Bérengère, « Émotions et mémoire », RP Ressources plurielles, publication numérique, unité Inserm V1077-EPHE-UNICAEN\, article paru dans Cerveau et Psycho n°81, 08/2016, « disponible à l'adresse : « <http://ressources-plurielles.com/emotions-memoire/> » « date de dernière consultation : 02/09/19 »

Eustache, Francis, Desgranges, Béatrice, Guillery-Girard Bérengère, « Comment les émotions forgent nos souvenirs », « Cerveau et Psycho » 21/09/2016, article en ligne « disponible à l'adresse : <https://www.cerveauetpsycho.fr/sd/neurobiologie/comment-les-emotions-forgent-nos-souvenirs-9248.php> » « date de dernière consultation : 02/09/19 »

Van Hoorebeke, Delphine, « L'émotion et la prise de décision » « Revue française de gestion 2008/2 », n°182, pp 33-44, Lavoisier librairie en ligne, « disponible à l'adresse : <https://www.cairn.info/revue-francaise-de-gestion-2008-2-page-33.htm> » « date de dernière consultation : 02/09/19 »

DIVERS

Bichler, Camille, « Prix Turing : le Français Yan Le Cun couronné avec deux autres pionniers de l'Intelligence Artificielle », France Culture, Podcast, 27/03/2019, « disponible à l'adresse : <https://www.franceculture.fr/sciences/le-prix-turing-couronne-trois-pionniers-de-lintelligence-artificielle> » « date de dernière consultation : 10/09/19 »

Brulois, Vincent, Robert-Tanguy, Philippe, « Le communicant, force fragile : une identité de métier en question », Sociologies pratiques 2015/1, n°30 PP 1-7, CAIRN INFO, « disponible à l'adresse : « <https://www.cairn.info/revue-sociologies-pratiques-2015-1-page-1.htm> » « date de dernière consultation : 03/09/19 »

Carcolse, Élodie, « La publicité digitale capterait moins “l'attention” que les autres médias », La Réclame, Journalisme web-print “Vie numérique”, nouvelles technologies et pop culture 14/01/2019, « disponible à l'adresse : <https://lareclame.fr/mymedia-alpha-attention-digital-212390> » « date de dernière consultation : 23/05/19 »

Colombain, Jérôme, « Nouveau monde. Faut-il craindre la reconnaissance faciale ? » France Info, 23/05/2019 « disponible à l'adresse : https://www.francetvinfo.fr/replay-radio/nouveau-monde/nouveau-monde-faut-il-craindre-la-reconnaissance-faciale_3434539.html » « date de dernière consultation : 09/06/19 »

Padiolleau, Agathe, « La Computer Vision, ce n'est pas si compliqué ! », 26/06/2018, « disponible à l'adresse : <https://deepomatic.com/fr/computer-vision-definition-et-exemples/> », « date de dernière consultation : 04/11/19 »

Von Abrams, Karine, « Ad blocking in France 2018 » « Combien de personnes en France utilisent des adblockers », Emarketer, 4/09/2018, « disponible à l'adresse : <https://www.emarketer.com/content/ad-blocking-in-france-2018> » « date de dernière consultation : 23/05/19 »

MÉMOIRES

Bélouis, Julia, « Les émoji, une pratique textuelle ordinaire ? Etude des mutations d'un objet trivial en tension entre pratique et encadrement » Mémoire, 07/07/2017

« disponible à l'adresse : https://www.academia.edu/34791274/Les_emoji_une_pratique_textuelle_ordinaire_Etude_des_mutations_d_un_objet_trivial_en_tension_entre_pratique_et_encadrement » « date de dernière consultation : 12/10/19 »

Ducrot, Alexandra, « Discours, objets et imaginaires autour des « technologies big data » : Le cas des consultants en technologie, la construction d'une expertise à travers la « raison algorithmique » Mémoire, promotion 2016-2017

ARTICLES SIC

Aïm, Olivier, « La transparence rendue visible. Médiations informatiques de l'écriture » fait partie d'un numéro thématique « Internet, optique du monde » pp. 31-45 Persée, 2006 disponible à l'adresse : « https://www.persee.fr/doc/colan_0336-1500_2006_num_147_1_4575 » « date de dernière consultation : 25/09/19 »

Bouveresse, Jacques, « Le mythe du progrès selon Wittgenstein et von Wright » « Revue Mouvements 2002/1, n°19, pages 126 à 141 » « disponible à l'adresse : <https://www.cairn.info/revue-mouvements-2002-1-page-126.htm> » « date de dernière consultation : 26/09/19 »

Galinon-Méléneq, Béatrice, Zlitni, Sami, « L'Homme-trace, producteur de traces numériques », CNRS 'éditions, pp.7-19, CNRS, 04/11/2014 « disponible à l'adresse : https://www.researchgate.net/publication/278759869_L'Homme-trace_producteur_de_traces_numeriques » « date de dernière consultation : 10/08/19 »

Jeanneret, Yves, « Faire Trace : Un dispositif de représentation du social », fait partie d'un numéro thématique « De la trace à la connaissance à l'ère du Web » Intellectica, 59, p.41-63, 2013 « disponible à l'adresse https://www.persee.fr/doc/intel_0769-

4113_2013_num_59_1_1085 » « date de dernière consultation :12/10/19 »

Jeanneret, Yves, Flon, Emilie, Davallon, Jean, Tardy, Cécile, Cécile Tardy, « Traces d'écriture, traces de pratiques, traces d'identités », hypermédiat 2009 (H2PTM 09), Rétrospective et perspective, 1989-2009, Hermès science publications, 2009, p. 181-191, « disponible à l'adresse : <http://www.gripic.fr/productions-scientifiques/traces-decriture-traces-pratiques-traces-didentites> » « date de dernière consultation :15/08/19 »

Latour, Bruno, *Les "Vues" de l'esprit*, Culture et technique, Réseaux, volume 5, n°27, 1987, Questions de méthode. pages 79-96, « disponible à l'adresse : https://www.persee.fr/doc/reso_0751-7971_1987_num_5_27_1322 » « date de dernière consultation : 12/10/19 »

Pélissier, Daniel, « Trace, vous avez vraiment dit trace ? », Présence numérique des organisations, OpenEditions, 19/06/2015, « Disponible à l'adresse : <https://presnumorg.hypotheses.org/76> » « date de dernière consultation : 12/10/19 »

Pélissier, Daniel, « De quoi les traces numériques sont-elles le nom ? », 27/08/2015, Présence numérique des organisations, OpenEdition, « disponible à l'adresse : <https://presnumorg.hypotheses.org/94> » « date de dernière consultation : 12/10/19 »

Rieder, Bernard, « Pratiques informationnelles et analyse des traces numériques : de la représentation à l'intervention », OpenEdition, Etude et communication, 01/12/2012, « disponible à l'adresse : <https://journals.openedition.org/edc/2249> » « date de dernière consultation 10/10/2019 »

Rosa, Harmut, « Accélération. Une critique sociale du temps », compte rendu d'Élodie Wahl, OpenEditions, 05/11/2013, « disponible à l'adresse : <https://journals.openedition.org/lectures/990> » « date de dernière consultation : 26/09/19 »

Rouvroy, Antoinette, BERNS Thomas, « Gouvernamentalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », Réseaux, 2013, n° 177, pp. 163-196. DOI 10.3917/res.177.0163 89

Rouvroy, Antoinette, Berns, Thomas, « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques », *Multitudes*, 2010, n° 40, pp. 88-103

Rouvroy, Antoinette, « Pour une défense de l'éprouvante inopérationnalité du droit face à l'opérationnalité sans épreuve du comportementalisme numérique », *Dissensus, Revue de philosophie politique de l'ULg*, N°4, Avril 2011, pp. 127-149

Sadin, Éric, *La vie algorithmique, critique de la raison numérique*, Paris, Éditions l'Échappée. DL 2015, 278 p. (quatrième de couverture), « disponible à l'adresse : <https://www.lechappee.org/collections/pour-en-finir-avec/la-vie-algorithmique> » « date de dernière consultation : 12/10/19 »

Sauvageot, Anne, « Quand Baudrillard nous aide à penser le monde contemporain », *Mondes sociaux, Magazine de sciences humaines et sociales*, 01/02/2017, mis à jour 24/01/2019, « disponible à l'adresse : <https://sms.hypotheses.org/8941> », « date de dernière consultation : 16/09/19 »

Souchier, Emmanuël, « Voir le Web et deviner le monde, La "cartographie" au risque de l'histoire de l'écriture », pp. 213-234, CNRS éditions, 2013 « disponible à l'adresse : <https://books.openedition.org/editionscnrs/21786?lang=en#ftn3> », « date de dernière consultation : 09/09/19 »

Soulié, Rémy, « Lectures. Le mythe du progrès » *Le Figaro*, 01/02/2011 « <http://www.lefigaro.fr/lefigaromagazine/2011/02/19/01006-20110219ARTFIG00581-lectures-le-mythe-du-progres.php> », « date de dernière consultation le 26/09/19 »

RÉSUMÉ

Datakalab lance en 2016 la « feel data© » et la présente comme une évidence qui va s'imposer dans la boîte à outil des professionnels. Celle-ci peut être considérée comme issue d'au moins deux tendances, le « big data » et le recours systématique des marques à l'émotion. La start-up doit alors montrer qu'il est possible de « mesurer les émotions » typiquement humaines grâce à la machine. La tâche s'avère d'autant plus difficile que le cœur du métier de Datakalab repose justement sur le procédé de construction de cette donnée pour lui donner un sens aux yeux des annonceurs. Pour alors convaincre que l'on peut lier deux domaines qui semblent *a priori* inconciliables, Datakalab se construit un « ethos neuroscientifique ». Datakalab cherche bien à convaincre qu'il est possible de ramener la captation de l'émotion à un traitement computationnel.

La création de technologie d'analyse des expressions faciales de Datakalab a amorcé la « bascule », d'une expertise « conseil » à une expertise « technologique ». Je suppose alors que la promesse a été conservée malgré la suppression de l'outil « bracelet connecté ». Si Datakalab jouait sur les imaginaires autour des savoirs neuroscientifiques lors de son lancement, en se repositionnant, elle utilise désormais ses nouveaux atouts : sa plateforme et ses algorithmes propriétaires, cherchant à développer son expertise « technologique » en se construisant un « ethos algorithmique ». La seconde hypothèse n'est qu'en partie validée dans la mesure où les « discours d'escorte » font varier cette promesse, passant de la mesure des différentes émotions à celle des réactions. Datakalab se concentre désormais de manière plus large sur l'ensemble des signes captés sur le visage.

Je me demande enfin si les discours modulés en fonction des besoins des nouveaux clients permettent d'accroître « l'opérativité » sociale des données émotionnelles. Datakalab entre en résonance avec les nouveaux impératifs du marché, à savoir, l'innovation, la performance, le chiffre et le calcul algorithmique, inventant de nouvelles applications à sa donnée émotionnelle. Elle adapte ses méthodologies et personnalise ses présentations et ses métriques pour différents secteurs. De fait, elle tente bien de répondre aux besoins de sa nouvelle cible, ce qui valide ma troisième intuition. Les « feel data© » promettent même de nouvelles utilisations, par exemple autour de la santé et de l'éducation. En ce sens, Datakalab a bien réussi à les instituer en tant « qu'indice » de la réception face aux contenus et opéré une véritable révolution en liant émotionnel et traitement computationnel.

MOTS CLÉS

« Prétentions communicationnelles », émotions, « facial coding », « feel data© », donnée émotionnelle, « indice », neuromarketing, « discours d'escorte », « gouvernamentalité algorithmique »

ANNEXE

ANNEXE 1

Retranscription de la conférence « Datakalab au centre Pompidou ! » du 15/09/2018

Vidéo postée le 01/09/18, Retranscription du début de la vidéo, de 00:00 à 3:46, « disponible à l'adresse <https://www.youtube.com/watch?v=GF-j5VRFlY> »

À l'occasion de l'intervention de Antonio Damasio lors du Colloque “Le Pouvoir des émotions” du 15/09/2018, Anne-Marie Gaultier alors Présidente de Datakalab parle de l'analyse des émotions grâce aux outils issus des neurosciences.

« Anne-marie Gaultier - je suis la « Présidente » de Datakalab qui est une start-up dans la brain-tech. Voilà donc, je vais vous parler de comment on mesure les émotions. Vous l'avez compris, aujourd'hui tout est émotions. On l'a vu, même les discours les plus cartésiens, même les discours de banquiers sont émotions — Sorry Mr Descartes. Aujourd'hui, on peut mesurer les émotions grâce à l'intelligence artificielle et grâce aux outils issus des neurosciences et notamment le facial coding. (...) »

Le *facial coding*, c'est un outil qui est basé sur les travaux de Paul Ekman. Qui est Paul Ekman ? C'est un chercheur Américain, qui dans les années 70 a répertorié toutes les micro-expressions du visage : le plissement des yeux, le froncement des sourcils... C'est la combinaison de ces micro-expressions qui permet d'identifier les 7 émotions universelles qui sont : la joie, la surprise, la peur, la colère, le mépris, la tristesse, [le dégoût]. Ces micro-expressions sont extrêmement brèves et quasi inconscientes. Par exemple je vais vous donner un petit truc pour reconnaître la joie. Un vrai sourire, c'est quand la commissure des lèvres remonte, les joues se lèvent et les yeux se plissent. Un faux sourire, pas de plissement des yeux, parce que la partie haute du visage est à commande involontaire. Deuxième exemple pour reconnaître la colère : les lèvres se pincent, les narines se dilatent, et les sourcils se froncent. Pour la surprise, elle dure moins d'une seconde. Si elle dure plus d'une seconde, c'est qu'elle est simulée. Donc là, on reconnaît la surprise et la bouche s'entrouvre, les yeux s'écarquillent et les sourcils se relèvent. En d'autres termes, un expert en micro-expressions est tout à fait capable d'identifier vos émotions. C'est exactement ce que nous faisons aujourd'hui grâce à la reconnaissance faciale,

grâce à l'intelligence artificielle et à nos algorithmes. Si nous filmons, nous sommes capables de démocratiser cette faculté d'identification des émotions.

Alors à quoi ça sert ? Ça sert par exemple à comprendre le ressenti des individus face à une pub ou à une communication, ou une émission de télévision. Mais je ne suis pas là pour vous parler de pub ou de com. Je suis là pour vous parler d'éducation, d'apprentissage, parce qu'on ne retient bien que ce qui engage chez nous une émotion positive ou négative. Les émotions vont stimuler le cerveau limbique en particulier l'hippocampe qui est responsable de l'encodage la mémoire courte vers la mémoire longue. Cet outil va nous permettre probablement de relever un des plus grands défis de la société d'aujourd'hui qui est l'éducation, la formation. Si je mets une caméra dans une salle et que je filme les élèves, je suis capable d'identifier quels moments engagent les élèves qui vont être la garantie de cette mémorisation, ou les moments qui laissent de marbre.

Maintenant j'aimerais vous parler de vous. En fait ce matin nous avons mené une petite expérience et nous avons filmé la salle. Et là c'est ce que vous voyez. Nous avons filmé la salle et là vous voyez les *faceboxes* qui identifient les visages et qui se colorent en fonction des émotions. Sur la droite de l'écran, vous voyez une courbe qui se construit. En fait cette courbe, c'est la courbe d'engagement émotionnel. Là c'est le pourcentage de personnes qui expriment une émotion positive ou négative lors de l'intervention.

ANNEXE 2

Pitch Frank Tapiro – Challenge *Start-me up!* KPMG & Fevad

Parmi les 56 dossiers déposés, Datakalab a fait partie des 5 start-up sélectionnées pour pitcher devant le top 100 e-commerce le 26 juin, lors des « Enjeux e-commerce » organisés par la Fevad. Frank Tapiro disposait de 5 minutes pour présenter la start-up et plus particulièrement son outil Datakalab Website.

« Vous vous attendiez à un jeune startupper, vous avez en échange un vieil entrepreneur. Bonjour je m'appelle Frank Tapiro et je suis co-fondateur et CEO, Chief Emotion Officer de Datakalab. Le vrai CEO, c'est mon associé Xavier Fischer qui est plus... jeune que moi.

Commençons par 2 chiffres pour poser l'ambiance : la production de contenus représente 300 milliards de dollars en 2019 et il y aura plus de 2,2 milliards d'acheteurs en ligne en 2021. Aujourd'hui les seules données utilisées en Ecommerce sont rationnelles : panier moyen, nombre de visites, nombre de clics... Ces données ne suffisent pas pour « mesurer » la performance. Il faut les enrichir avec des données émotionnelles. La conversation que nous avons et que vous avez avec vos clients tous les jours en magasin ou en ligne est avant tout relationnelle et émotionnelle. En magasin, on gère la relation, le merchandising, le visuel, l'ambiance olfactive, l'ambiance sonore... Sur un site Ecommerce les « drivers » sont encore trop rationnels. Les émotions sont responsables de tout, nos décisions, nos actions, nos souvenirs, nos achats. 90 % des stimuli émotionnels sont traités par notre inconscient. Ce qui laisse une petite part pour le conscient.

Si je vous demande de vous souvenir de toutes les émotions que vous avez ressenties seconde par seconde pour chaque pitch, vous n'y arriverez pas car nos cerveaux n'ont pas la capacité de stockage. C'est pour ça qu'on a créé Datakalab. Pour quantifier, structurer, analyser, stocker les émotions et les transformer en *data* qui sont l'or noir du commerce et demain, du Ecommerce. Pour que le E de Ecommerce veuille dire Emotionnel.

Nous proposons aux marques une Plateforme *B to B* qui leur permet de mieux comprendre leur audience pour optimiser l'expérience et la performance. Comment ? On a développé des algorithmes de *computer vision* qui nous permettent de détecter les émotions et les

comportements en filmant des gens en conditions réelles depuis chez eux pendant leur visite sur un site, avec leur accord bien sûr. Cet algorithme est plus puissant que celui de Facebook ou SenseTime et permet même de prendre en compte des aléas comme une main devant la bouche, un contre-jour ou une tête qui se tourne. Notre partenariat avec IBM Tealeaf nous permet d'avoir des données contextuelles de tout le parcours sur le site (taux de clics, durée de la session, nombres de pages visitées...). Dans le même temps, l'algorithme Datakalab transforme les vidéos en données émotionnelles.

Nous travaillons notamment pour le site de Éram avec une plateforme qui leur permet de voir les *data* en continu jour après jour. La première métrique c'est l'attention par page pour savoir si la personne regarde bien l'écran sachant que sur un site c'est vous qui *drivez* votre propre expérience à la différence d'un contenu vidéo qui vous est imposé. On peut voir que le *Lookbook* de printemps provoque peu d'attention mais beaucoup d'émotions positives. Éram doit donc optimiser l'attention au démarrage avec un visuel plus impactant. La 2^{ème} métrique est l'émotion par page. On observe que la page du *E-Lab* génère peu d'émotions. Il y a trop de texte et des marques pas assez connues. Éram a le choix de l'optimiser ou de la supprimer. Le blog (au bout à droite) ne génère ni attention ni émotion. Éram l'a supprimé. Pour la valeur de page, plus elle est élevée plus elle génère des émotions positives. On peut ainsi savoir combien d'euros vont être gagnés grâce à cette page en direct ou indirect. Nous avons aussi mesuré les émotions ressenties en fonction des couleurs de saison. [...]

Le classement émotionnel par type de cuir nous donne aussi de bonnes indications à la fois pour le site et les magasins, à condition que vous aimiez le *buffalo leather*, cuir de buffle. L'*emotional product ranking* nous permet de classer les produits par émotions ressenties. Éram peut ainsi présenter ces produits sur la *landing page* pour générer plus d'émotions, plus d'engagement et donc plus de performance. L'émotion est un prédicteur de vente et nous permet d'anticiper sur les produits à mettre en avant des semaines à l'avance. *Datakalab Website* permet d'optimiser le merchandising, l'acquisition client, la stratégie phygitale, l'AB testing et l'optimisation de sites et du produit marketing et surtout de cracker la boîte noire des émotions.

Je voulais aussi vous faire une confidence : on a filmé le jury pour savoir si j'ai provoqué de l'attention, de l'émotion et surtout de l'adhésion afin de prédire la note qu'ils vont nous donner. Réponse tout à l'heure. Je vous remercie. »

ANNEXE 3

Résultats de l'analyse terrain

Les entretiens qualitatifs semi-directifs ont fait ressortir plusieurs thèmes :

Pédagogie autour de l'émotion et revendication d'un savoir neuroscientifique

- Choix du centre *The Brain Place to Be*
- Impact des émotions au quotidien dans nos réactions
- Nécessité de capter « l'attention », nouveau nerf de guerre

Choix de communication de Datakalab

- Un lancement stratégique
- Datakalab, un nom chargé d'imaginaires
- Des « métriques » qui se veulent miroirs des besoins de l'annonceur vis-à-vis des consommateurs
- Une promesse de « mesure émotionnelle » altérée mais maintenue malgré le repositionnement
- Datakalab comme une révolution qui lance des tendances

Argumentaire commercial

- Les données comme ancres mémoriels

Appréhender la charge symbolique des émotions

- Universalité des émotions de base de Ekman
- Limite des analyses déclaratives : biais, difficulté de collecte des données
- Croisement des données émotionnelles et déclaratives, avantages et biais de ces deux méthodes de collectes

Prise de distance des professionnels face à la « mesure émotionnelle »

- Manque d'éducation en France autour du « neuromarketing »
- Discours divisés des professionnels autour de la notion « d'affective computing », miroir d'une logique capitaliste

- Logiques différentes recherche/vente
- Croyance ou non des différents professionnels à la « mesure des émotions » par la machine

Réglementaire et éthique

- Respecter le nouveau règlement sur la protection des données
- Consentement du client et respect des libertés personnelles
- Connotations négatives du terme « neuromarketing »

Cadre technique et technologique

- Choix du « facial coding » pour « mesurer l'émotion » et adaptation de la méthode Ekman
- Mise en place des études émotionnelles
- Transformation des données brutes en données intelligibles
- Collecte et le traitement des données émotionnelles
- Différentes technologies de captation et de mesure des émotions pour différents contextes et usages
- Nouveaux algorithmes propriétaires et innovants de Datakalab

Difficultés de construction d'algorithmes de « détection d'émotions »

- Possibilités de biais des algorithmes
- Difficulté de trouver des données spécifiques pour entraîner des algorithmes à la détection d'émotions
- Minimiser « l'erreur » dans la détection de pics d'émotions et « mesurer » la performance des algorithmes
- Problématiques de construction d'algorithmes spécialisés dans l'analyse faciale

Pilotage stratégique de Datakalab

- Principes business de mise à l'échelle des technologies sur un panel
- Changements dans la mesure émotionnelle suite au repositionnement et à la suppression de l'outil « bracelet connecté »
- Bascule de l'expertise « conseil » à « technologique »
- Stratégie de Datakalab et levée de fonds

- Un changement de cible anticipé
- Stratégies concurrentes
- Possibilité d'ouvrir la plateforme de collecte de vidéos des panélistes à tous les utilisateurs
- Datakalab projette de s'internationaliser

Application des produits Datakalab

- Datakalab crée ses propres normes
- La start-up creuse les autres applications du « facial coding »

Limites de l'analyse

Établir une vision des différents discours portés sur la donnée émotionnelle s'avère complexe, déjà dans la mesure où ces discours évoluent constamment. Les acteurs impactés par cette donnée sont nombreux : les employés de Datakalab eux-mêmes, commerciaux et ingénieurs. L'équipe rencontre d'autres acteurs extérieurs à Datakalab, les annonceurs, instituts d'études, médias, participants aux études, prospects. Il faudrait aussi les interroger sur l'appropriation qu'ils font de cette donnée. Cette étude pourrait être confrontée à une étude quantitative pour savoir quelles entreprises ont recours à la donnée émotionnelle. Elle devrait être anonyme pour ne pas être confrontée à la nature sensible de ce sujet et à la difficulté d'obtenir des témoignages de la part des annonceurs. Une observation ethnographique de l'utilisation de ces données dans des sociétés internationales pourrait encore être conduite.

ANNEXE 4

Entretien avec Frank Tapiro, Co-fondateur et Chief Emotion Officer, 26/06/2019, Paris 17^{ème}

Frank est publicitaire, ex-conseiller médiatique de présidents français mais avant tout un producteur d'émotions. Il a créé plusieurs entreprises dont Hémisphère Droit et On Broadway. Frank a une connaissance intime des plateformes de marque et a écrit le livre Pourquoi la vache qui rit ne pleure jamais ? paru chez Albin Michel sur le génome des marques.

Dans quelle mesure garde-t-on la même promesse de mesure émotionnelle alors que l'outil est différent ?

On n'est pas du tout sur la même promesse. J'ai commencé à travailler avec les « bracelets », ensuite, j'ai muté, grâce à Xavier sur le *facial coding*. La promesse n'est pas du tout la même, déjà d'un point de vue pratique. On a un outil qui est *scalable* d'un côté, le *facial coding*, puisqu'on peut en faire sur n'importe quelle caméra, c'est automatique, simple et rapide. D'un autre côté, on a le bracelet, qui est efficace, mais qui demande beaucoup d'opérations. Le bracelet n'apparaît pas tout seul dans la vie des consommateurs, il faut prendre des hôtesse qui doivent être briefées. Elles doivent aller voir les gens, leur mettre le bracelet. Il faut ensuite suivre la personne, taguer ce qu'elle est en train de faire. La charge opérationnelle est donc très lourde. C'est un outil fabuleux, qui arrive à donner des pics émotionnels, donc on sait s'il y a eu un engagement émotionnel ou pas. Le fait que les ressentis soient négatifs ou positifs, ce n'est pas si important, et on peut le savoir grâce aux questions qu'on pose ensuite aux participants. Au moins on sait qu'il s'est passé quelque chose. On cherche aussi à savoir quand il ne s'est rien passé. Quand il ne se passe rien, cela ne veut pas dire que c'est bien, cela veut juste dire que c'est neutre, et donc que les gens ne font pas attention à ce qu'il se passe. On aime que les gens soient engagés, qu'ils aient des pics émotionnels. Il faut ensuite traiter la donnée. Cela a très bien fonctionné, mais Datakalab était à l'époque principalement portée sur le conseil. On vendait ce produit cher, mais la marge était moins forte qu'aujourd'hui avec le *facial coding*. Il peut y avoir des sociétés qui ne font que ça, et au début, c'était le cas de Datakalab. On a commencé avec des grandes marques comme SNCF, on est passé au 20h plusieurs fois, et on a pu optimiser leur gare pilote SNCF d'Aix-en-Provence.

On s'est demandé si on devait poursuivre sur le conseil. Xavier, qui était à l'époque notre Chief Innovation Officer ne voulait pas être une énième société de conseil. Pourtant Anne-Marie venait de ce monde-là. Tout le monde parle du *facial coding* depuis assez longtemps. Mais comme on n'avait pas encore l'outil propriétaire, notre algorithme actuel, on prenait des briques qui venaient d'un peu partout. On a pris une décision : tout ce qu'on faisait avec le bracelet et la partie conseil allait financer le lancement de la technologie, ce qui nous a permis de ne pas faire de levée de fonds. Je ne voulais pas faire de start-up mais une *smart-up*. Une *smart-up*, c'est une boîte qui lève des clients au lieu de lever de l'argent. Ça nous a réussi puisqu'on a fait 650 000 euros en un an avec nos bracelets. Xavier est venu avec son frère, on a engagé des chercheurs pour développer l'algorithme. On savait que le jour où on aurait cet algorithme là, on basculerait sur un mode complètement « tech », complètement *scalable*. La bonne surprise a été de voir que ça a été plus rapide que prévu, que l'algorithme est monstrueux, et qu'on a récupéré les deux meilleurs chercheurs au monde là-dessus.

Tout était prévu.

Oui, c'était totalement stratégique. On l'avait anticipé depuis le démarrage, ce n'est pas du hasard. La seule coïncidence a été la rencontre avec Xavier. J'avais monté Datakalab depuis quelques mois, et il est venu me voir pour me présenter Emotient aux États-Unis. Cette société que je ne connaissais pas, faisait la même chose que moi mais sur le *facial coding*. Je lui ai ensuite présenté Datakalab. C'est comme ça qu'on est devenus très proches. J'ai immédiatement apprécié Xavier pour son esprit brillant et sa vision. Il me parlait d'une société que j'étais en train de monter.

Au bout d'un an on s'est associés, Xavier, Anne-Marie et moi. J'ai laissé mes anciens associés et on a relancé Datakalab en 2017 sous un nouveau format : du bracelet avec du conseil pour financer le lancement de l'algorithme. *In fine*, grâce à l'algorithme, le processus est rapide et fluide, ce qui permet à l'entreprise d'être beaucoup plus rentable, parce que les marges n'ont rien à voir avec ce que fait un bracelet. Ces nouvelles données qu'on récolte avec l'algorithme ne sont pas comparables avec celles du bracelet. Le bracelet nous dit uniquement les pics d'émotions, alors que notre algorithme nous dit tout, il décrypte le visage, il dit si on est un homme ou une femme. Il ne fait pas de la reconnaissance faciale, il détecte le clignement des yeux pour comprendre la charge cognitive du regard, l'orientation de la tête pour savoir si la personne est attentive ou non.

Le nerf de la guerre aujourd'hui c'est l'attention. S'il n'y a pas d'attention, tout le reste ne sert à rien. On a tendance à oublier que c'est l'attention qui fait tout, c'est le point de départ. S'il y a de l'attention, on peut mesurer l'émotion. Grâce aux « action units », on peut détecter ces expressions. L'algorithme aujourd'hui qui a été fait par Arnaud et Kevin est au-delà du réel, et surtout, c'est le nôtre. La grande prouesse de Xavier avec l'équipe, c'est qu'en deux ans, on a le meilleur algorithme au monde, devant Facebook et Facetime, et c'est prouvé avec des chiffres. L'algorithme est totalement en phase avec ce qu'on attend aujourd'hui : capter l'attention, capter l'émotion, capter l'adhésion. C'est une fierté absolue, on est au début de l'histoire.

J'avais lu tous les articles parus sur Datakalab quand j'étais encore au CELSA. Pourtant, tant qu'on est extérieur à Datakalab, on ne se rend pas compte de toute cette stratégie.

Tout est vrai, mais lorsqu'on raconte l'histoire, les gens sont sidérés. J'ai rencontré Xavier grâce à Jacques Séguéla, mais Xavier venait me voir pour Hémisphère Droit, pour tester mes pubs. Il y avait le logo de Datakalab sur la porte. Je lui ai demandé s'il n'avait pas vu le logo. Il m'a répondu que non. C'est moi qui lui ai présenté Datakalab, lui me parlait d'Emotient. Je lui ai dit : « un jour ou l'autre on travaillera ensemble, c'est sûr ». Le nouveau Datakalab a été créé grâce à Xavier et Lucas. J'ai eu la vision de Datakalab, créer les *feel data*© et avoir une nouvelle génération de données émotionnelles. Par contre au niveau des outils, c'est Xavier qui a eu l'idée en disant qu'il fallait arrêter les bracelets et de se concentrer sur l'algorithme.

Je sais que le nom vient du film *Bienvenue à Gattaca*, mais pourquoi ce choix ?

C'est une longue histoire sur le rapport aux sciences et à la médecine. Je voulais être chirurgien, mais j'ai oublié de déposer mon dossier en faculté de médecine par un acte manqué total. J'ai ensuite voulu me rabattre sur une fac de pharmacie, mais cinq minutes avant de rentrer dans la faculté, j'ai fait un malaise vagal. J'ai décidé que c'était un signe, et que ça n'était pas pour moi. J'ai fait une école de communication, une très bonne formation pour apprendre une autre de mes passions, le cinéma. J'ai créé la première agence de pub, j'ai travaillé chez RSCG avec Jacques Séguéla, j'ai aussi travaillé avec Nicolas Sarkozy et François Mitterrand. J'ai eu beaucoup de chance, je considère ça comme du fun qui a duré longtemps. Je devais faire un an de stage dans la pub, j'en ai fait 30. J'ai créé mon agence, Hémisphère droit, je m'intéressais déjà à la préférence cérébrale : c'est la partie du cerveau qui gère les émotions. Je considérais

déjà la pub comme de l'émotion. L'émotion, c'est de l'engagement, et l'engagement c'est la clé du succès.

J'ai souhaité inventer un outil pour décrypter les marques. Je ne me suis pas demandé si les publicités sollicitaient l'hémisphère droit ou l'hémisphère gauche. La plupart des pubs sont sur l'hémisphère gauche, et les publicitaires essayent de les faire basculer sur l'hémisphère droit. J'ai inventé un concept en rapport avec la biologie et la génétique « l'ADN des marques », qui était dans l'air du temps, et plus particulièrement le « génome des marques ». Tout comme il y a eu la cartographie du génome humain, j'ai inventé la cartographie du « génome des marques ». J'ai travaillé pendant 13 ans en recherche fondamentale et j'ai écrit un livre là-dessus. J'ai travaillé avec toutes les plus grandes marques, de Coca-Cola à Yoplait, et je me suis rendu compte que cet outil était extraordinaire pour réveiller les valeurs fondamentales sur lesquelles les marques étaient assises sans même les utiliser. J'ai pu voir que comme dans notre ADN, les marques ont des « Authentiques Différences Naturelles » elles sont la somme de ces valeurs fondamentales partagées avec tous les publics. Il y a de la génétique et il y a de l'émotion, puisque l'émotion, c'est de la transmission.

Déjà à l'époque, on ne comprenait pas pourquoi je parlais de génétique en pub. Je parlais des marques, sans parler de pub. La pub, c'est le langage. La marque, c'est un être vivant qui est animé par des valeurs, c'est émotionnel. J'ai été jusqu'à demander à Daniel Cohen, le père de la carte du génome humain, d'écrire la préface de mon livre. Il a complètement validé mon approche. Une marque, c'est une personne avec des fondamentaux, à la fois innés et acquis, avec des valeurs et des mutations en fonction du contexte. À l'époque, j'expliquais que les émotions se transmettaient génériquement, à chaque fois que je rencontrais des scientifiques, personne ne me contredisait sans pour autant me donner raison. Je travaille avec des médecins, avec des chercheurs, des scientifiques, et on parle de l'émotion, la chose la plus humaine et qui définit le plus l'humain. C'est ce qui nous différenciera toujours des machines.

J'ai décidé d'arrêter la pub puisque je considère que dans une agence on n'arrive plus à faire passer ses idées. On a qu'à regarder ces quatre dernières années les catastrophes créatives qu'il y a sur les murs et à la télévision. J'ai voulu inventer un outil scientifique pour prouver que les créatifs avaient raison. Ils ne sont jamais pris au sérieux, il fallait une donnée scientifique qui étaye, valide tout ce que pensent les créatifs et les artistes. Ce qu'on ne peut pas expliquer

autrement que par notre talent ou notre expression artistique, on ne peut pas l'expliquer parce que ça touche les valeurs, les émotions, l'ADN.

Bienvenue à Gattaca est mon film culte, il parle de génétique, de transmission, d'émotions, de transhumanisme. C'est le film qui résume toutes les passions de ma vie et je souhaitais lui rendre hommage. Datakalab, c'est un laboratoire de recherche qui agite le bocal pour prouver de manière scientifique que les créatifs ont raison : quand on crée de la différence, on crée de la performance. Pendant un an, j'ai travaillé sur ce projet avec des gens qui travaillaient avec de la donnée rationnelle. J'ai été le premier à inventer le *data capital*, j'avais écrit un papier pour expliquer que plutôt que de se battre contre l'utilisation des données, chacun devrait se les approprier, et les louer, les vendre pour créer un *data capital*. J'envoie le nom de Gattaca à mon avocat qui me dit que le nom est déjà déposé, et je pense au « data capital » et je me dis que Dataca serait un bon nom, d'autant plus qu'on pouvait dire : *Data Creative Agency*, la société qui prouve que les *data* aident à la création. Je dépose le nom, mais un *Dataca.com* était déjà déposé aux États-Unis. Je me suis alors dit que comme c'était la mode des laboratoires, on allait s'appeler Datakalab. On m'a dit que c'était compliqué, mais j'ai répondu que Datakalab sonnait un peu comme « abracadabra », cela renvoie à la magie. En plus, personne n'arrive à le dire, ce qui me permet de rebondir dessus, d'en faire une blague. Aujourd'hui, en télé, en conférence, tout le monde en parle. Qui ne connaît pas Datakalab aujourd'hui ? C'est une belle histoire vraie.

Est-ce que tu connais le terme « d'affective computing », et si oui, pourquoi est-ce qu'on ne l'utilise pas ?

Tu peux m'en dire plus ?

C'est une branche des sciences informatiques pour « mesurer » les données corporelles, le pouls, les expressions du visage pour connaître et interpréter les émotions.

Oui, j'aime bien ce terme, c'est vrai qu'on ne l'utilise pas. C'est une très bonne question, une bonne suggestion, on va en parler avec Xavier, mais *affective computing*, c'est un très bon terme. Il faut qu'on voie qui l'utilise, ce n'est pas loin de « feel data© ». *Affective computing*, c'est très beau, j'aime beaucoup, même si Affectiva est l'un de nos concurrents. On va en parler avec les chercheurs.

As-tu inventé et déposé le terme de « feel data© » avant que le monde entier ne l'utilise ?

C'est vrai ? Je ne le savais même pas que ce terme avait autant de succès. J'en suis fier. On l'a lancé en 2016 au salon Big Data. J'ai expliqué que dans un monde de big data, il était important que la « feel data© » arrive. La donnée émotionnelle, c'est aussi plus que l'émotion, c'est le ressenti, c'est aussi ce qui m'intéresse.

Oui, et « feel data© » peut aussi faire penser à un ordre : « ressentez la data ».

Décidément, très très bonne remarque, je n'y avais jamais pensé.

Si on reste sur le *naming*, j'ai entendu que vous vouliez renommer la « métrique » actuelle « d'émergence » en « impact ».

Oui, parce que l'émergence est encore un mot qui sonne « conseil ». L'impact, tout le monde comprend, parce que dans tous les domaines : la communication, le cinéma, s'il n'y a pas d'impact, il n'y a pas d'émotion. Je pense qu'il faudrait que nos métriques soient : l'attention, l'impact, l'adhésion, et le rejet. Je pense que c'est la dialectique de tout contenu. D'abord, tu crées l'attention. Ensuite, il y a l'impact émotionnel, et après, il faut savoir ce que ça laisse comme ressenti, positif ou négatif : adhésion, rejet, ou neutralité, mais quand il y a de la neutralité, c'est qu'il y a eu peu d'émotions, ou des émotions qui ont déclenché tellement de ressentis négatifs que ça provoque un rejet violent. Je dirais que même ça, au moins ça crée quelque chose. Je préfère une campagne détestée, conspuée, censurée à une campagne sans saveur, sans odeur. Au moins, on a brisé la barrière de l'indifférence. Aujourd'hui, on ne peut plus être en deçà de la barrière de l'indifférence, c'est le nerf de la guerre dans tout. Il faut créer de la différence. C'est de l'identité. Aujourd'hui, on en manque, il y a une banalité absolue de tout, des start-up, des *pitchs*, des présentations, des *buzzwords* stupides comme les « écosystèmes » pour « créer de l'activation des communautés ». Tout ça, c'est pour masquer le manque de talent et de créativité. Pendant que j'entends ce genre de discours, j'attends qu'on donne des idées. C'est du « bullshit » total.

Peux-tu me parler du terme de « brain tech » ?

C'est une catégorie. Il y a beaucoup de sociétés de bio-tech. « Brain tech », c'est le rapport entre la « tech » et le cerveau. Comment est-ce qu'on analyse le cerveau grâce à la « tech », l'intelligence artificielle, et la « computer vision » pour notre part. C'est la façon dont l'intelligence artificielle arrive à interpréter le cerveau. Le cerveau n'arrive lui-même pas à le faire, puisque 90 % des stimuli qui nous impactent sont traités par l'inconscient. L'intelligence artificielle sert à faire remonter le fond à la surface. Grâce à l'intelligence artificielle, qu'on accuse à tort, on sait ce qui se passe dans l'inconscient des gens. Ce n'est pas pour les espionner, c'est pour révéler des choses positives. Les gens se plaignent tout le temps, mais n'osent pas toujours parler, et n'ont pas toujours conscience de ce qui ne va pas. On pourra sans doute faire un jour de la psychothérapie grâce à l'intelligence artificielle. J'y crois sincèrement, ça va nous aider. Peut-être que grâce à l'intelligence artificielle on va se rendre compte de nouvelles choses. Il faut savoir écouter les gens, et surtout ce qu'ils ne disent pas. La personne pense qu'elle ne nous dit rien, mais en fait elle nous dit tout : sur son visage, sur son expression, sur son attitude. Je le dis souvent aux présentateurs télé, qui réagissent comme s'ils ne voulaient pas interagir avec leurs invités. On est dirigés par notre inconscient. Grâce à l'intelligence artificielle, on va aussi pouvoir libérer les choses qui nous angoissent, qui nous traumatisent.

Pourquoi le passage de l'ancien « menu » : *Data diagnostic, feel data, data telling*, à la *baseline, Cracking the emotion's blackbox* ?

Au début avec Datakalab, on disait : « la donnée sera émotionnelle ou ne sera pas ». Cela montre que plus on va avancer, plus on va aller vers l'homme et vers l'intelligence émotionnelle, et non pas vers la machine contrairement à ce qu'on dit. La machine va nous aider à être plus humain. Moi c'est mon but dans la vie. J'aimerais que le 90 % des stimuli traités par l'inconscient se transforme en 80 %. J'aimerais qu'on se rende compte qu'on a des transmetteurs émotionnels par nos gènes. Maintenant c'est prouvé, les émotions se transmettent génétiquement. Voilà pourquoi de générations en générations, on a des transmissions d'angoisses non résolues, on a des bagages génétiques qui ne sont pas à nous et qui nous viennent de nous aïeux, et on a leurs réactions alors qu'on ne les a jamais connus. Les émotions sont partout, leurs transmetteurs, ce sont les gènes, et grâce à l'intelligence artificielle, on va faire marcher différemment son cerveau et on va être un humain augmenté grâce à la technologie. C'est ce que j'avais écrit en 2016 au moment de lancer Datakalab. C'est Rabelais :

« Science sans conscience n'est que ruine de l'âme ». Ça dit tout : si on a la conscience, la science ne fera que du bien. Si on n'a pas de conscience, pas de morale, pas d'éthique, la science va nous tuer. La science a été inventée par l'homme, et c'est un outil humain qui utilise le cerveau. Aujourd'hui, le cerveau veut utiliser la science pour faire encore mieux fonctionner le cerveau, c'est tout le but de Datakalab. C'est mon objectif. Je travaille maintenant pour Datakalab avec l'Institut Rafaël, on va développer des outils pour améliorer l'interaction entre patient et traitant pour faire en sorte d'envoyer un peu plus d'énergie positive à des gens en rémission de cancer pour éviter qu'ils ne rechutent. Là je reviens à ma petite enfance, et je serai l'homme le plus heureux au monde si on y arrive. Je me dirai que je suis passé par un autre chemin mais j'ai réussi mon objectif, parce que j'aime aider les gens, les guérir, les soigner avec ce que je peux, mon mental, et mon énergie de vie, qui est un don. J'ai envie d'en faire profiter tout le monde. Le jour où on gagnera beaucoup d'argent, je ne ferai que de l'humanitaire et de la santé. Quand une personne atteinte du cancer va chez le médecin, ce n'est pas uniquement pour parler des analyses, c'est aussi pour parler de la vie, parce que sans ça il n'y a pas d'espoir, il va rechuter. Si quelqu'un passe un entretien d'embauche, il peut être jugé sur son stress alors qu'il serait un excellent élément dans l'entreprise. J'aimerais qu'on puisse réveiller le meilleur en l'humain. D'abord, on montre les compétences de Datakalab sur l'analyse, le stockage des émotions, et tout ce plan humain, ça sera mon Datakalab d'après.

ANNEXE 5

Entretien avec Soukeina Hassanaly, Ex Key Account Director chez Datakalab, 05/04/2019, Paris 17^{ème}

Avec 10 ans d'expérience dans le « retail », principalement au sein du Groupe Galeries Lafayette, Soukeina Hassanaly a travaillé dans les départements Marketing, Fidélisation et Connaissance Client. Elle a rejoint Datakalab dès sa création en 2017. Elle était responsable commerciale et gérait l'ensemble des études émotionnelles. Aujourd'hui, Soukeina a rejoint le cabinet de conseil Diametrix (Diamart Group) sur l'utilisation des différentes data (structurées ou non) pour une meilleure performance omnicanale.

Avant d'arriver chez Datakalab, tu travaillais dans le CRM aux Galeries Lafayette. Pourquoi est-ce que cette start-up t'a attirée ?

D'abord parce que c'est une start-up, et que je n'avais jamais travaillé dans cet environnement. Travailler dans les émotions, c'est dans l'air du temps. Donc c'était important pour moi de pouvoir comprendre cela pour avoir une plus grande valeur ajoutée. Je voulais savoir comment on exploitait toute la *data*, et ce qu'on pouvait en tirer si on rajoutait cette dernière brique émotionnelle. C'est elle qui permet d'avoir une vision 360 du client. Celui-ci ne peut pas être seulement considéré comme un panier moyen ou un chiffre d'affaires.

Est-ce qu'on pourrait dire que Datakalab a fait un « me too » de toutes les start-up qui naissent dans le neuromarketing, notamment aux États-Unis ?

Oui et non. Xavier a travaillé chez Emotient, donc oui, sauf qu'il est allé plus loin. Il s'est demandé quelles étaient les limites de Emotient. Quelle était la technologie utilisée, et comment améliorer les algorithmes ? C'est pour cette raison qu'on a encore recruté deux nouveaux chercheurs.

Et alors, quelles étaient les limites de Emotient ?

Xavier l'expliquera mieux que moi mais ce sont toutes les limites par rapport aux annonceurs, une brique marketing qui n'allait pas jusqu'au bout dans leur accompagnement.

Datakalab est une start-up dans la « brain tech ». Ce terme de « brain tech » peut paraître surprenant. Comment vous est-il venu ?

Lorsque Anne-Marie est revenue du CES²⁵⁶ et de la NRF²⁵⁷ en 2018 nous avons décidé d'utiliser ce terme, très courant aux États-Unis. En France il surprenait l'année dernière, maintenant les gens comprennent parce qu'ils s'intéressent davantage à tout ce qui touche aux neurosciences.

Pourquoi le choix d'être dans *The Brain Place To Be* ?

Toutes les entreprises qui y sont installées travaillent dans les neurosciences. À la base, l'objectif était aussi de pouvoir y faire des rencontres, du *networking*. Pour le moment nous avons réalisé un événement. Les trois entreprises qui y travaillent invitent leurs clients, et on parle tous de sujets liés aux neurosciences. On n'a pas eu le temps de s'y atteler, mais l'objectif est de devenir un lieu incontournable. On a fait le lancement de *The Brain Place To Be*, mais on ne l'a pas encore refait.

Datakalab surfe sur plusieurs tendances comme l'IA, les neurosciences, le neuromarketing, on parle même de neuro-consommateurs. Datakalab s'appuie-t-elle sur d'autres tendances ?

Nous ne surfons pas sur ces tendances, nous les avons anticipées. On parlait déjà de neurosciences et d'émotions avant que le marketing ne s'oriente là-dedans. On a été créés avant même qu'on en parle, donc on ne surfe pas sur la tendance. Je pense plutôt qu'on en a été précurseurs.

²⁵⁶ Salon à Las Vegas sur les nouvelles technologies

²⁵⁷ Premier salon à l'échelle mondiale dans le retail qui a lieu chaque année à New York

Pour améliorer une pub, pourquoi les « métriques » actuelles « d'attention, d'émergence et d'adhésion » sont plus parlantes à l'annonceur que de parler directement des émotions : joie, tristesse, colère par exemple ?

L'objectif des annonceurs aujourd'hui, c'est de faire la différence. C'est pour ça qu'on a cette notion d'émergence. Aujourd'hui, on est bombardés d'informations, donc il faut capter l'attention. Le consommateur est volage : je peux me promener dans la rue, je ne regarde même pas les abribus, les 4X3, parce je suis sur mon téléphone en train de discuter. Il faut créer une publicité qui ait suffisamment d'impact pour déjà attirer mon œil. Si oui, ça veut dire que j'ai rendu la personne attentive. Ensuite, on se demande si on a réussi à lui faire ressentir une émotion. On sait que lorsqu'on ressent une émotion, quelle qu'elle soit, cela stimule l'hippocampe qui est responsable de l'encodage de la mémoire. Donc il faut savoir que si on a réussi à faire ressentir quelque chose à la personne, elle devrait s'en souvenir. C'est pour ça qu'on utilise ces métriques, c'est très important pour l'annonceur. Aujourd'hui, tout le monde parle de ça. Comment capter l'attention ? C'est l'enjeu de tous les annonceurs, et de toutes les conférences.

Aurais-tu un exemple de « data telling » où la donnée déclarative dit une chose et la donnée émotionnelle une autre, et comment fais-tu dans ce cas ?

Je n'ai jamais de données qui soient contradictoires. Au contraire, c'est la donnée émotionnelle qui nous donne une indication de ce que ressent la personne puisqu'elle est prise directement dans la partie inconsciente, et le déclaratif vient enrichir cette donnée émotionnelle. Donc je ne me suis jamais retrouvée dans une situation où les données étaient contradictoires. Par contre, la donnée déclarative vient expliquer la donnée émotionnelle : « Pourquoi avez-vous eu ce pic émotionnel dans les toilettes ? » ça, c'est du factuel. La donnée déclarative est agrégée et nous donne l'explication. En l'occurrence, il y avait un pic émotionnel parce que c'était payant.

Dans la perception commune, c'est plutôt l'émotionnel qui vient enrichir le déclaratif. Là tu me dis que c'est l'inverse. Du coup comment convainrais-tu l'annonceur ?

Pourquoi on s'occupe des émotions ? Parce que c'est la partie inconsciente, alors que le déclaratif n'est que sur la partie consciente. Il y a 90 % des choses qui nous impactent dans une

journée qui ne sont traitées que par l'inconscient. Donc le déclaratif ne donne qu'une vision partielle de ce que pense le consommateur.

Au début Anne-Marie voulait que Datakalab fasse du conseil, mais aujourd'hui on veut se concentrer sur la partie « technologique ». J'ai entendu que l'on devait se démarquer des instituts d'études et qu'on voulait passer en « mode SaaS²⁵⁸ » avec un focus sur le « facial coding ». Est-ce qu'il y a d'autres raisons en dehors du fait que ce nouveau modèle est plus « scalable » ?

Si Datakalab est devenue une plateforme SaaS, c'est notamment parce qu'on ne veut plus s'occuper du bracelet connecté, dont le protocole est assez lourd : il faut être sur place, avoir des hôtesse, ce qui coûte cher. Or la partie d'analyse des expressions faciales est beaucoup plus facile à mettre en place parce qu'on a développé tout l'algorithme et la plateforme. C'est grâce à cela qu'on va être *scalable* rapidement. En ce qui concerne les instituts d'études, ils ont déjà une expertise pour la traduction et l'interprétation des *data*. Les instituts d'études produisent des études traditionnelles avec du déclaratif. Ils n'ont pas cette brique émotionnelle.

Sans le bracelet, on peut toujours capter les émotions par « facial coding » mais plus l'intensité émotionnelle. Ce sont d'autres « métriques » et on aura peut-être une vision moins globale de l'émotion. L'intensité émotionnelle ce n'est pas la même chose que les émotions positives par exemple.

Oui, mais on a quand-même une notion d'intensité quand on fait du *facial coding*. On a bien émotion par émotion une intensité qui est différente. La valeur ajoutée du bracelet se révèle lors du parcours client, quand on se déplace, parce qu'on ne peut pas mettre une caméra face à chaque personne, dans la rue, ou dans une vitrine. Donc c'est pour ça qu'on avait développé cet outil là, pour pouvoir traquer l'émotion lors d'un parcours. Le bracelet nous donne une indication sur l'intensité émotionnelle : le degré, le pic, l'intensité du pic. Mais cela ne nous dit pas si c'était de la joie, de la surprise, de la colère. Ça c'est la partie du *facial coding* qui nous donne une lecture directe de l'émotion parce qu'on a les micro-expressions. C'est pour ça que le protocole du bracelet est plus lourd, parce qu'on était obligés de faire une phase de *debriefing*, de déclaratif, pour avoir la bonne interprétation de l'émotion. Sinon en voyant des pics

²⁵⁸ Logiciel en tant que service, modèle de distribution de logiciel en ligne.

d'intensité émotionnelle dans les toilettes, on pouvait se dire : « c'est parce que ça sent mauvais, ou que c'est sale ». Mais ce n'est pas parce qu'on ne fera plus le bracelet qu'on n'aura plus une vision globale de l'émotion. L'émotion se traduit par différents signes extérieurs, et nous sommes focalisés sur les signes présents sur le visage.

Pour les différentes applications de ce nouveau type de données, on parle de la santé et de l'éducation. Est-ce que tu verrais d'autres domaines d'applications ?

Je trouve que le plus important est de pouvoir s'occuper des enjeux sociétaux. Donc tu l'as dit, il y a l'éducation, et il y a un gros travail à faire. Il y a la partie « santé » qui est aussi importante puisqu'aujourd'hui on a la *MedTech*. On a fait un partenariat avec l'Institut Rafaël, la maison de l'après-cancer. La question est de savoir comment on va accompagner les patients dans leur guérison. Parce qu'il y a la partie guérison médicale, mais il y a la partie aussi guérison dans la tête. Ensuite parmi les autres enjeux qu'on a essayé de développer, il y a la partie « sport ». On sait que le sport est un créateur de lien social et d'émotions. On a travaillé avec AccorHotels Arena pour pouvoir expliquer comment se différencier pour pouvoir faire venir des clients et driver de l'émotion.

Il y avait aussi la partie politique avec le suffrage émotionnel. Donc peut-être pour les prochaines présidentielles ?

Oui peut-être !

ANNEXE 6

Entretien avec Xavier Fischer, Co-fondateur et Chief Executive Officer de Datakalab, 26/07/2019, Paris 17^{ème}

Diplômé de l'école Centrale, Xavier a été en charge du développement chez Emotient, société spécialisée dans l'intelligence artificielle et l'analyse des expressions faciales rachetée par Apple en 2016. Il y développait la technologie de Apple de reconnaissance faciale.

Peux-tu me parler de ce que tu faisais chez Emotient, et en quoi tu as pu aller plus loin chez Datakalab ?

Chez Emotient, j'étais responsable du produit, donc je gérais les équipes d'ingénieurs pour faire converger tout ce qu'ils faisaient, notamment sur le développement de nouvelles *features*²⁵⁹. La différence avec Datakalab, c'est que maintenant j'essaie de créer quelque chose en fonction de ma vision. Emotient, c'était une entreprise qui développait vraiment que de la R&D. Chez Datakalab, nous sommes beaucoup plus proches du marché, dans le sens où on veut vraiment créer des business dès le démarrage pour ne pas dépenser trop d'argent.

En quoi est-ce que notre stratégie te semble meilleure que celle de nos concurrents ?

Le marché de l'émotion est très jeune. Les boîtes comme Affectiva lèvent 50 millions de dollars et elles embauchent 50 personnes, mais elles ne sont pas rentables, et ne génèrent pas d'argent. Tandis que Datakalab recrute les meilleurs, et a besoin de très peu d'argent : on est quasiment rentables aujourd'hui, ce qui nous permet d'avoir une durée de vie plus longue. Quand tu lèves 50 millions de dollars, tu es soit obligé de te faire racheter, soit de faire 10-15 millions de dollars annuel parce que les investisseurs veulent toucher leur argent, soit sous forme de dividende, soit sous forme de ré-achat.

²⁵⁹ « Fonctions » dans le langage ingénieur.

Penses-tu qu'il y ait pourtant des difficultés pour Datakalab, freins de la part de nos nouvelles cibles ?

Je pense que cette nouvelle cible constitue un bon marché. Tout le monde regarde ce qui s'y fait. Toute action prend du temps à se mettre en place mais quand ça s'enclenche, tout est lancé rapidement. Nous cherchons à enclencher ce processus. Grâce au partenariat avec Médiamétrie, on commence à bien avancer. Ensuite il y a d'autres leviers sur lesquels il faut commencer à réfléchir, dans le sens où plus tu collectes de data, mieux c'est. Les marchés où tu utilises les algorithmes tous les jours, toutes les minutes, sont encore plus intéressants pour nous.

En termes de stratégie, est-ce que tu verrais des faiblesses, voire des menaces pour Datakalab ?

Je pense que ce choix de ne pas lever trop d'argent, de prendre notre temps, peut être une faiblesse dans le sens où si quelqu'un veut nous copier, il peut aller vite en levant de l'argent. D'un autre côté, il faut qu'on soit réactif. Mon but consiste à être agile et donc à voir ce qui se passe, quelles sont les menaces. Il faudra peut-être qu'on accélère. Je pense qu'il y a plein de menaces, mais que le but est d'y réagir à chaque fois qu'elles arriveront. Beaucoup d'acteurs comme Microsoft pourraient se positionner sur ce qu'on fait, mais notre but et notre métier c'est d'analyser ce qui se passe et de voir comment y réagir.

Pourquoi faire le choix de ne pas communiquer sur le neuromarketing ?

C'est un terme qui est proche de la manipulation, et ce n'est pas ce qu'on fait, dans le sens où on détecte la position du visage, des déformations musculaires. Aller dans le sens du mot « neuro » est très fort. On ne dit pas que c'est du neuromarketing quand on aime une photo sur Facebook, alors qu'en fait, on se demande comme à Datakalab si la personne aime ou non. Il ne faut pas employer des termes qui font peur. Notre objectif, c'est de dédramatiser.

À l'époque de la « première version » de Datakalab, celle du conseil sous la direction de Anne-Marie, vous pensiez déjà à la seconde version, plus axée sur la technologie. Est-ce que vous pensez déjà à la « troisième version » de Datakalab ?

On pense à la troisième version de Datakalab. Le but, c'est d'accélérer de plus en plus, de développer de plus en plus de *features* technologiques, et d'utiliser le plus de fois et de minutes par jour nos algorithmes, dans la mesure où l'algorithme est à nous, il est financé, et plus on l'utilise mieux c'est. Le but, c'est vraiment de l'amortir.

ANNEXE 7

Entretien avec Lucas Fischer, Co-fondateur et Chief Data Officer de Datakalab, 08/07/19, Paris 17^{ème}

Diplômé à la fois de l'école Centrale Supélec et de Berkeley, Lucas a travaillé sur le comportement humain et sur la reconnaissance des émotions. C'est un entrepreneur passionné qui a de grandes connaissances en data mais aussi en pilotage de marque.

Tu es diplômé de Berkeley. Est-ce que ces études t'aident au quotidien ?

Oui, je ne serais pas aussi à l'aise dans le code sans ces études. J'y ai surtout appris l'entrepreneuriat. On fait partie de l'incubateur de Berkeley. J'ai passé six mois à apprendre l'entrepreneuriat à la Silicon Valley, ce qui m'a donné un *mindset* particulier, qui est différent de celui en France sur l'entrepreneuriat, sur les *business model*, sur la relation avec les clients, sur la manière de réfléchir aux questions du jour. Ça permet de penser différemment.

Frank m'a expliqué que vous vouliez vous lancer en tant que « smart-up », c'est-à-dire financer le lancement de la « deuxième version » de Datakalab, axée technologie, grâce aux revenus générés par la « première version » de Datakalab axée conseil.

Le bracelet, c'est un outil qu'on utilise de moins en moins, c'est un choix. On privilégie la technologie qu'on est en train de développer en propre ici. C'est une belle technologie, mais on veut se concentrer sur ce qui permet de payer les salaires, en l'occurrence, l'algorithme. On pense que des gens vont vouloir racheter le bracelet. On essaye de créer notre technologie, mais bien sûr on ne se finance pas uniquement avec les revenus du bracelet. On veut signer des abonnements.

Avant et après le pivot, on a maintenu la promesse de « mesurer les émotions » des consommateurs. Penses-tu que cette promesse ait évolué ?

On veut toujours donner plus d'informations sur les réactions des gens, mais nous n'étudions plus tant les émotions que les réactions. C'est l'attention qui a été revalorisée dans la « deuxième version » de Datakalab, même si on en parlait déjà beaucoup dans la « première

version ». Ensuite, ce qu'on sait maintenant détecter comme le clignement des yeux, « l'eye tracking » va nous donner beaucoup plus d'informations. Dans notre promesse, on a décidé de faire évoluer l'offre, et de ne pas se cantonner aux émotions. On explique maintenant que les émotions ne sont qu'une partie de l'information. On va essayer de tout détecter sur le visage, de tout comprendre sur le visage.

Cela implique aussi la détection de nouvelles émotions par rapport aux émotions de base de Ekman ?

Non, parce que ce n'est pas scientifique. On ne va pas créer de la science qui n'en est pas. Si on s'intéresse à des émotions qui ne sont pas universelles, on instaure un biais. Ce n'est pas quelque chose qu'on veut faire. On veut rester dans de l'objectif, avec les émotions de Ekman, telles que Ekman les a définies, les clignements des yeux, la respiration, le *eye tracking* : là où la personne regarde, les mouvements des yeux, la lecture, des éléments faciles à déterminer à l'œil nu. On ne va pas essayer de définir la confusion. On ne veut pas recréer la théorie de Ekman, on veut pouvoir l'appliquer, et appliquer tout ce qui est détectable sur le visage.

Qu'est-ce que tu penses des travaux de Ekman sur lesquels on se base ?

Il faut avoir l'idée de dire que les émotions sont universelles alors que cela va à l'encontre de tout ce qui est principe de mémorisation, tout ce qui est mémoire personnelle etc. Je trouve que ses travaux précurseurs sont plus intéressants que les travaux de sa fin de carrière. Dans les travaux de fin de carrière, il parle des émotions secondaires et dit qu'elles ne sont pas généralisables. Pour essayer de comprendre quelqu'un, c'est bien, et c'est l'un des premiers à avoir dit que pour comprendre quelqu'un il faut le regarder.

À l'heure actuelle, quelles sont les difficultés, enjeux et menaces pour Datakalab ?

On développe bien notre technologie, mais il faut vraiment avancer avec notre nouvelle cible. C'est en bonne voie. Il faudrait qu'on soit plus *sustainable*, être une vraie entreprise dans le sens où il faut faire du chiffre et être encore plus rentable. Il faut avoir ce *mindset*, se concentrer sur l'utilisation de la technologie actuelle pour avoir de la rentabilité. Ayant la rentabilité, il faut l'assurer sur le long terme pour pouvoir grossir de manière organique et intelligente. On a les meilleurs chercheurs, donc il faut développer ça. Mais il n'y a pas trop de peur à avoir.

Est-ce que tu peux me parler de la façon dont tu traites les données brutes issues du travail d'Arnaud et de Kevin ?

Je regarde la qualité du signal, si la qualité n'est pas bonne, je ne regarde pas la vidéo. Ensuite, je regarde si la personne est bien devant la caméra, si elle est bien attentive, j'essaie d'analyser chacune des émotions et les *action units*. Je détecte les moments importants avec une forte amplitude dans le signal. Si ces moments remplissent certaines conditions, je sélectionne ces moments comme des moments de fortes émotions. J'agrège ces informations, puis je donne l'information au niveau de l'étude.

En quoi es-tu particulièrement attentif pour qu'il n'y ait pas d'erreur ?

En *machine learning* on peut se demander quand on définit un moment d'émotion comme un vrai moment ou comme étant du bruit. La structure du visage fait que tu peux avoir du bruit. Parfois on prédit quelque chose qui n'est pas bon : on dit que c'est vrai mais c'est faux. Parfois on dit que c'est faux, mais c'est vrai. Je fais très attention à ce que je prédis comme vrai soit vrai. C'est une fonction qui est clé que je vais donner pour mes clients, et je veux leur donner une information pertinente et vraie. Je ne veux pas leur donner des « seuils » plus bas mais qui laisseraient passer plus d'erreurs. J'essaie d'être très attentif sur la qualité de la *data*.

Si Datakalab venait à s'internationaliser, comment pourrait-on être sûr que l'éthique soit toujours respectée ?

Le respect de l'éthique et de la RGPD est dans notre essence. Notre but est de respecter la vie privée, les gens. On est sur un sujet sensible et plus il y a de rigueur, mieux c'est. Il faut avoir le contrôle et faire attention à ce qu'on accepte. Datakalab aux États-Unis par exemple, c'est la même chose. Je ne vois pas de différence, on demandera toujours le consentement des gens. Certes, il n'y a pas la RGPD aux États-Unis, mais les gens prennent de plus en plus conscience de l'importance de la protection de leurs données. En Chine, c'est très différent, mais on n'y est pas encore.

Qu'est-ce que tu imagines pour le futur de Datakalab ?

Développer des technologies de pointe pour répondre à des « use cases » où l'humain intervient. Utiliser des technologies de très haute qualité comme on peut le faire sur des trackers de visages pour avoir toujours une information pertinente même dans des conditions compliquées.

Donc rester sur le visage ?

Oui, le visage est notre clé.

ANNEXE 8

Entretien avec Maureen Kahn, Ex Project Manager chez Datakalab, 02/07/ 2019, Paris 17^{ème}

Maureen Kahn, diplômée de l'ESCP Europe travaillait avec moi en tant que « Project Manager » chez Datakalab de septembre 2018 à juillet 2019. Nous réalisons ensemble les études émotionnelles et assurons le suivi auprès des clients ainsi que la gestion de la communication de Datakalab.

Pourrais-tu présenter Datakalab comme on le fait en salon ?

Datakalab fait de la mesure d'émotions sur tous les points de contact de la marque : spots publicitaires, affiches, bande annonce, mais aussi lors de conférences. Notre dernier outil *Web Experience* a été dévoilé au CES à Las Vegas et permet de faire de la mesure émotionnelle sur site web, en partenariat avec IBM Tealeaf.

Quelle est la différence entre ce que les annonceurs pensent du neuromarketing, et ce qu'on fait vraiment chez Datakalab ?

En salon, on rencontre encore beaucoup de personnes qui définissent le neuromarketing comme : « l'utilisation des outils médicaux pour entrer dans la tête des gens » mais chez Datakalab, ce n'est pas du tout ça puisqu'on fait de l'analyse des micro-expressions qui sont les signes extérieurs de l'émotion. Nous n'utilisons pas des électrodes ou l'IRM. Il y a un manque d'éducation sur le neuromarketing en France et les gens ont peur de s'engager dans quelque chose qui peut sembler non éthique.

Comment convainrais-tu un annonceur réfractaire à l'utilisation du « facial coding » pour optimiser les contenus publicitaires ?

On a beaucoup de « use cases » qui nous permettent d'être crédibles auprès des entreprises démarchées. Aujourd'hui, il y a de plus en plus de start-up. Les géants se tournent vers les petites entreprises qui sont spécialisées et qui se développent dans la tech notamment. Par exemple, Éram a besoin de nous pour se moderniser dans la mesure émotionnelle de site web.

Dans quelle mesure est-ce qu'on adapte notre discours au client ?

Le but est que les clients comprennent les résultats et puissent les utiliser. Autrement, le rapport serait inexploitable. Pour l'un de nos clients, on a dû adapter nos métriques au cinéma. Le client va ensuite montrer à ses propres clients notre présentation, et il faut qu'elle soit compréhensible.

Pourrais-tu m'expliquer comment nous établissons les normes dans le cadre de notre partenariat avec Exterion Media ?

Nous avons signé un abonnement sur un an avec Exterion Media. Plus on a de données, plus les normes sont fiables. Notre partenariat avec Exterion nous permet d'étudier des affiches régulièrement et donc de construire des normes de plus en plus fiables. En termes de corrélations, on catégorise l'affiche avec des critères « présence d'un enfant », « couleurs chaudes » ... pour ensuite savoir si un élément fonctionne mieux que d'autres.

Penses-tu qu'on devrait avoir des normes par secteur : par exemple, des normes qui s'appliqueraient spécifiquement à l'automobile ?

Il faudrait d'abord avoir des normes générales, et après entrer dans le détail des catégories.

Quelles recommandations ferais-tu à Datakalab ?

Il faut vraiment se concentrer sur des clients qui sont prêts à signer des abonnements et perdre moins de temps sur des clients moins importants. Mais c'est difficile de savoir directement si les prospects sont qualifiés ou non. Il faut avoir des revenus fixes. Peut-être que le fait d'ouvrir la plateforme permettrait de passer moins de temps sur les études au bureau. Il faudrait alors être sûrs que nos clients sachent utiliser la plateforme, et que cela ne décrédibilise pas Datakalab.

Penses-tu qu'on puisse vraiment « mesurer les émotions » propres à l'homme, avec les algorithmes, propres à la machine ?

Oui, je pense. On a vu que nos algorithmes étaient relativement avancés. Quand on les teste au bureau on se rend compte de tous les points qui suivent le visage. On sait qu'il est robuste et qu'il se développe de plus en plus.

Pourrais-tu me parler de ce qu'on fait avec l'institut Rafaël ?

On peut toucher à tout un tas de domaines avec Datakalab. Le médical est un sujet crucial, et c'est important pour une société de pouvoir aider aux avancées dans ce secteur, et de pouvoir optimiser la qualité de vie des patients. L'institut Rafaël est une maison de l'après-cancer. C'est un organisme particulier parce qu'il y en a très peu en France. Lors des entretiens médecins-patients, on aimerait savoir comment le patient réagit. Pour le moment on est en phase d'implémentation. Cela veut dire qu'il y aurait une caméra dans la salle, le patient en serait prévenu. Datakalab demande toujours l'accord. Si le patient n'est pas d'accord, soit on coupe la caméra, soit il se rend dans une salle où il n'y a pas de caméra. Dans le monde médical il n'y a pas beaucoup de temps. Il y a des relations entre le corps médical et le patient, mais ces relations ne sont pas assez poussées. On veut savoir comment le patient se sent pendant les consultations pour ensuite l'aider à suivre son traitement.

ANNEXE 9

Entretien avec Edouard Yvinec, Machine Learning Scientist Junior chez Datakalab, 14/06/2019, Paris 17^{ème}

Edouard effectue un master en « machine learning » à Sorbonne Université. Il est entré chez Datakalab pour développer un algorithme de détection du regard.

Pourrais-tu m'expliquer ton travail dans les grandes lignes et le lien avec la détection des points caractéristiques du visage ? Quelles sont tes difficultés au quotidien ?

Je fais du *gaze estimation* : à partir d'une image monoculaire qui va être récupérée via une webcam par exemple, il faut estimer où la personne regarde sur l'écran. On ne peut pas trouver de solutions générales quel que soit le contexte. Donc il faut faire plusieurs suppositions : où est-ce que la caméra se situe par rapport à l'écran ? Ça peut être problématique si quelqu'un pose sa caméra sur sa tour, ce n'est pas du tout la même chose que si la webcam est posée sur l'écran. Mais en général la webcam se situe sur l'écran et c'est la première hypothèse que l'on pose.

C'est très différent de ce que fait Arnaud, mais on utilise les mêmes méthodes, des réseaux de neurones artificiels *deep*, et essentiellement tout ce qui est « convolution », donc traitement d'images. Arnaud se concentre sur la détection de points. Il prépare le terrain, et tout notre travail s'appuie sur le sien. Arnaud nous donne une étude assez précise de la façon dont le visage est composé face à la caméra. À partir de ces points, j'arrive à récupérer une bande sur les yeux de la personne. C'est à partir de cette bande, et d'informations comme l'endroit où la personne se tient par rapport à sa caméra, que je dois être capable de dire où la personne regarde. Arnaud travaille sur une information qui peut être utilisée de différentes manières, moi je *process* pour obtenir une information, une sorte de conclusion.

Les réseaux de Arnaud sont très gros et sa tâche est très difficile. Moi je pars avec beaucoup d'informations, puisque le regard est déterminé par là où est la tête. Arnaud part avec une information bas niveau. Il devait déjà avoir un gros réseau. Mais pour entraîner un gros réseau, il faut beaucoup de ressources, et lui ne peut lancer les expériences qu'une à une. Il doit réduire

la taille de son algorithme. Arnaud part d'une information très brute et doit la rendre plus fine. Moi j'ai déjà une information fine que je dois rendre encore plus fine.

Quelle est la part d'automatisation effectuée par la machine en regard de la compétence humaine ?

Pour répondre, je vais t'expliquer rapidement comment fonctionne le *deep learning*. Imagine une fonction classique, la fonction carrée qui ressemble à un bol, à une cuve. Quand tu calcules une distance, tu calcules le carré de quelque chose. On cherche à minimiser la distance. C'est exactement ce que je fais. Quand je dois estimer le regard, je veux que la différence entre ma prédiction et la vérité soit la plus faible possible. Pour ça, il faut arriver vers le fond du U, le fond de la cuve. Si je suis à droite de la partie montante, la dérivée est positive, on cherche à descendre, à prendre la dérivée. Si on descend, on veut avancer, et ta dérivée est négative donc tu reprends « moins la dérivée ». Peu importe où tu es, tu prends la dérivée pour avancer. C'est ce qu'on appelle une « descente de gradient » ainsi nommée parce que pour la descente, il y a l'idée de descente dans la « cuve ». Dans l'exemple que je t'ai donné, la dérivée est le gradient dans la dimension 1. Mais tout ça peut être généralisé. Quand on a une image HD, on a 1080 dimensions x 1920 dimensions, environ deux millions de dimensions. La fonction ne va pas être tout à fait une cuve, elle va être bosselée. On ne veut pas se retrouver coincé dans une petite bosse par exemple, autrement, on n'est pas tout en bas. Si c'est le cas l'algorithme n'arrive pas à s'améliorer. Il a un biais d'apprentissage. Voilà l'idée générale de la descente de gradient.

L'algorithme a des paramètres. Il passe l'exemple sur les paramètres, fait une erreur, il calcule l'erreur et depuis cette erreur, il va arriver à l'utiliser pour corriger chacun des paramètres en arrière. En maths, c'est de la « dérivée de composée de fonction ». Ici le réseau c'est pas une fonction, c'est plein de fonctions que tu enchaînes. Ce n'est pas moi qui le fais, il y a des fonctions standards qui sont déjà implémentées. Maintenant, il faut savoir quelle fonction on compose et de quelle manière. C'est la première chose sur laquelle on peut jouer. Quand la cuve n'est pas belle, il faut parfois descendre petit à petit et d'autres fois, il faut descendre un peu plus vite pour ne pas se retrouver coincé dans une bosse. Donc on peut jouer sur le rythme d'apprentissage. J'ai pris deux exemples, mais il y en a plein d'autres. C'est ce qu'on appelle les « hyperparamètres », les paramètres sur lesquels l'algorithme ne joue pas, mais l'homme joue. On a donc d'abord l'architecture, ensuite, les hyperparamètres, et ensuite l'approche du problème.

Ici pour le regard, on pourrait dire que l'œil est comme une sphère. Donc si j'ai le centre de l'œil et la pupille, je trace une droite, et je sais dans quelle direction tu regardes. C'est ce qui se fait en imagerie médicale, sauf qu'en imagerie médicale on n'utilise pas des webcams de mauvaise qualité. Donc on peut se demander si avec une méthode qui repère les points de façon approximative, on pourrait extrapoler de manière assez précise, mais ça ne fonctionne pas.

Mais dans l'approche générale du problème, parfois, le *deep learning* n'est pas forcément nécessaire. C'est la compétence humaine. Après, l'interface entre la compétence humaine et la machine, c'est que la machine réalise vite et bien cette tâche. C'est comme une calculatrice. En *deep learning*, on entend parfois dire « on ne sait pas comment l'algorithme apprend », mais je viens de te l'expliquer, en soi on sait comment il apprend. On dit ça parce que dans un algorithme, on a beaucoup de paramètres, qui sont des chiffres liés à des paramètres de paramètres. C'est toute une imbrication. On sait comment l'algorithme a progressé, on pourrait le refaire à la main, mais une fois qu'on a les valeurs finales, et le modèle qui est prêt à être exploité, on ne peut pas le traduire en règle explicite. Une image est par exemple stockée sous forme de chiffres, et donc on ne comprend pas ce que c'est. On sait comment l'algorithme apprend, mais on ne connaît pas son jeu de règles. Par contre, on peut tester l'algorithme, mais on n'a pas de méthode automatique pour connaître les règles humaines. C'est là la difficulté, parce que parfois l'algorithme plante mais on n'a pas de moyen de savoir ce qui lui pose problème, par exemple, si la personne a tourné la tête.

Ici, ça crée du « bruit » ?

C'est ça, exactement. Le « bruit », tu peux le tester. Au début de l'apprentissage, l'algorithme trouve des vérités gros grains. Il faut tenir compte du « *dataset* académique », qui vérifie les bases de données de l'état de l'art pour savoir si les vérités qu'il a sont représentatives.

Est-ce que tu pourrais m'expliquer plus en détail comment tu entraînes l'algorithme ?

Il y a deux phases dans l'apprentissage. D'abord, dans la partie académique, on va utiliser ces « *dataset* académiques » qui sont utilisés par les autres chercheurs. Ces *dataset* se présentent comme des *crops* de la personne, on ne voit que ses yeux, et le *dataset* nous dit que la personne a regardé à tel endroit. J'ai mon exemple, et la « vérité terrain », ce vers quoi je dois me

rapprocher dans ma prédiction. J'ai deux « sets » : l'un pour l'entraînement et l'autre pour l'évaluation. Je donne à l'algorithme uniquement des « sets » d'entraînement et je vais lui dire de s'optimiser. Une fois qu'il s'est optimisé, je lui montre des exemples qu'il n'a inédits, et je regarde ce qu'il fait.

En apprentissage supervisé, tu dis à l'algorithme quelle est la « vérité terrain », si tu veux l'entraîner pour reconnaître des visages, tu lui dis ce qui est un visage et ce qui n'en est pas un. Là il est supervisé parce que je lui donne la réponse pour qu'à la fin, il puisse connaître son erreur. Il y a plusieurs méthodes pour entraîner l'algorithme. On a testé une méthode de *curriculum learning*. L'idée c'est d'y aller progressivement, d'abord on donne à l'algorithme des exemples faciles, et on lui rajoute petit à petit des exemples plus complexes. Pour nous, ça n'a pas fonctionné. Les exemples simples, comme lorsque la personne est centrée face à la caméra avec le visage tourné vers l'écran, ne l'ont pas aidé. Je pense que c'est parce que l'algorithme traite déjà facilement ce genre d'information de lui-même, donc ça ne change rien si on lui donne des exemples encore plus simples sur une tâche qui lui est déjà facile.

Aujourd'hui, ce qui fonctionne bien dans les conférences, ce sont les « données synthétiques ». On crée des faux visages qui ressemblent à des visages humains. On a un réseau qui va faire des prédictions et à côté, on a une prédiction qui dit que tel visage réel, ressemble beaucoup à tel visage synthétique. Sauf qu'on sait à quoi correspond la prédiction pour le visage synthétique. Du coup, on a à la fois la prédiction faite pour la vraie personne, et la prédiction faite pour le synthétique, qui est extrapolée. On peut ensuite comparer ces deux prédictions. L'algorithme peut alors combiner ces deux informations, par exemple grâce à une moyenne.

On peut aussi chercher à transformer l'image des yeux en carte de chaleur de là où la personne regarde. L'œil est une image particulière qui ne se prête pas spécialement au fait d'être transformée en carte de chaleur. On part d'une image, et on en va remodeler une autre, ici pour tâcher d'obtenir la carte de chaleur d'une vision, mais ce n'est pas évident pour l'algorithme. Toutes ces méthodes rentrent dans ce qui s'appelle des « algorithmes stochastique », c'est-à-dire probabilistes, aléatoires. Ils apprennent tous à partir de paramètres définis de manière aléatoire, parce qu'on ne veut pas que toutes les prédictions soient identiques. Grâce à cela l'apprentissage final est aléatoire. Aussi, à chaque fois qu'on le relance, on peut se demander s'il va apprendre de la même manière, mieux ou moins bien.

Pour la deuxième phase, il s'agit d'entraîner l'algorithme sur les données de Datakalab. Elles sont différentes des données académiques, parce que ces données académiques regardent toutes les zones de l'écran, alors que nous n'allons regarder que certaines parties de l'écran, ce n'est pas réparti de la même manière. Pour l'algorithme, le début de l'apprentissage influence ce qui suit, et il faut qu'il y ait la même distribution entre les données d'entraînement et les données de test. Si dans l'entraînement les personnes regardent toujours au milieu, l'algorithme va apprendre à regarder les gens qui regardent au milieu. Il faut que tout soit bien distribué et que tout soit bien renormalisé, mais ce n'est pas évident, parce qu'on manipule de très gros objets. Les gros réseaux de neurones ont 160 millions de paramètres. Même si la technologie progresse, on traite des objets énormes difficiles à appréhender comme un ensemble.

La *batch norm* est une fonction qui remet toutes les données à 0 avec une variance à 1. L'intérêt, c'est de tout ramener dans les mêmes valeurs pour que d'un point de vue probabiliste ça ait du sens, que tout soit indépendant et identiquement distribué. En probabilités, la loi la plus standard, c'est la loi de Gauss. Sa quintessence, c'est quand elle est « centrée réduite », en moyenne de 0, avec une variance de 1. Donc elle ne varie pas trop. Pour que les calculs restent standards, il faut les ramener à du 0/1.

Quelle est la particularité d'un algorithme qui détecte les « action units » si on le compare à un algorithme qui détecte les objets ?

C'est la même chose que la différence entre la mécanique newtonienne et la mécanique des fluides. Pour la mécanique newtonienne, si on prend une balle, et qu'on la jette, on veut savoir où elle est dans les temps qui suivent. C'est simple, parce que c'est un objet solide. Au contraire, dans la mécanique des fluides, on prend un tuyau avec de l'eau qui coule, on fait un geste ample, et on se demande où est l'une des gouttelettes d'eau. C'est compliqué, aussi parce que le mouvement d'un fluide est spécial. C'est la même chose avec les *action units*, les muscles qui bougent sur le visage, mais il n'y a pas de rupture de la continuité du visage. Donc d'un côté on a quelque chose de solide et de facile à repérer et de l'autre on a quelque chose de fluide à repérer dans une masse.

Est-ce que tu penses qu'on pourrait apprendre à l'algorithme à analyser des déplacements du corps et si oui, comment pourrait-on les utiliser ?

Tu peux lui apprendre à gérer n'importe quel déplacement du corps. Le problème, c'est qu'il va falloir donner à l'algorithme des exemples de gens qui se déplacent, mais pour que ça fonctionne, il faudrait lui apprendre sur tous les types de mouvements possibles, des mouvements de gens qui marchent, qui font de la gymnastique, du breakdance. Mais là, il faut annoter, dire où étaient tous les éléments à tel instant. Quand on montre à l'algorithme une photo ou une vidéo, il y a un fond derrière. Donc il faudrait lui donner des exemples avec un fond urbain, un fond rural. Pour estimer le regard, je n'ai que les yeux, donc je n'ai pas de fond. Si j'avais la tête, on peut penser que le fond est soit un mur, soit une porte, soit une fenêtre. Mais si on étudie le déplacement du corps, la personne peut se déplacer où elle veut, et il faut aussi prendre en compte des conditions de luminosité. L'enjeu, c'est que ça fait une quantité folle de données. À ma connaissance, il n'y a pas de donnée de l'état de l'art là-dessus.

Dans 10 ans, qu'est-ce que tu penses qu'on saura faire en plus dans la détection de l'émotion ? Je pense que ce qui pose le plus problème aujourd'hui c'est la compréhension qu'on a de l'émotion, d'un point de vue psychologique. Il n'y a qu'une émotion qui soit vraiment positive sur les sept de base de Ekman. Quand on demande aux gens de mimer ça, les gens font des expressions ridicules. Quand une personne est ironique, même l'humain ne le comprend pas toujours. En soi, l'apprentissage mathématique ne dépend pas du modèle psychologique, on se met d'accord sur une vérité terrain, et ensuite l'algorithme détecte certains éléments du visage pour faire des conclusions. Je pense qu'on ne comprend pas encore bien les expressions humaines. Certaines personnes ne sont pas expressives, d'autres biologiquement sont nées avec des sourcils plus froncés, d'autres sont paralysées d'une partie du visage. On arrivera à être au top de ce que les modèles psychologiques proposent. La question, c'est : est-ce que les modèles actuels suffisent, ou est-ce que dans 10 ans on n'aura pas besoin de modèles plus précis ?

ANNEXE 10

Entretien avec Kevin Bailly, Head of Research chez Datakalab, 19/07/2019, Paris 17^{ème}

Enseignant-Chercheur à Sorbonne-Université, Kevin Bailly a rejoint Datakalab fin 2018 pour créer les algorithmes propriétaires. Il a remporté de grands concours internationaux, et l'un de ses papiers de recherche a été sélectionné pour être présenté à ICCV 2019, l'une des principales conférences en « computer vision ».

Peux-tu m'expliquer ton travail chez Datakalab ?

Mon travail consiste essentiellement à mettre en place des algorithmes pour l'analyse des expressions faciales. Je construis ces algorithmes, je fais de la programmation puis je publie les résultats. C'est un travail de recherche assez standard. On se demande aussi comment ce qui sort de nos algorithmes peut être utilisé par Datakalab. Là on rejoint le travail de Lucas, avec la détection de pics, la manière dont on les agrège, quelle métrique est pertinente pour le client. Enfin, je cherche des gens à recruter pour la partie « tech ». À terme, il s'agit plus de superviser que de coder au jour le jour. Pourtant, coder me fait voir les problématiques de l'implémentation qui ont beaucoup évolué avec les nouveaux *framework* en *deep learning*.

Beaucoup de termes circulent autour de ta discipline comme « l'affective computing ». Que penses-tu de ces termes à la mode ?

Ce type de termes est vendeur mais assez creux, pourtant ce sont aussi des termes qui permettent de fédérer une communauté. Par exemple, pour le *deep learning*, il y a ceux qui l'utilisent, ceux qui contribuent à sa construction, et ceux qui codent des *framework* qui vont être utilisés par toute une communauté. Ce ne sont pas du tout les mêmes métiers.

Depuis le repositionnement, as-tu pu remarquer des changements dans ton métier ?

Non, le pivot est survenu au moment où je suis moi-même arrivé. Le centre de gravité de la start-up a changé. Avant, on était axé sur le conseil, mais on sous-traitait la partie « algorithmes » par Microsoft ou autres. Datakalab collectait les vidéos en amont avec sa

plateforme propriétaire, et interprétait les résultats. Mais pour la partie traitement et analyse faciale, Datakalab n'avait pas de solution propre.

Que penses-tu des travaux de Ekman ?

J'ai toujours fait en sorte de ne pas m'intéresser à la pertinence de ce modèle-là. Le modèle des *action units* est pertinent, il est objectif et juste relié à l'activation musculaire, il n'y a pas d'interprétation pour savoir si le fait de froncer les sourcils fait que la personne est en colère ou pas. On se demande juste si elle fronce les sourcils et à quelle intensité. Des experts doivent annoter cela dans ce code FACS, mais ils coûtent chers, et il y en a peu, donc on n'a pas beaucoup de données sur lesquelles on peut apprendre. On essaye d'apprendre sur des modèles inspirés du FACS mais qui ne demandent pas d'être formés au États-Unis pendant trois semaines pour savoir annoter les données.

Tu peux m'en dire plus sur ces données qui sont inspirées du FACS ?

Ce qui nous intéresse, ce n'est pas de savoir si c'est « *l'A.U. quatre* » ou « *l'A.U. douze* » qui est activée, mais on veut savoir si la personne a levé ses sourcils, si elle a souri, si elle a plissé son nez. C'est la même chose que les *action units* les plus faciles à détecter mais sans qu'on l'estampille avec le label *action units*. Ekman a fait deux choses importantes, il a codifié chaque activation musculaire par une *action units*. Il a ensuite essayé de faire un lien entre les *action units* et les émotions. On n'utilise ni l'un ni l'autre mais on utilise un dérivé.

Edouard m'a expliqué que pour détecter les « action units » tu utilisais une méthode de calcul de distance entre les points. Peux-tu m'en parler ?

Oui, cette technique est assez standard dans le domaine de l'analyse émotionnelle. Elle combine distance et apparence du visage. L'une des méthodes les plus simples consiste à prendre le visage, découper l'apparence d'un visage, la mettre à l'entrée d'un réseau et il apprend tout seul. Ce qui pourrait bien marcher si on avait énormément de données, mais ce n'est pas le cas. Donc on est obligés de faire du prétraitement, de détecter des points caractéristiques et de calculer des distances. Ça nous donne une information sur la « géométrie du visage ». Alors on donne des distances plutôt que des points directement parce que ce sont des descripteurs qui sont invariants en rotation. Si je tourne ma tête, la distance entre les points reste la même, alors que

la position du point dans l'image a bougé. On prémâche le travail de l'algorithme en lui disant d'utiliser les distances. On lui donne plein de distances et ensuite c'est à lui de trouver quelles sont les celles pertinentes pour prédire l'émotion ou l'activation musculaire.

Est-ce que tu connais les différences entre nos algorithmes et ceux de nos concurrents ?

C'est un secret bien gardé. Il n'y a pas beaucoup de publications chez nos concurrents, et elles sont sur des points spécifiques à chaque fois. On ne peut pas savoir ce que font les concurrents, on peut juste dire où ils en sont par rapport à l'état de l'art académique. Il y a des publications sur les points, la détection de points caractéristiques, mais pas tellement sur l'émotion. C'est aussi une question de communautés. Ceux qui s'attaquent à la localisation de points sont plutôt des gens qui viennent de la « vision ». Ils ont des problèmes de « vision » : « comment on détecte un piéton, une chaise, etc. » mais ils en viennent à s'intéresser à la détection de *landmarks*, des points caractéristiques du visage. Pour détecter ces points du visage, il faut trouver les points des yeux, de la bouche, ce genre de choses. Récemment, il a eu beaucoup de bases de données annotées pour la communauté. Du coup, beaucoup de gens qui faisaient de la « vision » pure s'y sont intéressés parce que ça marche bien sur des algorithmes *deep learning*, qui ont besoin de beaucoup de données pour apprendre. L'émotion, c'est compliqué, parce qu'il y a beaucoup moins de bases. Les annotations sont plus hétérogènes, et il y a plus de bruit dessus. Un clin d'œil est facile à définir, mais pour la joie, c'est plus compliqué. Cela reste sujet à l'interprétation. Même les bases les plus conséquentes sont mal annotées. Les algorithmes sont spécifiques au domaine, et c'est pour ça qu'il y a moins de gens qui travaillent là-dessus.

Pourquoi a-t-on si peu de données entraîner les algorithmes à la détection d'émotions ?

Déjà parce qu'il n'y a pas beaucoup de monde qui travaille sur les émotions, les expressions faciales, ou les points. Comme peu de personnes travaillent dessus, on a peu d'annotations. Ensuite, les émotions sont difficiles à annoter parce qu'il y a différentes manières de représenter une émotion et une « expression faciale ». Il y a les expressions de Ekman, mais il y en a plein d'autres, on peut « mesurer » les intensités de « valence », ou les intensités de « arousal », « dominance expectance ». Les émotions peuvent être décrites de plusieurs manières. Le modèle des émotions de Ekman est restrictif, « l'arousal » est difficile à annoter. La « valence », ce n'est pas si dur parce qu'on voit si la personne est plutôt contente ou non. À ce moment là cela revient juste à savoir si elle est joyeuse. On peut ramener ça à de la joie et à de la tristesse

avec un curseur entre les deux. Après, tout le reste est compliqué à annoter. En plus de cela, ceux qui annotent le font en continu : ils regardent la vidéo, et si la personne sourit, ils vont dire que la « valence » a augmenté mais ils ont mis deux secondes à réagir donc ça veut dire que quand on met les données en entrée de la machine, on va avoir des images où il y a un sourire, mais qui va être annoté comme une « valence » faible, parce qu'on n'était pas encore sur le moment où la personne a réagi. Si on demande à dix personnes d'annoter, on se rend compte que la « valence » dans l'annotation est assez forte. Non seulement on a peu de données, mais en plus l'annotation est très bruitée.

Les micro-expression durent 1/15 à 1/25 de seconde, est-ce que nos algorithmes les détectent vraiment ?

Dans « micro », on comprend le côté rapide et le côté subtil, c'est une faible variation. Les vraies micro-expressions, l'algorithme ne va pas les capter. Il aurait le temps de les voir, mais il faut ensuite être capable de les différencier du bruit. Nos bases de données sont souvent capturées dans des bonnes conditions. Ce n'est pas le cas de ce qu'on va récupérer des personnes qui font ça chez elles avec leur webcam, qui peut être différemment orientée, elles peuvent faire ça dans le noir, surexposées si elles sont devant leur fenêtre. Il y a vraiment un écart entre nos données annotées et les images qu'on récupère. Ce n'est pas le cas pour les *landmarks*. Même dans des conditions difficiles, on est toujours capables de détecter ces points et on peut collecter plein d'images qu'on peut annoter en cliquant. L'algorithme pourrait donc détecter les micro-expressions, mais dans des contextes beaucoup plus contrôlés, peut-être avec des cas d'utilisation pour Datakalab avec des caméras plus haute fréquence.

Pour la sécurité par exemple ?

Oui, en l'occurrence, la personne sera plus coopérative, on lui demandera de se placer devant la caméra pour être capable de capter des mouvements subtils. Pour l'instant, on ne le fait pas. Souvent la résolution et le débit de la connexion des personnes ne permet pas d'avoir 30 frames par seconde, or les micro-expressions durent 1/15 de seconde. Avec des caméras normales, on est à 25 ou 30 images par seconde donc on pourrait les avoir mais on n'a même pas ça dans la plupart des cas. Pour « mesurer des réactions » face à des pubs, on n'a pas besoin de détecter la pupille qui se dilate, puisqu'on peut agréger ça sur 300 personnes. Si les personnes sourient face à une pub, parce que le moment est drôle, il y a de bonnes chances pour qu'une bonne

proportion des gens sourient. Si on fait la somme de toutes les personnes qui ont souri, on va être à 60 ou 70 %. Par contre si c'est une erreur de détection, ou si la personne a souri parce que quelqu'un est entré ou lui a fait un signe il y a peu de chance que pour les 100 personnes, il y ait eu un événement qui les fasse rire qui soit indépendant de la pub. Du coup même si on a de fausses détections, comme ça n'arrivera pas au même moment pour tout le monde, si on fait la somme, on aura au maximum 10 %, donc 10 personnes qui auront souri au même moment ou parce qu'il y a eu un événement extérieur ou parce que l'algorithme s'est trompé. Il y a peu de chance que ça soit significatif par rapport au vrai pic où tu as 80 % des personnes qui vont sourire parce que c'était drôle.

ANNEXE 11

Entretien avec Arnaud Dapogny, Machine Learning Scientist chez Datakalab, 27/06/19, Paris 17^{ème}

Diplômé de CentraleSupélec, spécialisé en « machine learning » et en « computer vision », Arnaud était Enseignant-Chercheur à Sorbonne Université. Il a rejoint Datakalab fin 2018 et optimise aujourd'hui l'algorithme de détection de points caractéristiques du visage. Son dernier document de recherche rédigé avec Kevin Bailly a été sélectionné pour être présenté à ICCV 2019, l'une des principales conférences mondiales en « computer vision ».

Dans les salons, on vous présente toi et Kevin comme « les meilleurs chercheurs au monde ». Comment es-tu arrivé à devenir l'un des premiers chercheurs dans la détection de points ?

Je ne pense pas être l'un des « meilleurs chercheurs au monde ». Il existe plusieurs types de personnes : ceux qui ont une idée bien précise de ce qu'ils veulent faire, et ceux qui tâtonnent pour trouver leur voie. Je fais partie de cette seconde catégorie. J'ai fait médecine puis une école d'ingénieurs généraliste qui étudiait les systèmes, l'électronique, le traitement du signal. J'avais des cours de « traitement de l'image » en troisième année à Rennes avec Renaud Séguier. L'un des projets de son cours m'avait beaucoup plu. J'ai décidé de refaire un master après l'école d'ingénieur pour me spécialiser. Je ne me jugeais pas compétent sur ce genre de thématiques au regard de mes anciens stages. J'ai effectué un master de traitement d'images dans les années 2012-2013 à la Sorbonne. C'était le début du *machine learning*. J'ai fait mon stage à l'École des Mines, puis j'ai fait une thèse à Jussieu avec Kevin. C'est donc une suite de hasards qui fait que tu vas dans un domaine que tu ne connaissais absolument pas au départ. Suite à cela j'ai publié dans des conférences internationales et j'ai pu rencontrer la communauté.

Tu t'intéresses aujourd'hui à la détection de points sur le visage.

Avec la mode du *deep learning*, l'exemple le plus connu est celui des thématiques de classification/catégorisation d'objets. Il existe un *benchmark* avec dix millions d'images et mille catégories : des avions, des voitures, des chats, des chiens, et tu dois les classer une par une. Ce sont des domaines où il y a énormément de *data* et les développements architecturaux

sont très lourds. Avec les *data* du visage, on a des problématiques d'anonymisation, de labellisation. Cela coûte plus cher de faire annoter précisément 100 points sur le visage que de dire si c'est un chat ou un chien. Tu as moins de *data* et plus de place pour innover dans les algorithmes.

Sais-tu quels algorithmes Datalab utilisait avant ton arrivée ?

Avant, la start-up était plus axée « conseil » que « tech ». Quand j'ai rencontré Xavier, il faisait le lien entre des sociétés comme Microsoft ou Affectiva, qui proposent des solutions qui permettent d'analyser un contenu émotionnel, et leurs clients, des agences de publicité qui font tester le contenu. Datalab ne fournissait pas le service technique en tant que tel, mais plutôt la mise en relation. Maintenant, on peut développer des algorithmes en interne, ce qui permet d'être plus proche de cette attente-là. Les clients ne s'intéressent pas spécifiquement aux *action units*. Ils peuvent vouloir savoir si la personne porte un chapeau, ou si elle met du maquillage, on peut essayer de développer des solutions pour détecter ce genre d'éléments. On peut difficilement dire à Microsoft de modifier son API pour détecter telle ou telle chose pour un client inconnu. Pour les clients, cela permet de pouvoir interagir directement avec les propriétaires de la technologie. Cela permet aussi à Datalab de trouver de nouveaux marchés que la start-up ne pouvait pas adresser.

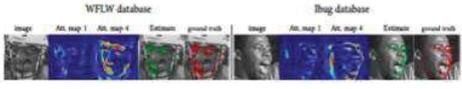
Sur ce « slide », on explique que la technologie Datalab est meilleure que celle de Facebook. Peux-tu me l'expliquer ?



One of Datalab research paper is selected at ICCV 2019. ICCV is one of the most important venue in computer vision and deep learning worldwide.

This Paper was written by two of our researchers Arnaud Dapogny, Machine Learning Scientist, and Kevin Bailly, Head of Research and Teacher at Sorbonne University along with Matthieu Cord, who also took part in this work.

Our Research is better than Facebook's



WFLW database: image, All. map 1, All. map 4, Estimation, ground truth

FRW database: image, All. map 1, All. map 4, Estimation, ground truth

Table 2: Comparison in terms of Mean error (lower is better), AUC (higher is better) as well as failure rate (lower is better), on WFLW.

metric	method	all	head pose	expression	illumination	make-up	occlusion	blur
ME (%)	CFSS [10]	9.07	21.36	10.00	8.30	8.74	11.76	9.96
	DVFN [10]	10.84	46.93	11.15	7.31	11.65	16.30	13.71
	LAB [10]	5.27	10.24	5.51	5.23	5.15	6.79	6.32
	Wang [1]	5.11	8.75	5.36	4.93	5.41	6.37	5.81
	DeCAF	4.62	8.11	4.65	4.41	4.63	5.74	5.38
AUC@0.1	CFSS [10]	0.366	0.060	0.316	0.385	0.369	0.269	0.305
	DVFN [10]	0.456	0.147	0.389	0.474	0.449	0.379	0.397
	LAB [10]	0.532	0.235	0.465	0.543	0.539	0.449	0.463
	Wang [1]	0.554	0.310	0.496	0.541	0.558	0.489	0.492
	DeCAF	0.563	0.392	0.546	0.579	0.575	0.485	0.494
FRW (TW)	CFSS [10]	20.56	49.26	23.25	17.34	22.84	32.88	23.07
	DVFN [10]	10.84	46.93	11.15	7.31	11.65	16.30	13.71
	LAB [10]	7.56	28.83	6.37	6.73	7.77	13.72	10.74
	Wang [1]	6.00	22.70	4.78	4.30	7.77	12.50	7.78
	DeCAF	4.84	31.4	3.73	3.22	6.18	9.26	6.61

<https://arxiv.org/pdf/1904.02549.pdf>

<https://arxiv.org/pdf/1907.03248.pdf>

<https://arxiv.org/pdf/1910.09450.pdf>

© Copyright Datalab 2019

On a un certain nombre de bases de données sur ce slide. Ici ce sont des modèles d'alignements de points. On va « mesurer l'erreur » point à point. On va aligner un point qui correspond à l'œil et donner à l'algorithme une « vérité terrain ». Ce sont des bases annotées, avec des personnes qui ont cliqué pour définir l'emplacement de la pupille. L'algorithme va te la prédire un petit peu plus loin. Donc tu vas « mesurer » cette distance-là entre tes deux points. Plus tu es proche de la « vérité terrain », plus tu es précis. Si tu as une « erreur » de 0, c'est que tu recopies exactement la « vérité terrain ». Plus ton « erreur » est petite, plus ton algorithme est performant. On mesure donc une « erreur » dans la prédiction de l'algorithme.

Quel est le degré de précision des algorithmes de Datakalab ?

« L'erreur » est une donnée connue de la communauté. On mesure une « erreur » point par point sur 100 points pour cette base-là. On fait une moyenne sur l'ensemble des points. Après, on normalise ça par la distance entre les deux yeux. On fait ça parce que si un visage est très loin et l'autre très proche, on ne veut pas qu'il y en ait un qui compte plus que l'autre. Ensuite l'opération se fait sur une grosse banque d'images. Sur celle-ci, il y a 2 500 images annotées. On a là des erreurs de 5 %, pour les humains, on est sur du 5 %. La machine est plus précise que l'humain pour prédire les points. C'est le cas aussi pour la reconnaissance d'identité.

Tu pars d'une « donnée brute » pour atteindre une granularité assez fine.

Oui, c'est ça. Les approches *deep* reposent sur ce postulat. On peut donner une « *data* brute » et la machine va apprendre une chaîne de transformation. Elle va abstraire sémantiquement la donnée. Pour l'ordinateur, une image est juste un tableau 100x100 pixels. Dans ce tableau, il doit trouver une suite de représentations pour en déduire s'il s'agit d'un visage qui sourit, par exemple.

Qu'est-ce qu'un « algorithme faiblement supervisé » ?

On utilise cette expression quand on apprend sur des données faiblement annotées. Annoter 100 points sur une image prend du temps, c'est à la fois cher et fastidieux. On veut que les algorithmes apprennent à prédire 100 points en utilisant dans leur phase apprentissage des données qui coûtent moins cher et qui sont annotées « plus faiblement », de manière moins

précise. Par exemple, on peut demander à l'algorithme de trouver une correspondance entre cinq points et cent. C'est grossièrement ce qu'on fait. La métrique qu'on a proposée dans cette conférence sert à intégrer cette logique d'apprentissage à partir d'un plus petit nombre de points.

Est-ce que tu connais le terme « d'affective computing », et si oui, pourquoi on ne l'utilise pas ?

Je connais ce terme, mais je trouve qu'il ne veut pas dire grand chose. Il n'y a pas lieu de faire la différence entre *machine learning* supervisé et *affective computing*. Deuxièmement, on a des *benchmarks* qui essaient de reproduire des robots qui interagissent de manière naturelle avec l'humain. C'est quelque chose qui ne m'intéresse pas plus que ça puisque quand tu n'as pas de métrique d'évaluation, tu peux faire n'importe quoi.

Tu considères que c'est un terme trop vaste ?

Trop vaste, et aussi politique. Même dans la recherche, on n'échappe pas à la logique capitaliste. Il faut vendre quelque chose. Si tu es chercheur, il faut que tu vendes ta recherche et tes idées. Si tu ne présentes pas ce que tu fais, que tu ne vulgarises pas, que tu ne te vends pas, personne ne t'écoute. *L'affective computing* se rapproche d'un terme « bullshit » politique pour vendre. C'est un nouveau terme comme celui de « brain tech ». À mon sens, tous ces termes veulent dire la même chose. Je considère que la recherche devrait être débarrassée de ces problématiques de traitement de l'information. Des mathématiciens comme Poincaré faisaient de grandes découvertes mais n'inventaient pas ce genre de termes pour faire passer leurs idées pour plus intéressantes qu'elles ne le sont.

Selon toi qui construit ces termes, les chercheurs, les communicants, les deux ?

C'est un peu des deux, mais ce n'est pas tant du côté communicant. Quelques scientifiques restent dans leur tour d'ivoire. Je caricature, mais certains chercheurs présentent une fois tous les trois ans leurs avancées. Aujourd'hui, sur du *machine learning* et de la « vision », il existe des applications industrielles directes, comme la voiture autonome ou ce qu'on fait là pour du marketing. Dès que cela intéresse un peu plus des entreprises et dès que ça rapporte plus, il faut se vendre. Les grandes conférences autour de ces domaines se déroulent dans des lieux

prestigieux comme Hawaï avec plus de 5 000 personnes, et le déplacement coûte cher. D'un côté, on peut y faire circuler des idées, de l'autre, on y crée des concepts anti-scientifiques.

Justement, quand tu vois nos présentations et nos « métriques », es-tu satisfait de ce qu'on rend d'un point de vue scientifique ?

À chacun son travail. Frank et Xavier me disent des choses qui me font parfois tiquer et je leur explique quand ce n'est pas tout à fait correct. Leur logique n'est pas la même qu'en recherche. Aussi nous devons essayer de créer des systèmes performants. Du côté marketing/communication, vous devez rendre ces données intéressantes pour le client. L'un de mes amis a fait ses études à Cambridge et travaille sur l'énergie houlomotrice. Cela fait des années qu'il travaille dans sa cave. Comme c'est un vrai scientifique, il ne veut pas avoir à se vendre. De ce fait, personne n'est au courant de ce qu'il fait, alors que ses idées pourraient faire avancer son domaine.

Est-ce que tu penses qu'avec la « data » propre à la machine, on peut « mesurer les émotions » qui sont propres à l'Homme ?

Le paradigme du *machine learning*, c'est que si l'humain sait le faire, il n'y a pas de raison que la machine ne sache pas le faire. Dans la plupart des cas, les émotions servent à entretenir un lien social, à comprendre les gens. Il y a émotion, ce qui se passe dans ton cerveau, et « expression », comment tu le montres. Cette manifestation-là a aussi un sens. C'est communiquer pour mieux t'intégrer, pour te faire comprendre et pour que les gens puissent raisonner avec toi.

Des émotions ne sont pas forcément exprimées par des « micro-expressions ».

Oui, pas mal de gens ont fait des recherches sur la manière dont le *body langage* et la voix expriment des émotions. Tout ne passe pas par ces canaux, mais c'est un vecteur. D'un point de vue scientifique, on s'est concentrés là-dessus parce que sous certaines conditions, si on est globalement de face, qu'on filme bien, c'est mesurable par l'homme et par la machine de manière assez satisfaisante. Ensuite, il ne faut pas faire dire aux *data* ce qu'elles ne disent pas. Si tu as pris un modèle avec 200 personnes de face et qu'on le teste en étant de biais, cela ne va

pas fonctionner. On peut perdre de l'information en se concentrant exclusivement sur des visages, plutôt que de prendre en compte le contexte, les mouvements des épaules, etc.

Quelles autres « métriques » pourrait-on imaginer, et quelles seraient leur utilité pour Datakalab ?

Ça dépend de ce que veut le client et du *use case*. Dans le meilleur des mondes, on pourrait mesurer beaucoup de choses, comme les micro-mouvements oculaires. Actuellement, on ne va pas jusqu'à ce degré de précision. C'est possible de le faire avec des images de très haute résolution, si on calibre correctement la géométrie du visage. Quand on apprend à analyser les émotions quand on est un bébé, on reconnaît à peine le visage de notre mère. Notre vue n'est pas performante, alors on passe par beaucoup d'autres canaux comme les odeurs, le toucher. En grandissant, on voit des milliards et des milliards d'images.

On parlait tout à l'heure du « faiblement supervisé ». On a des méthodes d'apprentissage qui n'impliquent pas de supervision. Pour schématiser, la supervision c'est « essai-erreur ». Dans « l'apprentissage non supervisé » on a plein d'images, on ne sait pas quoi en dire, et la question est « Qu'est-ce qu'on en fait ? ». Quand on apprend, on voit beaucoup de visages, on les catégorise, on les garde dans notre cerveau et on les enregistre. Surtout si ce qu'on a vu est original. On stocke cette information sous une certaine forme, avec des représentations internes dans notre cerveau. On n'a pas encore capté tous les fonctionnements dans nos systèmes.

Dans 10 ans, où penses-tu qu'on en sera dans la création d'algorithmes de mesure émotionnelle ?

Si on continue de se focaliser sur le négatif comme on le fait actuellement, on en sera au même point. La prédiction de l'émotion est un domaine qui n'avance pas très vite. Un thésard avait fait le challenge FERA²⁶⁰ sur la reconnaissance *d'Action Unit*, où on mesure « l'erreur » par rapport à la « vérité terrain ». Les méthodologies étaient bonnes, voire meilleures que celles de beaucoup de papiers qui sortent actuellement. Certains domaines, comme la voiture autonome, progressent vite. D'autres moins, puisqu'on s'enferme dans des termes comme *l'affective*

²⁶⁰ *Facial Expression Recognition and Analysis Challenge*

computing. On ne se pose pas des questions aussi simples que « pourquoi ça ne fonctionne pas ? ».

L'un des gros challenges du domaine, c'est le *cross-domain*, qui consiste à apprendre avec une base de données de personnes qui font des *action units* dans des laboratoires et après on veut les tester dans des conditions différentes, à l'extérieur, ou avec un visage qui est un peu tourné par exemple. Aujourd'hui, c'est quelque chose qui ne fonctionne pas pour la reconnaissance d'expressions. L'algorithme a appris dans un contexte et c'est compliqué de généraliser. Beaucoup reste à faire dans ce domaine, mais peu de chercheurs travaillent sur le sujet. En France, une dizaine de personnes en sont spécialistes. Dans le milieu scientifique, il faut arrêter de vouloir se vendre et être plus honnête quant aux capacités de nos systèmes. On peut dire quand on pose des jalons, mais il faut rester humble et dire quand nos systèmes ne sont pas fiables. Par exemple tout ce qu'on fait avec les points, ce sont des choses qui fonctionnent bien, qui sont matures et quantifiables très facilement. La pose du visage, c'est pareil. On mesure avec seulement 3-4 degrés d'erreur l'angle du visage, c'est précis.

ANNEXE 12

Entretien avec Éric Daoud, Ex Data Scientist & Machine Learning Engineer chez Datakalab, 04/06/2019, Paris 17^{ème}

Diplômé de l'ECAM Lyon et du Mastère spécialisé SIO « Systèmes Informatiques Ouverts » à CentraleSupélec. Il a ensuite travaillé comme « Data Scientist » chez ManoMano, une plateforme de vente en ligne de produits de bricolage et de jardinage ; puis chez Doctrine, une startup Legaltech, avant de rejoindre Datakalab.

Pourrais-tu me décrire une journée type ?

Tout change assez souvent en termes de missions. Je pars de tous les chiffres qu'Arnaud et Kevin me donnent grâce à l'algorithme de reconnaissance de visage, et je fais en sorte de transformer ça en résultats intelligibles. Je travaille sur la détection de pics, c'est du traitement de signal. Je fais des graphes et des corrélations pour montrer au client ce qu'on trouve dans le signal brut des émotions.

À quoi ressemblent les données à partir desquelles tu travailles ?

Arnaud et Kevin traquent des *action units*, des points sur des muscles du visage. Ils traquent le déplacement de ces points et sortent des chiffres qui évoluent dans le temps. Ce sont des séries temporelles. À partir de ces données temporelles on a une part de signal, donc une part *d'action units*. On en exploite une dizaine. En outre on a l'orientation de la tête et les angles du visage. À partir de ces données on veut essayer de savoir si le contenu a plu ou pas. Grâce au nombre de pics, on peut le savoir. Par exemple, si on a un grand nombre de pics d'une « *A.U.* » corrélée à de la joie, on a plus de chances que cela ait plu. Dès qu'on a une émotion, cela nous crée un pic, parce qu'on bouge nos « *A.U.* ».

Est-ce qu'on les bouge vraiment de manière réflexe ? Je sais que pour Paul Ekman, oui, mais qu'en penses-tu ?

La question est de savoir si on montre ou pas nos expressions. On aura du mal à détecter ça chez une personne qui ne réagit pas de manière expressive. Quand tu as une expression, ça te fait un pic, et après il faut savoir si ce pic-là est un vrai pic ou seulement un bruit.

Justement, qu'appelle-t-on le bruit ?

Cela correspond aux pics qui n'ont pas de sens. Si par exemple la personne se tourne, l'algorithme pense que ça génère un pic, mais ce pic n'est pas une vraie émotion. C'est juste un bruit, quelque chose auquel l'algorithme ne s'attendait pas, et du coup il croit que ça a un sens et nous dit que c'est un pic. Un signal qui a du bruit, c'est un signal qui a plein de points qui n'ont pas de sens.

Comment faites-vous alors pour savoir que c'est vraiment du bruit ? Faites-vous ça par un système de filtrage ?

Oui, c'est ça. En gros on considère qu'un pic est un vrai pic si l'orientation de la tête est bonne. Il faut déjà que la personne soit bien en train de regarder le contenu. Ensuite, il faut savoir si le pic détecté se distingue assez clairement ou pas. Nous regardons aussi si l'attention était bonne au moment du « pic ».

Je sais que tu travailles en binôme avec Lucas, avez-vous chacun votre spécialité ?

Je pars de ce qui a été fait à la base par Lucas, et je rends le code plus propre, en ajoutant quelques fonctionnalités. J'ajoute des tests. J'ai créé des choses tout en m'inspirant de ce qui existait. Quand on fait du code, il faut faire attention à ce que le code continue de fonctionner. Si le code est « sale » et qu'il n'est pas testé, il va finir par « crasher » et on ne saura pas pourquoi. Il y a aussi deux ou trois traitements que j'ai repris. En ce moment je suis en train de terminer le traitement des pics. Au tout début je faisais beaucoup d'études et de graphes, on les montrait à nos clients et partenaires, Éram et OpinionWay. Je faisais de la *data analysis* : des graphes et des stats pour donner du sens à nos données, et après je suis parti sur le nettoyage du code. Je travaille sur ces deux pôles.

Avec Lucas, on fait à peu près la même chose. Lui a plus travaillé sur la création de *features*, et moi sur l'intégration de ces *features* notamment avec des tests et des *pipelines* assez clairs de manière à ce que si plus tard on veut ajouter des éléments, on puisse le faire de manière simple. En faisant plein de tests, si on change quelque chose et que ça conduit à un bug, il faut qu'on ait un test qui plante. Grâce à cela, on sait qu'on vient de créer un bug. Plus on met de tests, et plus on est sûrs que lorsqu'on crée une fonctionnalité, elle ne va pas faire planter ce qui existe déjà. On ne veut pas développer un code à l'aveugle, puis le mettre « en prod » et se rendre compte ensuite que tout plante.

ANNEXE 13

Entretien avec Alexandre Macedo, Software Engineer chez Datakalab, 03/06/2019, Paris
17^{ème}

Diplômé de Polytechnique, Alexandre a travaillé pour Thales. Il travaille aujourd'hui chez Datakalab et optimise la plateforme propriétaire de collecte des vidéos des panélistes avec Lucas Lugão Guimarães.

Peux-tu me parler d'une problématique au quotidien sur l'amélioration de la plateforme ?

La semaine dernière, nous avons changé la manière dont les utilisateurs se connectaient à la plateforme. Il y avait un problème de configuration et on n'arrivait pas à faire des tests de connexion. Cela n'a pas fonctionné durant une journée. En ce moment, on tente d'utiliser un outil de *monitoring* pour aller plus vite dans le traitement.

Cela veut dire que pendant toute la durée du problème on ne pouvait pas recevoir les vidéos des panelistes ?

On demande à l'utilisateur de télécharger un fichier sur notre serveur. Les utilisateurs n'avaient pas accès à ces documents à cause de la politique de sécurité instaurée par Google Chrome. Il fallait faire des modifications dans le code pour que cela fonctionne. On a aussi dû changer des configurations pour les questions. Sur le *recorder* par exemple, on a plusieurs versions pour mieux identifier les résultats en les comparant. Nous avons énormément changé les « structures » de la plateforme, et nous voulons continuer à l'améliorer, notamment sur la partie *monitoring*.

« Monitoring », à quoi ça se réfère ?

Le *monitoring* permet de savoir si la plateforme fonctionne bien. Nous nous occupons à la fois de la partie « test » et de la partie *monitoring*. Lorsqu'on fait un changement, nous faisons un ensemble de tests pour valider cette version. La deuxième étape, c'est le *monitoring* : une fois que le changement est établi et que le client utilise la dernière version, on peut s'assurer que le *recorder* fonctionne bien.

Suite à l'entretien avec Lucas, j'ai compris qu'on pouvait envisager d'ouvrir la plateforme à tous les utilisateurs, comme le font Affectiva et Real Eyes. Penses-tu qu'on devrait l'ouvrir à tous les utilisateurs ?

Oui, à mon avis, ça peut être intéressant pour notre produit. Ensuite cela entraînerait aussi beaucoup de difficultés techniques. Comme nous utilisons la plateforme uniquement en interne pour l'instant, s'il y a un problème de sécurité par exemple, nous pouvons le régler facilement et rapidement. Je pense que pour les utilisateurs, c'est plus intéressant de pouvoir analyser directement leurs contenus, mais il y a beaucoup de travail à faire pour qu'on puisse y parvenir.

Quelle est l'importance de la plateforme pour Datakalab ?

La plateforme a une grande importance, dans la mesure où elle lie toutes les parties de l'équipe : celle qui travaille sur l'algorithme et son amélioration, celle qui travaille sur la plateforme, celle qui travaille sur le traitement de la donnée et l'équipe marketing/communication.

Lucas me parlait de « back-end » et de « front-end », peux-tu m'expliquer ces termes plus en détail ?

Le *front-end*, c'est la partie que tu vois sur ton écran, le *back-end*, c'est plutôt la partie « serveur ». Les deux communiquent, mais ils n'ont pas les mêmes droits. La partie *front-end* envoie des informations à la partie *back-end*. Par exemple : « Qui s'est connecté ? C'est telle personne ». À ce moment là, c'est la partie *back-end* qui autorise l'action ou non.

Pour l'instant, je me suis rendu compte qu'on pouvait uniquement « mesurer les émotions » exprimées, qui ne sont pas forcément les émotions ressenties. Penses-tu qu'on puisse réellement « mesurer les émotions » ?

On peut « mesurer l'émotion » aussi bien qu'une personne le peut. En ce moment on se concentre sur l'analyse de l'image. Si on enregistre une personne en train de parler, elle ne va pas forcément sourire, parce qu'elle s'adapte aussi à son interlocuteur. Mais quand tu es face à une machine, c'est plus simple de ne pas réagir et de faire les choses de manière instinctive. On est plus « authentique »

À quelles autres utilisations de la plateforme est-ce que tu penses ?

Il y a beaucoup de domaines où on peut utiliser notre technologie. Le domaine principal est celui de la publicité, mais on peut aussi penser aux films, aux séries. D'autant plus qu'aujourd'hui on est de plus en plus devant nos écrans. Nous avons pu adapter la plateforme pour pouvoir récolter des vidéos à partir des téléphones portables et non plus seulement des ordinateurs.

ANNEXE 14

Entretien avec Lucas Lugão Guimarães, Software Engineer chez Datakalab, 24/05/2019, Paris 17^{ème}

Diplômé de Polytechnique, Lucas a d'abord travaillé pour une filiale de Google avant de rejoindre Datakalab pour optimiser la plateforme propriétaire de collecte de vidéos des panélistes.

Comment as-tu connu Datakalab et pourquoi as-tu choisi cette start-up ?

J'ai connu Datakalab par un ami qui m'a parlé du projet. C'était l'un des projets de XProjet à Polytechnique. Il fallait traiter des vidéos en temps réel donc je me suis dit que j'allais essayer de mettre ça en place. Puis on a commencé à travailler ensemble et ça a bien marché. J'ai fait un stage à Datakalab durant ma deuxième année à Polytechnique. Pendant la troisième année j'ai fait un stage chez Verily Life Science, une filiale de Google. J'ai travaillé avec des interfaces biologiques. Je suis revenu un mois après chez Datakalab.

J'aime travailler dans une start-up. J'ai eu une expérience par le passé dans une grande entreprise et je peux comparer. J'aime le fait qu'on ait beaucoup de liberté. On peut choisir quelle technique on utilise. On discute de la manière dont on va procéder parce que la technique doit être alignée avec le business. Donc il y a quand même des contraintes, mais je trouve que dans le monde de la start-up on a plus de liberté qu'ailleurs.

Pourrais-tu décrire une journée type ?

On travaille sur différents projets en même temps. Par exemple en ce moment je travaille sur le *recorder*, c'est grâce à lui qu'on récupère les données. On a une liste de problèmes, *issues*, à résoudre. Dès mon arrivée à Datakalab, je regarde la liste pour prioriser et pour réfléchir à quelles solutions proposer. Je travaille en binôme avec Alexandre, qui me dit ensuite s'il est d'accord avec la solution proposée ou s'il a quelque chose à ajouter.

Une *feature* est une fonction. Par exemple, pour l'intégration on a dû changer le texte. Avant on ne pouvait pas changer le script. Donc on a créé la fonctionnalité de pouvoir changer le texte

selon chaque intégration, ce que les clients voulaient également. Les *features* et les « problèmes » sont très proches. Pour un problème il faut faire des corrections dans le code, et pour les *features*, il faut y rajouter quelque chose. Pour le *recorder*²⁶¹, il y a plusieurs aspects. Certains concernent le client, d'autres la plateforme que vous utilisez pour créer des expériences. On a le *back-end* pour récupérer les vidéos. Il y a plusieurs aspects, ce qui nous oblige à travailler sur différents points. Il faut être organisé et avoir des listes de problèmes. Nous traitons aussi des projets plus expérimentaux, on essaie de créer de nouvelles solutions d'enregistrement. J'essaie d'implémenter cela mais pour ce faire il n'y a pas de *guidelines* spécifiques. Donc pour résumer je dirais que dans une journée il y a plein de choses à résoudre, à créer, à expérimenter. Il faut aussi aider l'équipe communication à intégrer le script dans les questionnaires. Souvent, on ne travaille pas directement avec le client, c'est plutôt une entreprise qui travaille pour lui. Donc on n'intègre pas forcément avec la technologie du client. Dans ce genre de cas, la communication peut être compliquée.

Tu travailles en binôme avec Alexandre. Est-ce que chacun a sa spécialité ?

Là on essaye de rendre le *recorder* le plus robuste possible. On travaille avec les mêmes technologies. J'ai créé tout seul la plateforme donc je connais tous ses détails. Il est arrivé après moi, alors qu'on était dans un *refactoring*²⁶². C'est un terme technique pour qualifier la réécriture d'un code pour le mettre au propre. Quand Alexandre est arrivé, j'étais en train de faire le *refactoring* du *recorder*, donc il a pu aussi y assister. La plateforme a besoin de beaucoup de changements, notamment sur le design. Aujourd'hui on collecte de plus en plus de vidéos. Elle est adaptée pour gérer plusieurs vidéos, mais pas dans la même expérience. Donc on essaye de l'adapter pour qu'elle puisse gérer plus de vidéos en même temps.

Quand les panélistes regardent les vidéos sur la plateforme, ils ne sont pas dans les mêmes conditions que quand ils les voient dans la rue. Penses-tu que cela influence leur réaction, et que cela puisse venir biaiser l'étude ?

C'est une opinion, mais j'ai l'impression que ça dépend de la publicité. Comme j'ai fait des tests d'intégration, je regarde tous les contenus testés. Certains contenus sont tellement bien faits qu'on en oublie l'étude en cours. On ressent ce que la pub veut nous faire ressentir. Pour

²⁶¹ Plateforme propriétaire de Datakalab de collecte des vidéos des panélistes

²⁶² Refonte de l'architecture du logiciel

les images, les affiches, j'ai l'impression que les données sont moins intéressantes. Mais on voit aussi que plus les vidéos sont longues, plus il y a d'informations. Par exemple quand on analyse des *trailers*, j'ai l'impression que la personne a plus de temps pour réagir.

Peut-être que cela influence la « métrique d'attention » ?

Il y a une influence, mais elle est réduite, si l'affiche est bien faite et attire d'elle-même l'attention.

En quoi selon toi la plateforme rend-elle notre technologie plus « scalable » et en quoi nous permet-elle d'avoir un avantage par rapport à ceux qui ne l'ont pas ?

On est une jeune start-up, donc nous n'avons pas focalisé tous nos efforts dans une seule et même direction. Aujourd'hui notre propre plateforme est uniquement utilisée en interne. Elle n'est pas encore ouverte à la différence de celle de certains de nos concurrents. Par exemple, tout le monde peut utiliser la plateforme de Affectiva, et de Real Eyes. Il suffit de faire un compte et de s'inscrire pour tester n'importe quel contenu. Nous sommes encore dans la partie « test », où on essaye de comprendre ce qu'on peut faire avec les données et améliorer nos métriques. C'est compliqué de comparer deux technologies, parce que je ne connais pas la technologie utilisée en interne par les autres entreprises.

Peut-être que sur le long terme c'est ce que Datakalab voudra faire ?

Voilà, il s'agit d'une décision *business* qui a de l'impact sur beaucoup d'autres choses, il faut y réfléchir. On n'a pas encore la réponse, mais c'est l'un des objectifs possibles d'être une entreprise très *scalable* avec une plateforme ouverte, simple à utiliser, et accessible à n'importe qui. Il faudrait alors demander aux utilisateurs de se créer un compte et d'entrer leur carte bleue. C'est l'une de nos possibilités. Aujourd'hui, on travaille directement avec les clients, on est plutôt dans le *B to B*, pas dans le *B to C*.

J'imagine alors qu'on insisterait encore plus le côté « technologique », et moins sur le côté marketing. En faisant cela les utilisateurs pourraient voir directement l'analyse du contenu en temps réel.

Oui, sans doute.

Penses-tu qu'on puisse « mesurer les émotions » qui sont propres à l'homme avec de la donnée, qui elle, est propre à la machine ? Peut-on concilier ces deux domaines ?

Quand on parle de *machine learning*, on fait apprendre la machine. On apprend une notion qui est normalement humaine. Un logiciel qui fait de la traduction automatique (par exemple Google Traduction) ne comprend pas bien le langage naturel. Mais on essaie de faire apprendre la machine pour qu'elle puisse apprendre un texte, et le traduire. La machine peut faire relativement bien quelque chose relatif à l'humain. Si les machines peuvent reconnaître les choses comme les humains, alors je dirais qu'on peut concilier ces deux domaines.

Chez Datakalab, on analyse les vidéos pour voir quelles sont les émotions exprimées par les personnes. Pas les émotions vraiment ressenties. Normalement, quand on a une réaction involontaire, elle est directement liée à ce que l'on ressent. Mais par exemple dans le cas d'un psychopathe, il ne va pas montrer d'émotion, peut-être parce qu'il n'en ressent pas. Je pense qu'on peut « mesurer les émotions » qui sont montrées par le visage humain. Mais ce n'est pas la même question si tu me demandes : « Est-ce qu'une machine peut mesurer une émotion qui n'est pas montrée ? ». Là ça devient plus compliqué et on ne peut pas « mesurer » ce qui n'est pas mesurable. Le *machine learning* n'est pas magique, et il n'y a pas de données cachées dans le big data. Les machines peuvent stocker des données immenses. Mais je pense que si un être humain peut comprendre quelque chose, une machine va aussi pouvoir le faire plus ou moins bien. Notre discussion s'approche du test de Turing. Peut-on créer une machine qui peut se faire passer pour un être humain ? Xavier dit tout le temps en pitch de Datakalab : « on ne se souvient pas de toutes les émotions ressenties ». C'est le point clé de nos technologies. L'idée est de « mesurer ». Parce qu'il faut distinguer « mesurer » et « reconnaître ». Par exemple, un être humain peut savoir s'il fait chaud ou froid, mais avec un thermomètre, on peut « mesurer » à quel moment il fait chaud et à quel moment il fait froid. Si on peut « mesurer » on peut quantifier, et là on peut faire un autre type d'analyse.

À titre personnel, penses-tu qu'on exprime de la même façon une émotion en fonction de notre nationalité ?

Certains anthropologues disent que certaines émotions ne sont pas associées à la culture. C'est la même question pour la classification des couleurs. Certaines cultures n'ont pas de vert. Quel pourcentage de nos expressions faciales viennent de notre culture, et de notre côté humain ? C'est une question compliquée. Aujourd'hui le monde est connecté, j'ai l'impression que ça converge pour les pays connectés. Mais pour des personnes sur une île isolée, on ne sait pas comment ils vont réagir. Je suis Brésilien, et je sais qu'on ne réagit pas de la même manière. Par exemple, en France il est rare de voir les gens « rigoler en silence », alors qu'au Brésil c'est assez courant. Les différentes cultures ont différents gestes. J'ai l'impression qu'il y a des différences, aussi des similitudes.

Qu'imagines-tu pour le futur de Datakalab ?

Tout. On peut explorer différents chemins. On n'a pas encore fait de levée de fonds pour s'engager dans une direction spécifique²⁶³. J'ai l'impression que Xavier veut que Datakalab soit plus « tech » que conseil. On va avoir une grande équipe, il va falloir savoir où concentrer nos efforts.

²⁶³ Entretien du 24/05/2019

ANNEXE 15

Entretien avec Jessica Verres, Consultante-psychologue, Arnava, et Clément Delvigne, Consultant-Psychologue chez SBT Human(s) Matter, 05/07/2019, Paris 17^{ème}

Diplômée de Paris Descartes, Jessica est consultante Psychologue chez Arnava. Cette société de conseil et d'évaluation des compétences fait partie du groupe SBT. SBT, Unyck et Datakalab forment ensemble The Brain Place to Be, le premier centre dédié aux neurosciences cognitives. Dédié à l'échange pluridisciplinaire, le groupe transforme le rapport des entreprises et des marques aux neurosciences.

Clément a étudié à l'École de Psychologues Praticiens, PsychoPrat'. Lui aussi Consultant-Psychologue chez SBT Human(s) Matter, il travaille sur des évaluations, des bilans de personnalité et de compétences cognitives.

Pouvez-vous me présenter votre parcours, vos postes respectifs chez SBT et le rôle de la psychologie dans votre parcours ?

Jessica

Je suis consultante-psychologue. Je m'occupe des *assessments*, c'est-à-dire de l'évaluation des individus par une mise en situation. On leur donne des études de cas, on fait des jeux de rôle et des tests psychométriques comme des tests d'aptitudes, de personnalité pour ensuite décrire leur profil. J'analyse ce type de test pour les candidats, je fais le *débriefing* avec eux et j'écris une synthèse. J'ai suivi une formation scientifique à Paris Descartes. En me concentrant sur la psychologie du travail, je n'ai pas tant étudié le psychisme de l'individu que la façon dont il interagit dans son environnement, ce qu'on appelle la psychologie sociale. J'étudie comment un individu réagit dans son milieu professionnel. Comment gérer les changements organisationnels grâce aux recherches sur l'humain et ses relations à autrui.

Clément

Contrairement à Jessica, mon parcours est plus axé « psychologie intégrative ». Jessica a un ancrage très « thérapie cognitive/comportementale ». J'ai fait l'École de Psychologues Praticiens, PsychoPrat'. On a un ancrage assez psychanalytique et intégratif avec 4 ans de clinique, du RH et de la psychologie d'entreprise en dernière année. J'ai travaillé en insertion

socio-professionnelle. J'ai fait de l'accompagnement de salariés et j'ai travaillé sur tout ce qui tourne autour des risques psycho-sociaux au siège de la CPM. Je travaillais à la fois sur des études quantitatives et sur de l'accompagnement individuel. J'ai ensuite travaillé chez la SNCF, il fallait sélectionner des profils d'ingénieurs. J'y ai fait de *l'assessment*, pour les futurs directeurs d'établissement. Je suis arrivé chez SBT début 2018 pour faire des évaluations, des bilans de personnalité et de compétences cognitives. Je travaille sur des projets comme la refonte du dispositif d'évaluation pour entrer à l'EMLyon. Dans mon travail, il n'y a pas de livrable direct sur les émotions, mais je les utilise au quotidien.

Quel est l'impact des émotions dans notre quotidien ?

Jessica

Les émotions sont à l'origine de tout. On retient mieux ce qui déclenche une émotion. Elles nous donnent l'impulsion pour agir. Les émotions créent aussi des biais. Quand on est concentré, quand on a peur, ou qu'on est stressé, on peut manquer des informations. Plus les émotions sont fortes, et plus notre attention est mobilisée.

Que pensez-vous des études déclaratives pour « mesurer » la performance des contenus si on les compare aux études émotionnelles de Datakalab avec le « facial coding » ?

Clément

Si on cherche à comprendre les représentations de la personne face à un contenu, et ce que cela signifie pour elle, le déclaratif est indispensable. Si on cherche à objectiver une réaction face à un stimulus, dans ce cas-là, on peut utiliser la solution Datakalab. Pour analyser les réactions du plus grand nombre face à un discours par exemple, le *facial coding* me semble particulièrement pertinent, puisqu'on récolte beaucoup de données d'un seul coup.

Jessica

Oui, mais surtout, dans la société, la « désirabilité sociale » intervient tout le temps. Si on prend l'exemple des questionnaires déclaratifs, quand un individu fait un choix, il se demande si son choix est désirable socialement. Il se demande si on ne va pas le juger d'avoir répondu d'une certaine manière, même s'il n'y a personne. C'est ce qu'on essaye de combattre en psychologie quand on construit un questionnaire. Les individus veulent plaire aux autres et ne vont pas forcément dire la vérité. C'est en ce sens que l'étude des émotions me semble moins biaisée

que le déclaratif. Quand on répond en déclaratif, on fait attention à ce qu'on dit, on tempère. D'autres personnes se connaissent mal, et répondent de manière incorrecte sans le vouloir.

Pensez-vous que les études de Datakalab puissent vraiment « mesurer les émotions », qui sont propres à l'homme, avec des algorithmes, propres à la machine ? Peut-t-on appréhender la charge symbolique des émotions pour les individus ?

Clément

Ici, vous traitez des données informatiques plus que des émotions, ce qui ne vous oblige pas à avoir dans votre équipe des spécialistes de l'étude émotionnelle. Quand on observe les mêmes réactions sur un grand nombre de personnes, on peut déduire ce qu'elles ressentent à un moment donné, mais je pense que l'émotion ne se limite pas à ça. Pour comprendre une émotion, il faut aussi savoir au niveau singulier ce qu'elle évoque et ce qu'elle représente pour une personne, et si c'était feint, sincère ou non. Autant de données que la machine ne peut pas prendre en compte. Dans certains discours politiques, le politicien a un côté du visage qui exprime une émotion, et l'autre côté en exprime une autre. Si on filme son visage pour détecter ses émotions, que fait la machine et que va-t-on en déduire ? Ici, on a besoin de l'intelligence humaine pour guider la machine et aussi pour interpréter en fonction du contexte.

Selon Paul Ekman, les émotions sont universelles : un Américain, un Japonais, un Européen, exprimeront de la même façon les sept émotions universelles. Qu'en pensez-vous ?

Clément

Certaines émotions sont universelles puisqu'elles font partie d'un système de survie. Le ressenti de la peur par exemple est forcément universel puisqu'autrement, on ne fait pas le mort, on ne combat pas, ou ne fuit pas quand il faut le faire. Certaines expressions des émotions sont universelles, mais d'autres dépendent aussi de la culture. On ne va pas forcément exprimer notre dégoût de la même manière ici ou en Chine par exemple.

Jessica

Je pense que tout le monde ressent bien des émotions mais ne les exprime pas forcément de la même manière. Les Japonais, par exemple, ont beaucoup de retenue dans l'expression des émotions. On peut dire que c'est universel parce que tout le monde a des émotions, mais il

existe aussi des facteurs culturels : on ne les exprime pas de la même manière, et surtout, pas au même moment. Des sujets vont plus toucher dans une culture que dans une autre, je pense par exemple au respect, qui est très important pour les Japonais. On a tous les mêmes émotions, mais en fonction de notre culture, on ne les exprime pas avec la même intensité. Quand on regarde les pubs japonaises, on voit bien qu'elles ont un côté plus extraordinaire que les publicités européennes. Pour autant, elles restent de simples pubs pour les Japonais.

Comme on agrège les données, on ne s'intéresse pas à la charge symbolique de chacun, mais on se concentre sur les vérités consommateurs dans leur ensemble. Pensez-vous que les « micro-expressions » peuvent quand même appréhender la charge symbolique des émotions ?

Jessica

C'est tout le paradoxe de la psychologie. On suit une loi normale, une courbe de Gauss, une distribution symétrique autour d'une seule valeur, qui est à la fois le mode, la médiane et la moyenne. Une publicité drôle fera sourire une majorité de personnes. Même si les émotions sont individuelles, une majorité des panélistes vont avoir le même type d'émotions au même moment. On apprend avec les autres et en ce sens on est les autres. On sait que ça ne fonctionne pas pour absolument tout le monde, mais on a déjà plus de chances de faire réagir.

Selon vous, quelles sont les possibilités pour le produit Datakalab ? Quelles autres applications de notre technologie imaginez-vous ?

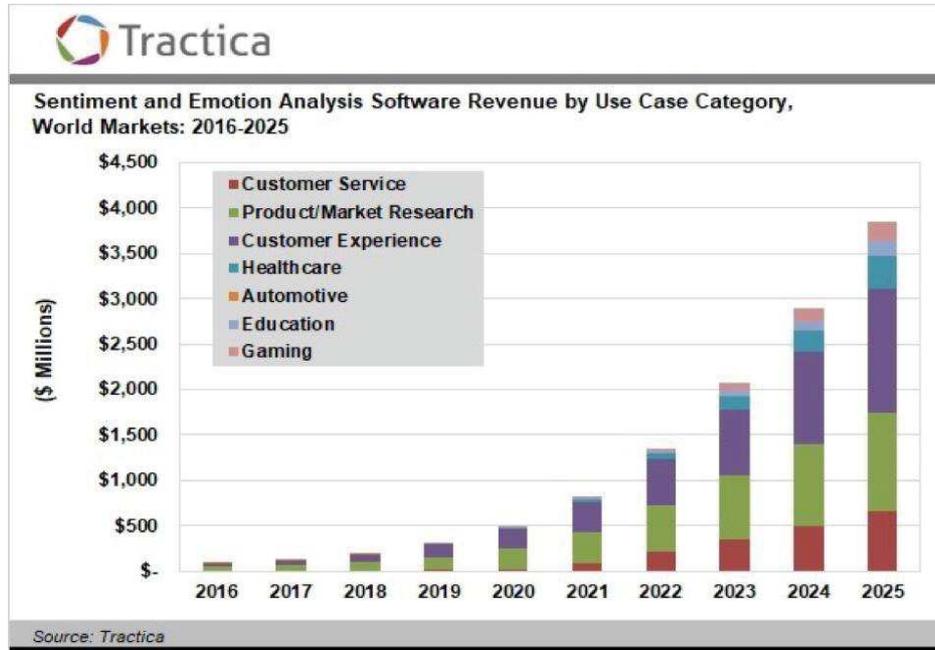
Clément

Spontanément, je pense à l'expression des émotions sur le spectre autistique. On dit que les autistes n'ont pas d'expression émotionnelle particulière, qu'ils ne captent pas les émotions. Le vécu d'une émotion va être totalement différent d'une personne à l'autre. Par exemple, la culpabilité qui est plus un sentiment qu'une émotion — mais qui s'en rapproche quand-même — va être vécue différemment en fonction des personnes. Dans le judiciaire, la solution Datakalab peut être utile dans les expertises où on a besoin d'éléments catégoriques, de pouvoir dire si la personne est psychopathe ou non. Le psychopathe a du mal à feindre les émotions. On pourrait utiliser le *facial coding* au moment de l'analyse psychologique notamment.

Jessica

Oui, il y a vraiment beaucoup de possibilités. De mon côté, je suis attachée à l'écologie, alors j'aimerais qu'on puisse sensibiliser la population à ce sujet en véhiculant plus d'émotions. L'intelligence émotionnelle se développe aussi en entreprise. C'est le fait d'identifier ses propres émotions, et celles des autres et de savoir y répondre. Peut-être pourrait-on apprendre à s'intéresser aux émotions, et faire plus de pédagogie pour comprendre quelles actions déclenchent quelles émotions.

Étude Tractica sur le marché de la donnée émotionnelle : *Emotion Recognition and Sentiment Analysis Market to Reach \$3.8 Billion by 2025*²⁶⁴

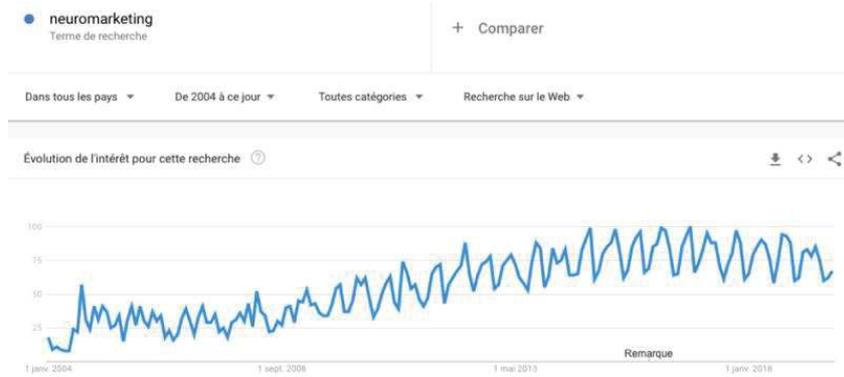


²⁶⁴ « Je traduis : Le marché de la reconnaissance des émotions et des sentiments devrait atteindre les 3,8 milliards de dollars en 2025 »

ANNEXE 17

Google Trends, recherches dans le monde de 2004 au 05/09/2019

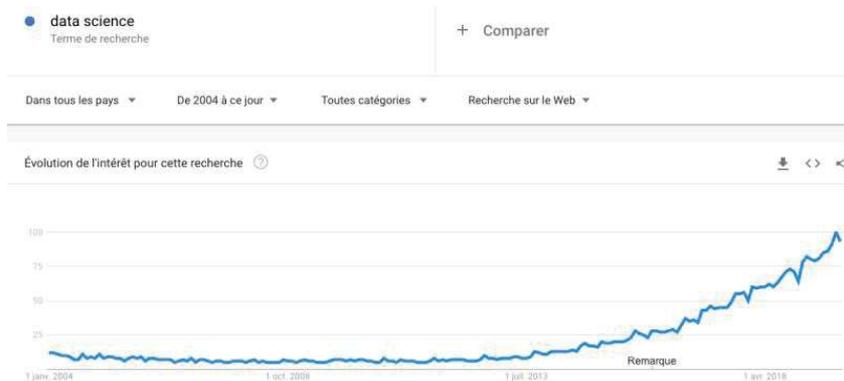
« neuromarketing »



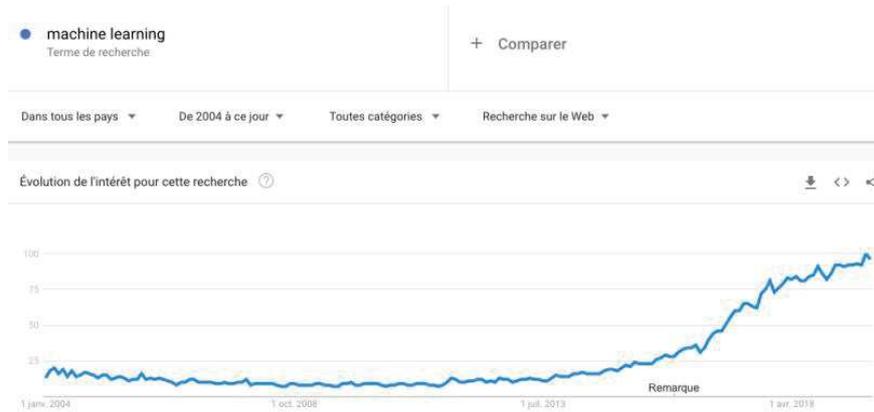
« big data »



« data science »



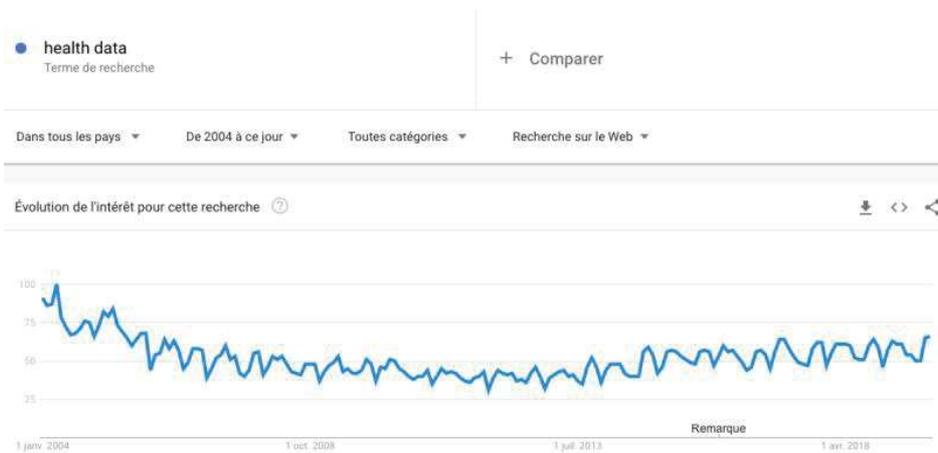
« machine learning »



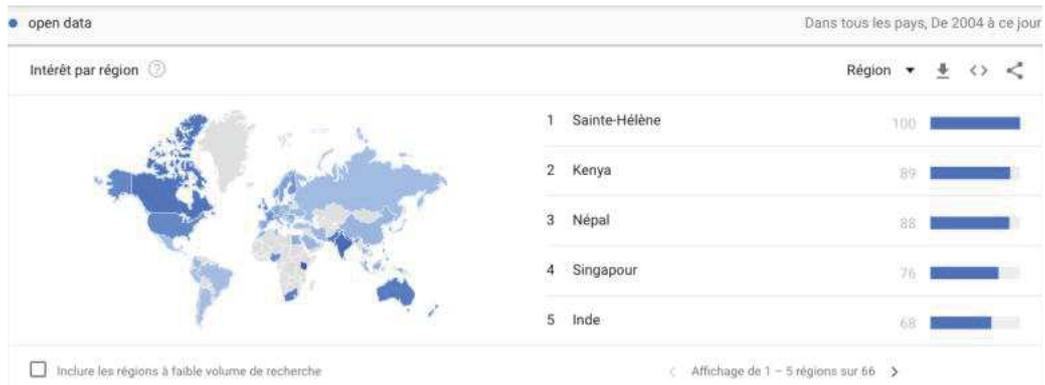
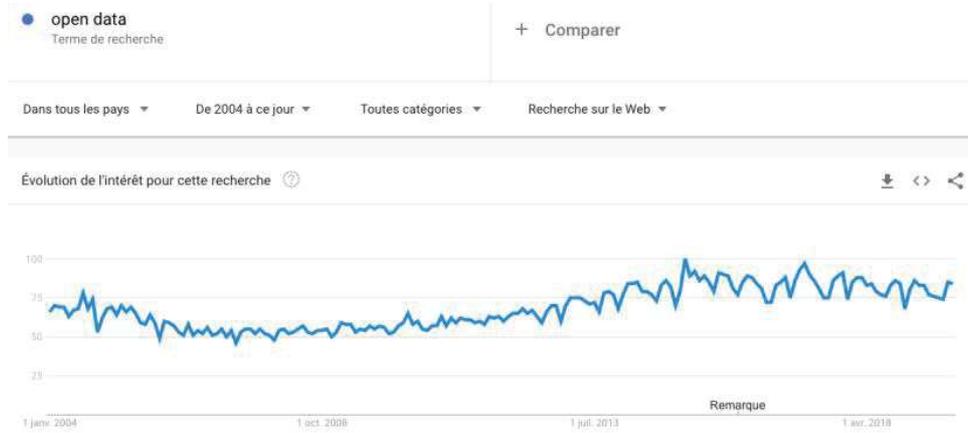
« data scientist »



« health data »



« open data »



« Innovation » et « performance » par pays



ANNEXE 18

Corpus d'étude - Captures d'écran du site internet de Datakalab au 01/06/2019

Page « ACCUEIL » de Datakalab, « disponible à l'adresse : <https://www.datakalab.com/> »



Page « LE LAB », « disponible à l'adresse : <https://www.datakalab.com/lab> »



NOTRE MISSION

Datakalab est une start-up dans le Brain Tech qui détecte les émotions des consommateurs à l'aide d'outils issus des neurosciences pour nourrir les stratégies de transformation et d'innovation des marques.

Datakalab croise les data transactionnelles et digitales et les éclaire avec les data émotionnelles et chaudes «**Feel data®**» pour optimiser la connaissance client et mieux piloter les stratégies d'innovation des marques et des entreprises.

“*Nous voulons
donner
un visage humain
aux data*”



_NOTRE VISION

Face à l'accélération du digital et de la robotisation, il est important de redonner un visage humain aux data.

Les data transactionnelles, digitales ou relationnelles ne permettent pas suffisamment de comprendre les motivations et les actions des clients et des individus. Aujourd'hui, 90% des actions et des décisions d'achat sont conduites par les émotions.

Les stratégies de transformation et d'innovation pérennes sont «client centric» et commencent avec une compréhension intime des clients :
«la client intelligence».

_NOTRE PROMESSE

“*La rencontre
du marketing,
des neurosciences
et de la tech*”

FEEL DATA®

Datakalab mesure les émotions des clients et les transforme en données émotionnelles et chaudes : les **feel data®**.

DATA CROSSING

Datakalab croise les data transactionnelles, digitales ou relationnelles pour les enrichir avec les **feel data®** et identifier des corrélations vertueuses et créatrices de valeur.

DATA TELLING

Datakalab visualise les datas grâce à sa plateforme et interprète les résultats fort de son expertise marketing et neuroscientifique.

DATA PRIVACY

Datakalab est privacy by design. Les données sont anonymes et agrégées. Le consentement des clients est géré par Datakalab.

Rubrique « Nos outils »

NB : L'outil « bracelet connecté » a été supprimé du site le 03/06/2019

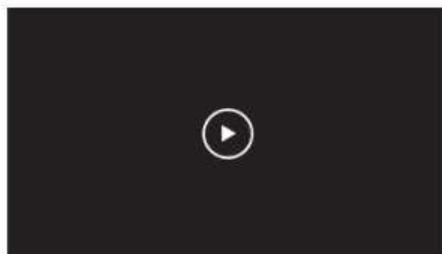
_NOS OUTILS



Le Bracelet Connecté

Le bracelet est muni de capteurs et d'un accéléromètre pour mesurer les signes physiologiques de l'émotion en temps réel :

- Microsudation (mesure électrodermale)
- Fréquence cardiaque
- Variabilité de la fréquence cardiaque
- Température extra-cutanée



Le Facial Coding

Créé à partir des travaux de Paul Ekman en 1970, le Facial Coding permet de détecter les micro-expressions du visage et d'identifier les 7 émotions universelles : la joie, la surprise, la peur, la tristesse, la colère, le dégoût, le mépris. La technologie d'Eye Tracking permet de trouver les zones qui attirent le regard, les zones d'intérêt pour l'utilisateur.

_NOS PRODUITS



DATAKALAB CONTENT

mesure les pics d'émotions pour un contenu de marque grâce au Facial coding et à l'Eye tracking.



DATAKALAB WEBSITE

mesure les niveaux d'attention et les émotions pendant le parcours clients. Ces données émotionnelles sont croisées avec les données transactionnelles collectées par notre partenaire IBM Tealeaf.



DATAKALAB EVENT

mesure les pics d'émotions sur une période donnée lors d'une conférence, d'un événement ou d'un débat grâce au Facial coding.

_NOTRE ENGAGEMENT

Parce que **Datakalab** place l'éthique et la protection des données à caractère personnel au cœur de ses engagements, nous souhaitons utiliser vos données de façon transparente afin de bâtir une relation de confiance avec vous.

Datakalab prend en compte les principes de protection des données, dès la conception des projets (Privacy by design).

En conséquence, ne sont collectées que des informations pertinentes, adéquates et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées. Nous ne disposons, et ne voulons pas disposer, des données nominatives ou d'identification vous concernant, comme votre nom, votre adresse, votre email ou votre numéro de téléphone.

Pour plus d'informations, vous pouvez consulter notre **[Politique de confidentialité](#)**



Hello Bank Timeline emotionnelle

Grâce au facial coding, nous avons pu mesurer l'attention, l'engagement émotionnel et l'adhésion lors du visionnage de 2 spots promotionnels. Nous avons corrélé les datas émotionnelles ainsi recueillies aux datas déclaratives récoltées après écoute. Cela nous a permis de mesurer les émotions ressenties par les panélistes lors des différents visionnages.



Sixième Son Timeline emotionnelle

Grâce au facial coding, nous avons pu mesurer l'attention et l'engagement émotionnel lors de 6 pantones vocaux. Nous avons corrélé les datas émotionnelles ainsi recueillies aux datas déclaratives récoltées après écoute. Cela nous a permis de mesurer les émotions ressenties par les panélistes lors des différentes écoutes.



Dassault Systèmes Timeline Emotionnelle

Grâce au facial coding, nous avons pu mesurer l'attention, l'engagement émotionnel et les émotions ressenties lors du visionnage des différents contenus. In fine, nous avons déterminé le brand content avec la meilleure émergence émotionnelle et la meilleure mémorisation.



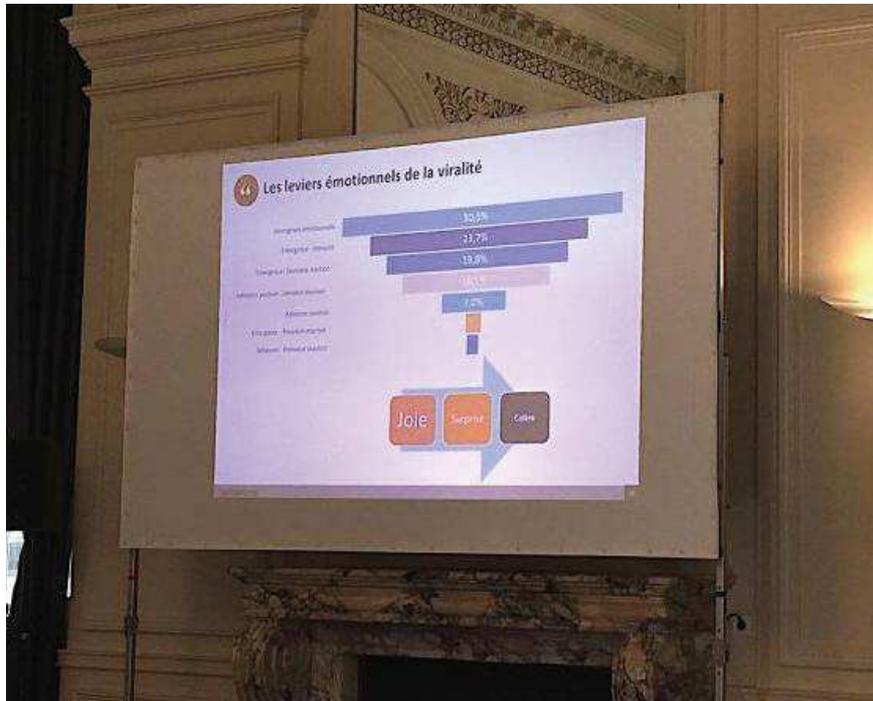
SNCF Gare & Connexions Cartographie Emotionnelle

Faisant suite à nos celle de 2017, nous avons réalisé une cartographie émotionnelle de la gare SNCF grâce au bracelet connecté. Nous avons mesuré la qualité du moment passé en gare et dans les différents espaces pour ainsi définir les endroits qui génèrent le plus de bien-être.

ANNEXE 19

Photos de la conférence « OpinionWay x Datakalab », Printemps des Études, 04/2019

OpinionWay présente l'importance des émotions dans la viralité



Support de présentation PowerPoint de Datakalab 2019

— Qui sommes-nous ?



- Une start-up dans la Brain Tech
- Nous mesurons les émotions à travers des outils issus des neurosciences.
- Nous optimisons la connaissance client et nourissons les stratégies de transformation et d'innovation des marques.



Copyright Datakalab 2019

— La réalité d'aujourd'hui



- Aujourd'hui, le cerveau humain est submergé d'informations.
- Si on considère la publicité au sens large, incluant le sponsoring, le placement de produit dans les films, les enseignes et devantures de magasins, les publicités sur distributeurs de boissons, les displays et autre présentoirs dans les magasins, les logos bien identifiables sur vêtements etc...
- Nous sommes exposés à pas moins de 15 000 stimuli commerciaux par jour.
- Face au bombardement des informations traitées par jour, les individus sont atones, en majorité exprimant un état neutre (80% du temps).



Copyright Datakalab 2019

— Les enjeux



- Dans cette société de sursollicitation, il est primordial de s'assurer que les contenus :

- Captent l'attention,
- Créent de l'émergence, garantie d'une meilleure mémorisation,
- Génèrent de l'adhésion afin de s'assurer que cette émotion soit au service de la marque.

Da

Copyright DataLab 2019

— L'inconscient domine notre expérience

- **90%** des stimuli qui nous impactent dans une journée sont traités par l'inconscient*
- Cela signifie que les études quanti ou quali qui s'appuient sur des réponses déclaratives ne mesurent que partiellement ce que les consommateurs perçoivent
- Les études déclaratives ne suffisent pas à « mesurer » objectivement les ressentis des consommateurs

*Gerald M. Edelman, Prix Nobel de la Médecine, 1972

Da

Copyright DataLab 2019

— Datakalab Content



Datakalab Content mesure les pics d'émotions sur une période donnée lors :

- d'une publicité (print, affichage ou TV),
- d'un contenu de marque sur les réseaux sociaux,
- d'un pre-roll/mid roll,

À travers nos outils de Facial coding et de predictive eye tracking.

Da

Copyright Datakalab 2019

— Les normes



- 80 annonces presse et affichage
- 55 vidéos et films
- Une vingtaine de conférences
- Plus de 30 000 vidéos de visages
- Plus de 10 millions d'images d'expressions faciales collectées et analysées

Da

Copyright Datakalab 2019

— Les métriques

ATTENTION

% de personnes ayant été attentives durant tout le spot

ÉMERGENCE

% de personnes ayant ressenti au moins une émotion

ADHÉSION

% de personnes ayant ressenti au moins une émotion positive

Da

Copyright DataLab 2019

— Nos convictions



- Les émotions constituent un vecteur d'efficacité non biaisé.
- Ces indicateurs permettent de mesurer le ressenti d'un public seconde par seconde pendant une conférence, un événement ou un débat.
- À la différence des questionnaires de satisfaction, nos mesures permettent d'objectiver le niveau d'attention du public et le niveau d'émergence tout au long de la conférence.
- Cela permet d'identifier :
 - Les sujets qui captent l'attention du public,
 - Les sujets qui suscitent des émotions,
 - Les sujets qui suscitent de l'adhésion.

Da

Copyright DataLab 2019

— Les problématiques

- Mesurer le niveau d'attention et l'émergence seconde par seconde.
- Objectiver le ressenti d'une audience versus le questionnaire de satisfaction.
- Corréler les données émotionnelles avec les données déclaratives.

Da

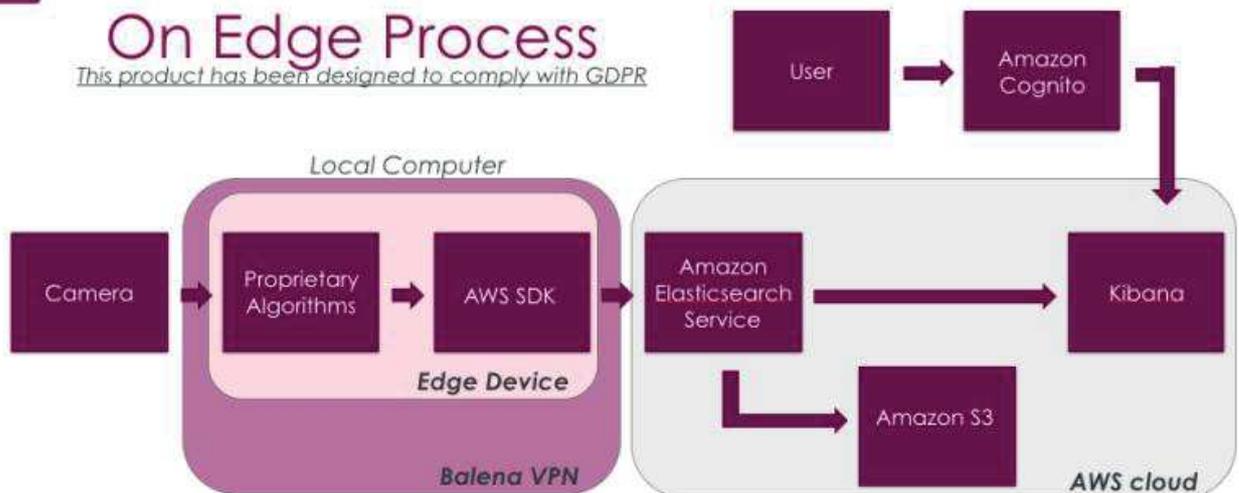
Copyright Datakalab 2019

Démarche de respect de la loi de la protection de la donnée personnelle, présentation de octobre 2019

11

On Edge Process

This product has been designed to comply with GDPR



Security :

Video :

- Frames are processed in RAM and deleted after 100 ms.
- Frames are not saved nor broadcasted.

Data :

- Data is sent to servers by batch of 10 people to protect individual privacy.

© Copyright Datakalab 2019

Sélection de mails envoyés aux 2000 abonnés à la « newsletter » Datakalab en 2019

1. Mail du jeudi 18/04/2019 sur la présentation des algorithmes Datakalab



La semaine dernière, nous vous envoyions notre premier papier de recherche sur l'estimation de la géométrie du visage par nos algorithmes. Aujourd'hui nous vous en faisons la démonstration.

Ici, le visage de Tanguy Massart est traité par nos algorithmes. Le réseau de neurones profond a appris à focaliser progressivement ses traitements sur les zones les plus importantes du visage.

Une centaine de points sont analysés sur le visage pour détecter l'orientation de la tête, le clignement des yeux, le mouvement de la pupille ainsi que les émotions et leur intensité.

À propos de Datakalab

Datakalab est une start up dans la Brain Tech qui mesure les émotions et le comportement des clients sur les différents points de contacts de la marque (pub, contenus de marque, affiches, sites internet, événements, magasins...) à travers d'outils issus des neurosciences et de l'IA. Datakalab a développé ses propres algorithmes et ses plateformes et a analysé plus de 10 millions d'images d'expressions faciales collectées lors de mesures d'affiches, d'annonces presse, de vidéos et films, de conférences. En tout, plusieurs milliards de data émotionnelles (Feel data©). Elle met son expertise au service de clients tels que TF1, France TV Publicité, M6, Crédit Mutuel Arkéa, Novotel, AccorHotels Arena, Exterion Media, LVMH, Disney, Mr Bricolage, SNCF Gares & Connexions, BNP Paribas, BPI France, Disney, Unibail Rodamco Westfield, Groupon, Groupe Figaro...



2. Mail du 23/07/2019 sur la sélection du papier de recherche Datakalab à ICCV 2019, l'une des deux plus grandes conférences en « computer vision » dans le monde



One of Datakalab research paper is selected at ICCV 2019. ICCV is one of the most important venue in computer vision and deep learning worldwide.

This Paper was written by two of our researchers Arnaud Dapogny, Machine Learning Scientist, and Kevin Bailly, Head of Research and Teacher at Sorbonne University along with Matthieu Cord, who also took part in this work.



This paper showcases a method to estimate facial geometry, which unifies the two predominant approaches for the alignment of feature points by deep learning.

This approach allows to finely estimate the geometry of the face by benefiting from weakly annotations via an optimized architecture and weakly supervised learning techniques.

Discover the draft of the paper below:

<https://arxiv.org/pdf/1904.02549.pdf>

Datakalab research strategy is to develop state of the art deep learning algorithms to better understand human face and to empower it's network of partner with a brand new generation of data.

Da



ANNEXE 22

Sélection de prise de parole Facebook de 2016 à 2019

Articles de 2016 postés dans la rubrique « Ils parlent de nous »

Ils parlent de nous



Prise de parole du 05/06/2019 sur les attributs que le nouvel algorithme sait détecter



575 vues

Frank Tapiro

4 juin

My personal demo of DatakaLab algorithm. 100 points are detected on my face. Once combined, the action units at the top right allow us to measure people's expres...

[Afficher la suite](#)

[Voir la traduction](#)

Prise de parole du 02/07/2019 sur la participation à la finale du challenge organisé par la Fevad et par KPMG sur les start-up innovantes dans le e-commerce



DatakaLab

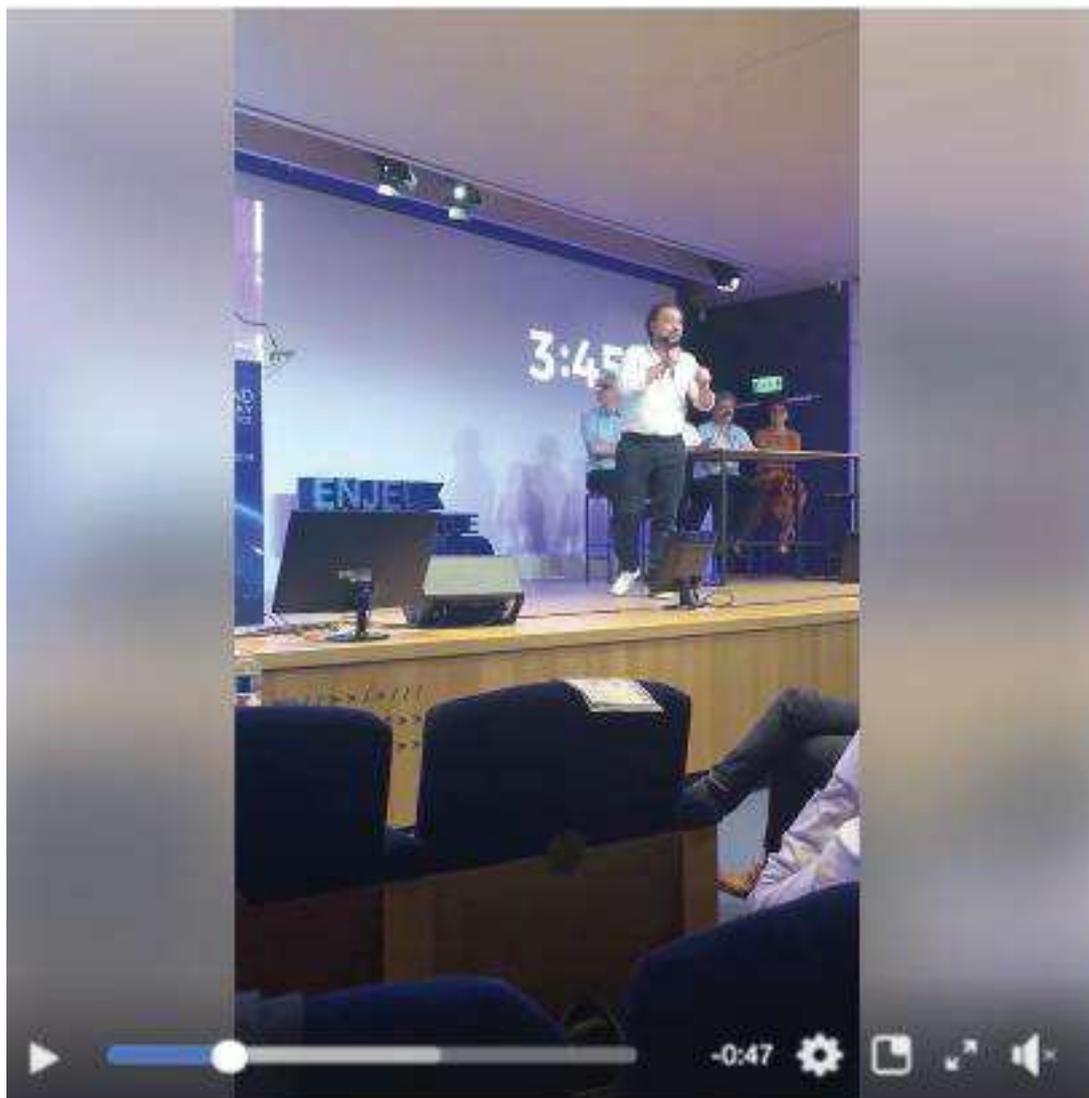
Publié par Claire Latourne · 2 juillet ·

Datakalab reached the finals of the "Start-me Up!" Challenge. The competition was organized by Fevad : Fédération e-commerce et vente à distance and KPMG to select European startup that boost e-commerce. We are glad to be ranked in the top 5.

Discover Frank Tapiro's pitch below 😊

#Ai #deeplearning #facialcoding #emotions #data

[Voir la traduction](#)



1 554

Personnes touchées

176

Interactions

[Booster la publication](#)

Prise de parole du 11/07/2019 sur le partenariat Datakalab x Médiamétrie



DataLab

Publié par Claire Latourre [?] · 11 juillet ·



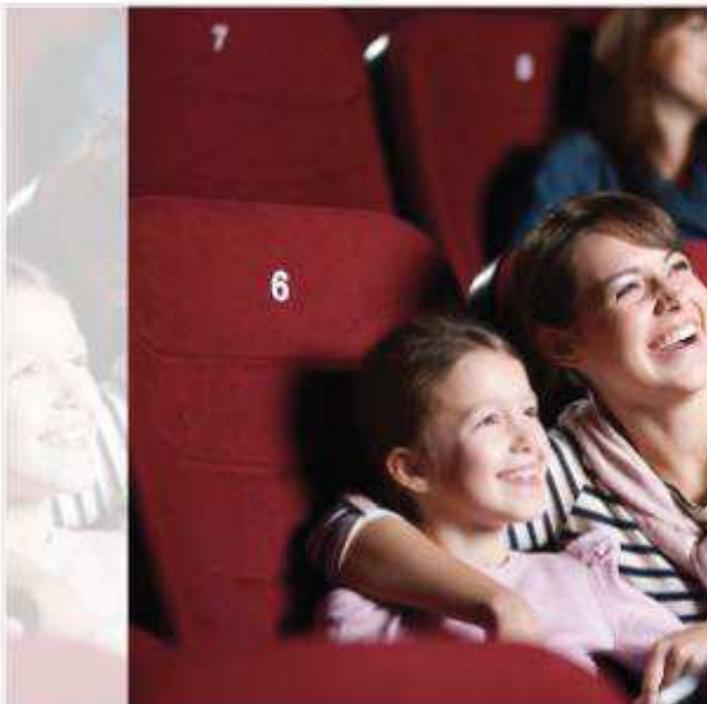
Proud of this collaboration with Médiamétrie . Médiamétrie has chosen Datakalab's technology to analyse audience's reactions to contents.

Discover the press release below:

<https://www.mediametrie.fr/.../lemotion-des-spectateurs-au-co.....>

[Afficher la suite](#)

[Voir la traduction](#)



L'émotion des spectateurs au cœur de la mesure cinéma de Médiamétrie - Médiamétrie



L'émotion des spectateurs au mesure cinéma de Médiamét

259

Personnes touchées

43

Interactions

[Booster la publication](#)