



MBA
MCI
LET'S GO DIGITAL

Denis LEFEBVRE
THÈSE PRO 2021

IA MARKETING DURABLE

Thèse Professionnelle de **Denis LEFEBVRE**

IA MARKETING DURABLE

Comment exploiter l'Intelligence Artificielle de façon durable pour sur-optimiser votre stratégie marketing

Tuteur de thèse : Thierry PIRES
Date de remise : 03 octobre 2021

L'auteur



Je suis **Denis Lefebvre**. Je suis le **co-fondateur de Actialuna**, qui développe depuis plus de 10 ans des logiciels innovants pour le livre.

Notre dernier projet est un outil d'aide à la traduction de bandes dessinées, commercialisé auprès des éditeurs, des traducteurs et des studios de lettrage en France et à l'étranger.

Notre solution Geo Comix exploite une Intelligence Artificielle que nous développons en partenariat avec l'Université de La Rochelle. Pour en savoir plus : www.geocomix.com

J'ai souhaité compléter mon expertise en marketing digital en rejoignant la **promo 2020-2021 en Part-Time du MBA MCI à l'Institut Léonard de Vinci**, élu meilleur Master depuis 10 ans dans sa catégorie (classement Eduniversal ¹).

J'ai également rejoint au cours de ce cursus le **collectif ia4marketing.fr**, et le groupe de travail sur l'IA durable au sein de **l'Institut du Numérique Responsable** (cf. p.7).

N'hésitez pas à **me suivre ou me contacter sur :**



@lefden



/denislefebvre



lefden@me.com



ia4marketing.fr

¹ <https://www.meilleurs-masters.com/master-e-business/pole-universitaire-leonard-de-vinci-mba-specialise-marketing-et-commerce-sur-internet-mci.html>

Introduction

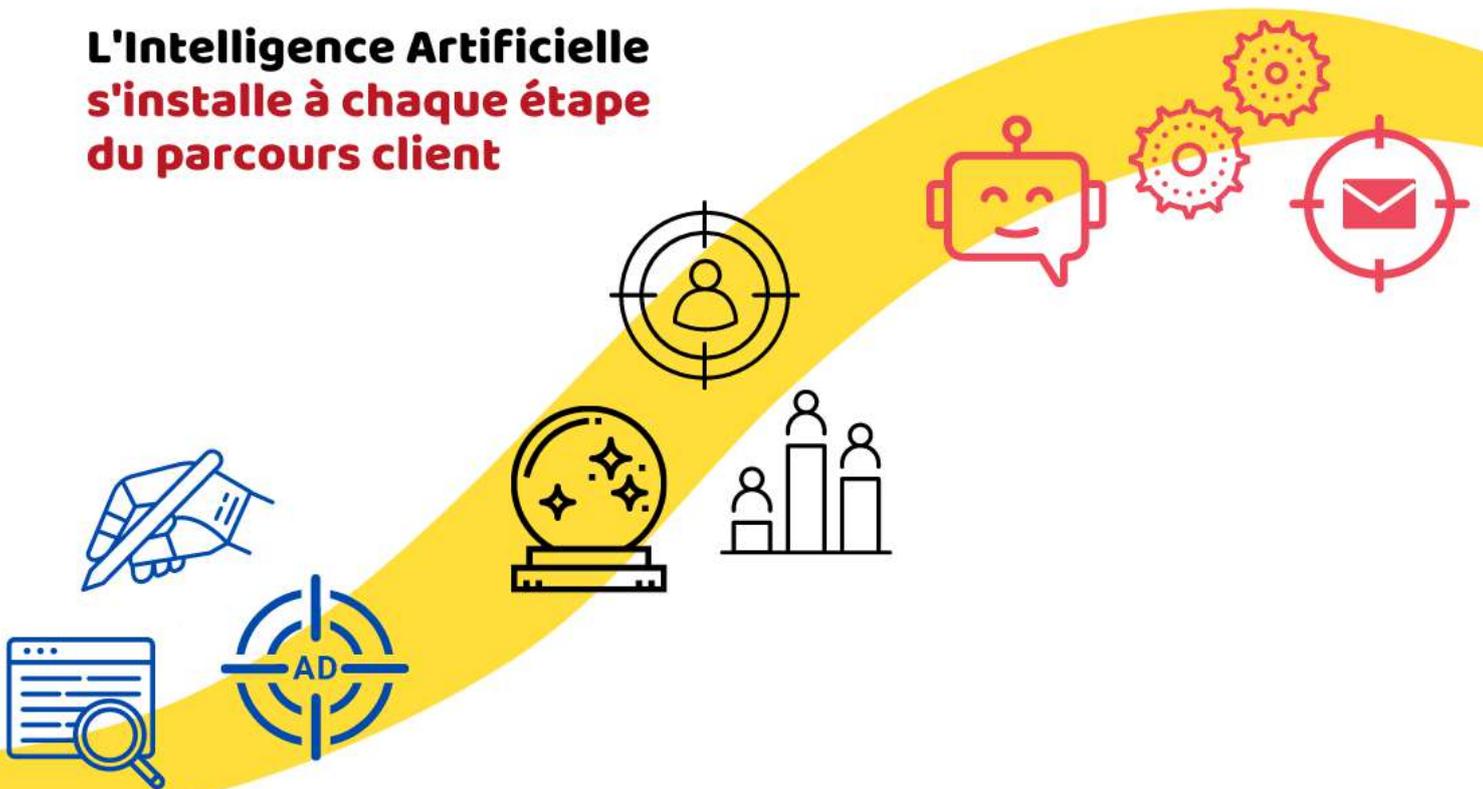
Cette thèse professionnelle est écrite dans le cadre de mon MBA en marketing digital et e-commerce à l'Institut Léonard de Vinci.

Cette formation complète aborde tous les aspects du marketing digital, avec à sa base ses fondamentaux, mais aussi et surtout ses évolutions permises par l'émergence de nouvelles technologies : Réalité augmentée, No Code, Blockchain...

Parmi celles-ci, il en est une des plus prometteuses : **l'Intelligence Artificielle**. Elle **offre de très nombreuses opportunités au marketeur** pour analyser toujours plus finement le comportement et les attentes de ses cibles, et optimiser leur expérience tout au long du parcours d'achat.

À l'heure où nous entrons toujours plus loin dans l'ère de la Big Data, avec un volume de données aujourd'hui 25 fois plus important qu'il y a 10 ans, et qui devrait encore être multiplié par 45 d'ici 2035, il devient en effet indispensable d'**automatiser sa connaissance clients et ses leviers marketing**.

**L'Intelligence Artificielle
s'installe à chaque étape
du parcours client**



L'IA est ainsi devenue une réalité tangible et incontournable pour tout Manager du marketing digital : 84% des professionnels du métier déclarent déjà utiliser en 2020 des outils enrichis à l'IA, soit +186% en 2 ans ². **Le marketeur « augmenté » fait de l'IA son allié** pour améliorer son référencement, son ciblage, son parcours de navigation... et bien plus encore. Voici donc **l'ère de « l'IA marketing »**.

Mais une autre étude sur les tendances 2021 du marketing digital a également retenu toute mon attention : celle de Hubspot, qui place l'IA en bonne position dans son top 10, tout en allouant à son premier rang... **le marketing éco-responsable** ³.

À l'heure de l'intensification des catastrophes naturelles un peu partout dans le monde (incendies, inondations...), la **nécessité d'une « révolution verte »** devient effectivement une réalité de plus en plus pressante. La prise de conscience des consommateurs est accueillie et relayée par les marques, dans leurs pratiques et leur communication.

Dès lors, **la problématique se pose : l'IA marketing est-elle en phase avec le marketing éco-responsable, ou bien risque-t-elle d'alourdir la facture environnementale par l'usage intensif des capacités de calcul de l'IA, et des serveurs qui les hébergent ? Les marques peuvent-elles adopter l'Intelligence Artificielle pour optimiser leur stratégie marketing tout en garantissant leur éco-responsabilité ?**

Alors que la transformation digitale des organisations est amplifiée par la crise de la Covid-19, la consommation énergétique liée à la collecte et l'exploitation de données toujours plus nombreuses, et qui ne cesseront de croître avec la démocratisation des objets connectés (IoT), devrait pousser la part du numérique dans les émissions de gaz à effet de serre bien au-delà de ses 4% actuels ⁴.

Si le « Green IT » se développe déjà assez largement au sein des entreprises, la question d'un « **Green AI** », pour sa part, commence à peine à émerger. Pourtant, à l'heure où le marché mondial de l'IA a déjà été multiplié par 10 en 5 ans, et devrait encore croître de

² <https://www.salesforce.com/fr/blog/2020/08/nouvelles-tendances-nouveaux-defis-marketing.html>

³ <https://blog.hubspot.fr/marketing/tendances-marketing>

⁴ <https://www.greenit.fr/2020/10/06/4-des-emissions-de-ges/>

+120% d'ici 2024 ⁵, la problématique de l'empreinte écologique d'algorithmes voraces, tels que ceux du Deep Learning, ne peut plus être ignorée.

Comment rendre le modèle de l'IA marketing durable ? C'est la question à laquelle je me suis attaqué dans le cadre de cette thèse professionnelle.

Il est important de préciser que j'ai choisi de limiter le **périmètre de mon étude à l'enjeu environnemental lié à la notion de « durabilité »**, qui recoupe aussi des enjeux sociaux et économiques parmi les 17 Objectifs de Développement Durable définis par l'ONU. Il s'agit bien au global de tendre vers une IA digne de confiance, d'une IA « responsable » dans tous ses aspects : gestion des biais, gouvernance...

Focaliser ainsi la présente étude sur une « IA verte » permet de **mettre en lumière tout le paradoxe de la place de l'IA face à la question environnementale, à la fois risque mais aussi opportunité** pour lutter contre le réchauffement climatique. L'IA pourrait bien être notre meilleur allié dans le nouveau défi de ce siècle... pour peu qu'elle s'applique à elle-même la voie de la frugalité.

Ainsi, **après avoir exposé les bénéfices de l'IA marketing pour sur-optimiser la stratégie marketing** des entreprises en acquisition, conversion et fidélisation (plus de 40 outils opérationnels listés), **nous tenterons de dresser un bilan carbone de l'Intelligence Artificielle**. Nous mettrons en balance, d'un côté sa facture énergétique, en expliquant pourquoi et comment elle consomme, et de l'autre sa contribution pour la sauvegarde de l'environnement. Enfin, **nous expliquerons la nécessité d'adopter une démarche d'IA marketing durable à l'heure du marketing éco-responsable, en dressant un mode d'emploi pratique** pour que vous, marketeur, puissiez déployer une IA marketing durable adaptée, selon que vous soyez une grande ou une petite entreprise.

⁵ <https://itsocial.fr/enjeux-it/enjeux-innovation/intelligence-artificielle/en-cinq-ans-le-marche-mondial-de-lia-a-ete-multiplie-par-dix/>

Remerciements

Je tiens tout d'abord à remercier l'Institut Léonard de Vinci, et plus particulièrement le directeur pédagogique du MBA MCI, Alexandre Stopnicki, qui m'a accueilli au sein de cette formation d'excellence. Merci également à Christophe Dané et à l'ensemble des intervenants, qui nous ont transmis leur savoir avec passion et bienveillance.

Je souhaite ensuite remercier chaleureusement mon tuteur de thèse, Thierry Pires, qui m'a non seulement soutenu dans ce travail, mais m'a aussi donné l'opportunité de **créer avec lui et deux autres membres le collectif ia4marketing.fr, lancé en mars 2021**. J'ai pu y publier plusieurs articles et interviews que vous retrouverez en annexe, et qui sont venus alimenter ma thèse. **Ma thèse pourra d'ailleurs être téléchargée le jour même de ma soutenance sur ma page dédiée** : <https://ia4marketing.fr/theses-professionnelles/ia-marketing-durable-denis-lefebvre/> !

ia4marketing.fr

L'autre pendant de ma thèse concerne l'éco-responsabilité de l'IA, et **j'ai pu rejoindre l'Institut du Numérique Responsable et son groupe de travail sur l'IA durable**. Merci à Vincent Courboulay, à Eric Benoits et à Fabien Abrikh pour leur accueil, ainsi qu'aux 85 participants auprès desquels je vais continuer d'apporter ma contribution.



Je remercie aussi bien évidemment les **professionnels avec lesquels j'ai pu échanger dans le cadre de mes interviews** (profils sur la page suivante) :

- Claude de Loupy **CEO de Syllabs** (génération automatique de textes)
- Xavier Paulik **CEO de GetQuanty** (ciblage prédictif B2B)
- Jonathan Kam **COO de Botmind** (chatbot hybride)
- Laurent Cervoni **Directeur Recherche et Innovation chez Talan** (cabinet de conseil)
- Jérôme Totel **Vice-Président de la Stratégie de Data4** (Data Centers)
- Maurice Ndiaye **Co-Président de l'Adetem** (Association Professionnels du Marketing)
- Gwendal Bihan **CEO de Axionable** (cabinet de conseil en IA durable)
- Thomas Solignac **CEO de Golem.ai** (IA frugale en analyse automatique du langage)
- Matthieu Ruault **Directeur Marketing Data & Performance** chez Oui.SNCF

Merci à mes camarades de la promotion 2020-21 du MBA MCI, avec qui j'ai pu partager des moments de partage et d'entraide intenses, et ce en dépit de la situation sanitaire qui nous aura majoritairement tenue éloignés.

Je souhaite enfin et surtout remercier ma femme Céline et mon fils Simon pour leur amour et leur soutien tout au long de cette année. Merci aussi à mon associé avec qui j'ai co-fondé la start-up « Actialuna ».



syllabs



GetQuantity
Predictive B2B Scoring



botmind



Talan[★]
Consulting



DATA4!
SMART DATACENTERS
AT SCALE



Adetem
association nationale
professionnels marketing



axionable



Golem.ai



Oui_{.snCF}

Table des matières

Introduction	4
Remerciements	7
Table des matières	10
Résumé	14
Summary	16
L'IA au service de votre stratégie marketing	18
L'IA pour booster votre Acquisition	19
1. Dopez votre création de contenus	19
1.1. L'IA pour vous assister dans votre rédaction	20
Collectez des informations et générez des idées en quelques secondes	21
Optimisez votre rédaction	24
1.2. L'IA pour générer des contenus automatiquement	26
Des contenus textuels en automatique	26
Des contenus visuels en automatique	31
2. Sur-optimisez votre référencement	33
2.1. SEO : l'IA au service de votre sémantique	34
Toujours plus d'IA dans les algorithmes de Google	34
Comment exploiter vous-mêmes l'IA pour votre SEO ?	37
2.2. SEA, Display, Social Ads : ciblez juste et au meilleur prix	42
Un ciblage publicitaire toujours plus intelligent	42
Du programmatique au prédictif dans un monde cookieless	46
L'IA pour booster votre Conversion	52
1. Devenez omniscients pour convertir vos visiteurs	52
1.1. Traquez vos leads les plus chauds	53
Identifiez vos visiteurs ayant le plus de chance de passer à l'acte	53
Convertissez davantage grâce au ciblage prédictif	59

1.2. Prédisez les comportements de vos clients	63
Scorez l'appétence de vos clients pour un ciblage plus fin et efficace	63
Anticipez les ventes, le churn et vos prochaines offres	67
2. Hyper-personnalisez en automatique	74
2.1. Offrez une expérience client sur-mesure	74
De la personnalisation à l'hyper-personnalisation	75
Un parcours client repensé à l'aune de la data	78
2.2. Accompagnez vos clients 24h/24 7j/7	82
La révolution des chatbots IA	82
Les chatbots au service de la relation client	88
L'IA face au défi de son bilan carbone	101
Une IA de plus en plus énergivore	103
1. Qu'est-ce qui fait que l'IA consomme ?	104
1.1. IA, Machine Learning et Deep Learning	104
Une brève histoire de l'IA	104
Le fantasme de l'IA moderne : imiter le cerveau humain	106
1.2. Des modèles d'IA au besoin énergétique croissant	110
Comment se construit une IA ?	110
Où est la consommation d'énergie ?	115
2. Mesurer l'empreinte carbone de l'IA	118
2.1. Quel est le coût CO2 de l'Intelligence Artificielle ?	118
Une évaluation du bilan carbone de l'IA	118
L'IA dans le Big Bang du Big Data	122
2.2. Le risque d'une croissance exponentielle ?	127
La folle tentation du 1% supplémentaire	127
Une bataille économique, éthique et géo-politique	130
L'IA au service de l'environnement	138
1. L'IA pour lutter contre le changement climatique	138
1.1. L'IA pour dépolluer	140
Dépolluer les mers	140
Dépolluer les sols	144
Dépolluer l'air	145

1.2. L'IA pour préserver la biodiversité	149
Préserver les espèces menacées	149
Préserver les environnements menacés	154
1.3. L'IA pour prévenir les catastrophes naturelles	156
Anticiper la détection des feux de forêt	156
Anticiper les tsunamis, les séismes...	157
2. L'IA au service de la transition énergétique	160
2.1. L'IA pour améliorer l'efficacité énergétique	161
Améliorer la rentabilité des énergies à faible impact carbone	161
Réduire la dépense énergétique des data centers	164
Autres axes d'optimisation énergétique	172
2.2. L'énergie des données : « l'IA est la nouvelle électricité »	174
Des serveurs en guise de radiateurs dans les HLM	175
Chauffer les hôpitaux et les églises aux Bitcoins	177
Une IA durable pour un marketing éco-responsable	180
L'IA marketing à l'ère de l'éco-responsabilité	181
1. La déferlante du marketing éco-responsable	181
1.1. De nouvelles habitudes de consommation	182
L'environnement : un enjeu majeur pour les Français	182
Les marques s'adaptent à ce nouveau paradigme	185
1.2. Coup de com ou véritable engagement ?	188
Les dangers du Greenwashing	189
Les bénéfices d'un marketing responsable sincère et transparent	193
2. L'IA marketing au cœur du numérique responsable	199
2.1. Accélérer son éco-responsabilité grâce à l'IA	199
L'IA pour optimiser le cycle de vie du produit et son marketing	200
L'IA, outil marketing de la transparence : pour un « Yuka du CO2 »	203
2.2. Green IT : où en est le numérique responsable ?	206
GAFAM : des Géants bientôt Verts ?	206
Green IT en entreprise : réconcilier écologie et économie	215
Comment rendre l'IA marketing durable ?	218
1. Évaluez les modalités de recours à l'IA	220

1.1. Cas n°1 : vous développez votre propre modèle d'IA	221
Questionnez l'utilisation des modèles d'apprentissage selon vos besoins	222
Évaluez l'environnement IT le plus sustainable	225
1.2. Cas n°2 : vous adoptez une solution IA clé en main	239
Définissez vos objectifs marketing pour cadrer votre besoin en IA	240
Évaluez l'éco-responsabilité des solutions d'IA sur le marché	241
2. Rationalisez vos méthodes dans une démarche ROIste	252
2.1. Faites les bons choix en terme d'optimisation	252
Cas n°1 : optez pour l'éco-conception et le recycling de l'apprentissage	253
Cas n°2 : ciblez le travail de l'IA pour mieux rentabiliser vos campagnes	257
2.2. Réduisez vos jeux de données	263
Passez de la Big Data à la Smart Data	263
Organisez et vidangez votre Data Lake	267
3. Analysez vos résultats, Communiquez sur votre impact	270
Conclusion	274
Recommandation synthétisée	277
Boîte à outils	278
Annexes	279
Bibliographie	280
Vidéo de teasing de thèse	283
Mes articles et podcast #MBAMCI	284
Mes articles sur ia4marketing	286

Résumé

L'Intelligence Artificielle devient un allié incontournable de votre stratégie marketing en vous aidant à mieux connaître vos cibles, mieux les segmenter et mieux les adresser. Ce qui importe à présent, ce n'est plus tant le *persona* que le comportement du visiteur en temps réel sur votre site. Pour capter l'ensemble de ses « signaux faibles » révélateurs d'une appétence pour vos produits, d'une intention d'achat ou d'abandon, **adoptez une approche *data driven*** en vous appuyant sur les capacités de collecte et d'analyse de l'IA.

Elle vous permettra de **sur-optimiser toutes les étapes de votre parcours client**, de votre acquisition en content marketing, référencement naturel et payant... jusqu'à augmenter vos ventes grâce à vos nouveaux super-pouvoirs de prédiction. Grâce à eux, vous allez pouvoir envoyer le bon message à la bonne cible, au bon moment, sur le bon canal, et ainsi hyper-personnaliser son expérience tout en limitant la pression commerciale.

Mais **votre stratégie marketing doit répondre à un autre enjeu grandissant : celui de la prise de conscience écologique des consommateurs**, accentuée par l'intensification des catastrophes naturelles et la crise de la Covid-19. Le marketing éco-responsable devient plus qu'une tendance : il devient **un impératif** auquel chacune de vos actions marketing doit répondre, en communiquant de façon sincère et transparente pour y faire adhérer.

Or, **l'éco-responsabilité des marques va se heurter à la consommation énergétique de l'IA** au sens large, et de l'IA marketing en particulier. L'hébergement de grands volumes de données et l'apprentissage de modèles en quête de toujours plus de performances, en font une technologie qui risque de devenir bien trop énergivore en s'industrialisant.

Pourtant, **l'IA pourrait aussi être notre meilleur allié dans la lutte contre le dérèglement climatique**, en mettant à notre disposition ses capacités d'analyse et de prédiction pour anticiper les risques, assurer la transition énergétique et préserver la biodiversité.

Pour que l'IA soit un vrai moteur de la révolution écologique il faut donc **veiller à assurer sa propre frugalité**. Pour cela, il revient aux entreprises d'adopter les bonnes pratiques.

Le mode d'emploi pour une IA marketing durable que je vais vous proposer mettra en lumière **5 étapes clés à suivre, selon 3 axes**, et adapté selon que vous vouliez créer votre propre modèle, ou adopter une solution SaaS clé en main :

- **Évaluez** vos objectifs marketing pour cadrer le périmètre de votre besoin en IA, ainsi que l'environnement IT le plus éco-responsable pour le faire ;
- **Rationalisez** le développement et l'entraînement des algorithmes, tout en réduisant vos jeux de données pour distinguer la donnée utile de la donnée futile, dans une démarche ROIste de gains aussi bien écologiques que économiques ;
- **Communiquez** sur l'impact de vos actions, en valorisant la performance de vos KPIs environnementaux, et les labels venant certifier l'éco-responsabilité de votre IA.



Découvrez ma vidéo de teasing de thèse

Summary

Artificial Intelligence becomes an essential ally of your marketing strategy, helping you to better know your targets, to better segment them and to better address them. What matters now is not so much the "persona", but the visitor's behavior in real time, on your site. To capture all the "weak signals" that indicate an appetite for your products, an intention to buy or to abandon, **adopt a data-driven approach** based on the collection and analysis capabilities of AI.

It will allow you to **over-optimize all the steps of your customer journey**, from your acquisition in content marketing, natural and paid search... to increasing your sales thanks to your new predictive superpowers. Thanks to them, you will be able to send the right message to the right target, at the right time, on the right channel, and thus hyper-personalize their experience while limiting sales pressure.

But **your marketing strategy must respond to another growing challenge: the one linked to the ecological awareness of consumers**, accentuated by the intensification of natural disasters, and the Covid-19 crisis. Eco-responsible marketing is becoming more than a trend: it is becoming **an imperative** to which each of your marketing actions must respond, by communicating sincerely and transparently so to make people adhere to it.

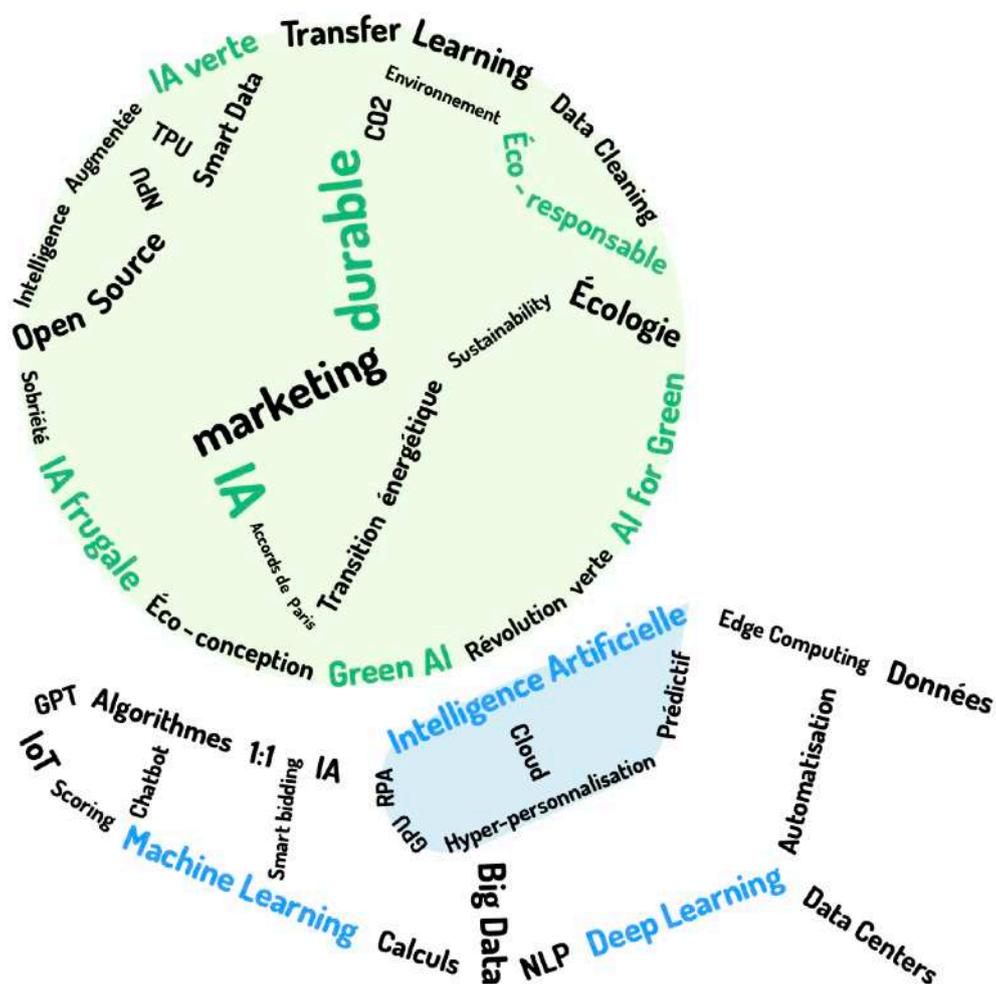
However, **the eco-responsibility of companies will come up against the energy consumption of AI** in the broadest sense, and of marketing AI in particular. The hosting of large volumes of data, and the learning of models looking for ever higher performance, make it a technology that risks becoming far too energy-intensive as it becomes industrialized.

However, **AI could also be our best ally in the fight against global warming**, by providing us with its analysis and prediction, capabilities to anticipate risks, ensure the energy transition and preserve biodiversity.

For AI to be a real driver of the ecological revolution, **we must therefore ensure its own frugality**. For this, it is up to companies to adopt good practices.

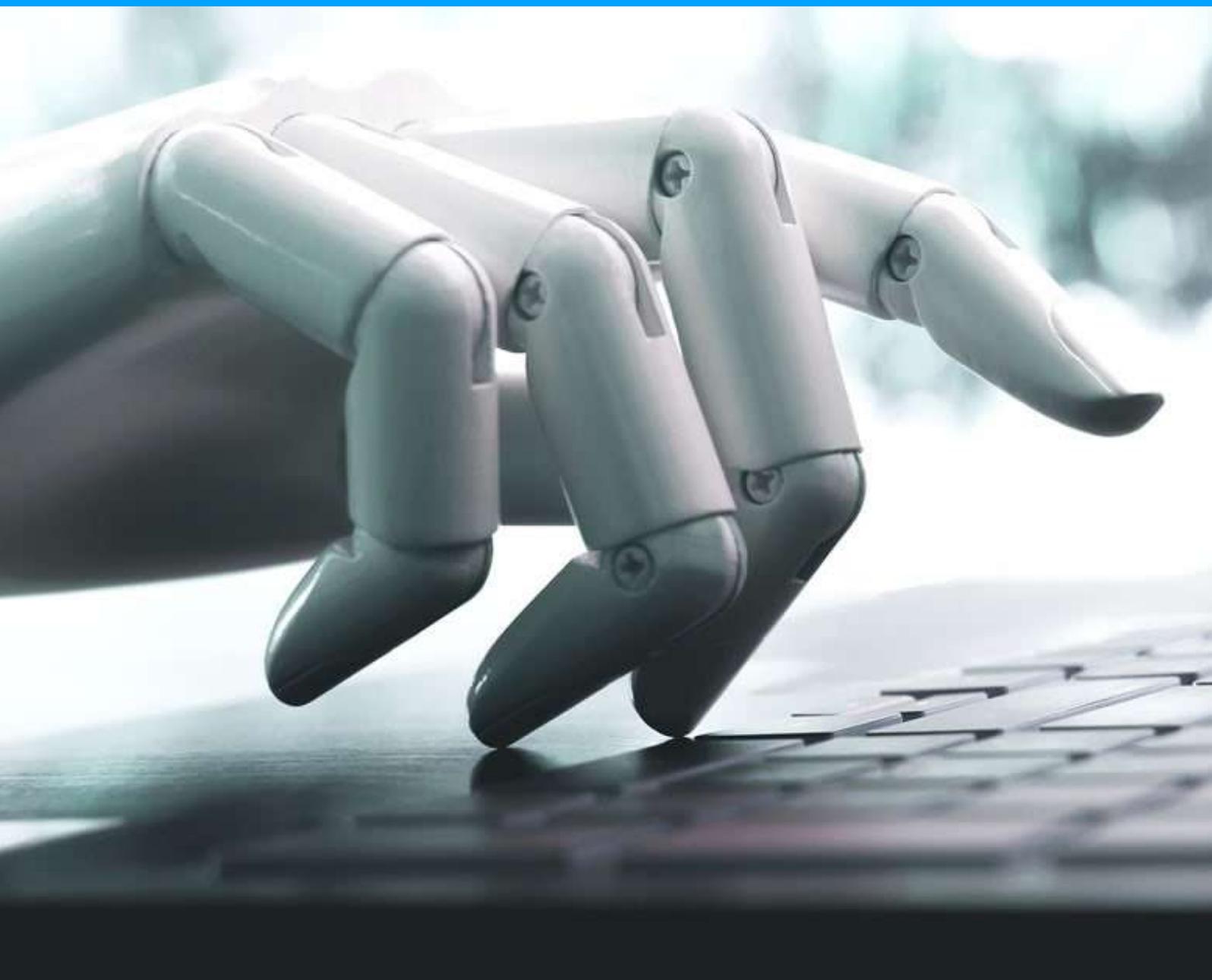
The instructions for a sustainable marketing AI that I will propose will highlight **5 key steps to follow, according to 3 axes**, and adapted according to whether you want to create your own model, or adopt an existing SaaS solution:

- **Evaluate** your marketing objectives to define the scope of your AI needs, as well as the most eco-responsible IT environment to do so;
- **Streamline** the development and training of algorithms, while reducing your datasets to distinguish useful data from futile data, in a ROI approach of ecological as well as economic gains;
- **Communicate** on the impact of your actions, by highlighting the performance of your environmental KPIs, and the labels that certify the eco-responsibility of your AI.



CHAPITRE 1

L'IA au service de votre stratégie marketing



L'IA pour booster votre Acquisition

Principaux pôles d'utilisation de l'IA par les CMO américains



© Denis LeFevre, inspiré par MarketingCharts.com
"The CMO Survey" réalisé auprès de 122 CMOs américains d'entreprises américaines (1er trimestre 2018)

Les usages de l'IA pour votre stratégie marketing sont à la fois nombreux et efficaces. Voyons déjà comment **l'IA va vous aider à devenir un vrai aimant à leads**, en dopant votre content marketing, et en sur-optimisant votre référencement naturel et payant.

1. Dopez votre création de contenus

On le sait, l'un des enjeux clés pour accroître la visibilité de sa marque ou de son produit, c'est de créer du contenu. Toujours plus de contenu. Et pour une stratégie d'inbound marketing efficace, vous devez les adapter et les diffuser sur tous les canaux pertinents pour votre cible : blog, emails, push, réseaux sociaux, publicités, SMS...

Pas facile à gérer... d'autant que face à la multitude d'informations, les internautes n'y prêteront attention que s'il est engageant, personnalisé, et qu'il répond à leurs besoins.

Mais bonne nouvelle : **pour publier du contenu rapidement et à grande échelle, trouver des idées de contenus intéressants, bien positionnés face à la concurrence et pour votre SEO, vous pouvez à présent compter sur l'Intelligence Artificielle.**

Des entreprises comme Orange, eBay, Expedia, Sephora et bien d'autres l'ont compris, et y ont déjà recours depuis des années pour alimenter le contenu de leurs campagnes.

Alors **dopez vous aussi votre marketing de contenu à l'IA !**

1.1. L'IA pour vous assister dans votre rédaction

Le copywriting, c'est l'art de convaincre avec les mots, et c'est un art exigeant :

- Vous devez sélectionner parmi une multitude de choix possibles les bons termes, ceux qui inciteront à cliquer et acheter.
- Vous essayez plein d'accroches différentes, et vous faites varier vos textes selon le canal, le segment d'audience, la période, et bien d'autres paramètres.
- Vous avancez à tâtons en *test and learn*, et s'il est satisfaisant de voir vos résultats progresser, tout ce travail reste extrêmement chronophage et laborieux.

Mais en tant que copywriter, vous votre « truc », c'est surtout de jouer avec les mots, d'adopter un style percutant et persuasif qui saura susciter de l'émotion... et donc de l'engagement.

Alors pourquoi vous épuiser à faire ce que l'Intelligence Artificielle pourrait faire pour vous ? Vous avez mille fois mieux à apporter à vos clients : **focalisez-vous sur l'émotion qui fait vendre, sur la stratégie et la performance de vos campagnes. Et appuyez-vous sur l'IA pour le reste.** Vous ne pourrez compter que sur votre talent pour faire preuve de créativité, mais vous pourrez compter sur l'IA pour vous décharger de toutes les tâches fastidieuses.

Alors **faites équipe avec votre assistant d'écriture virtuel pour accélérer votre travail et exceller dans la rédaction de vos contenus !**

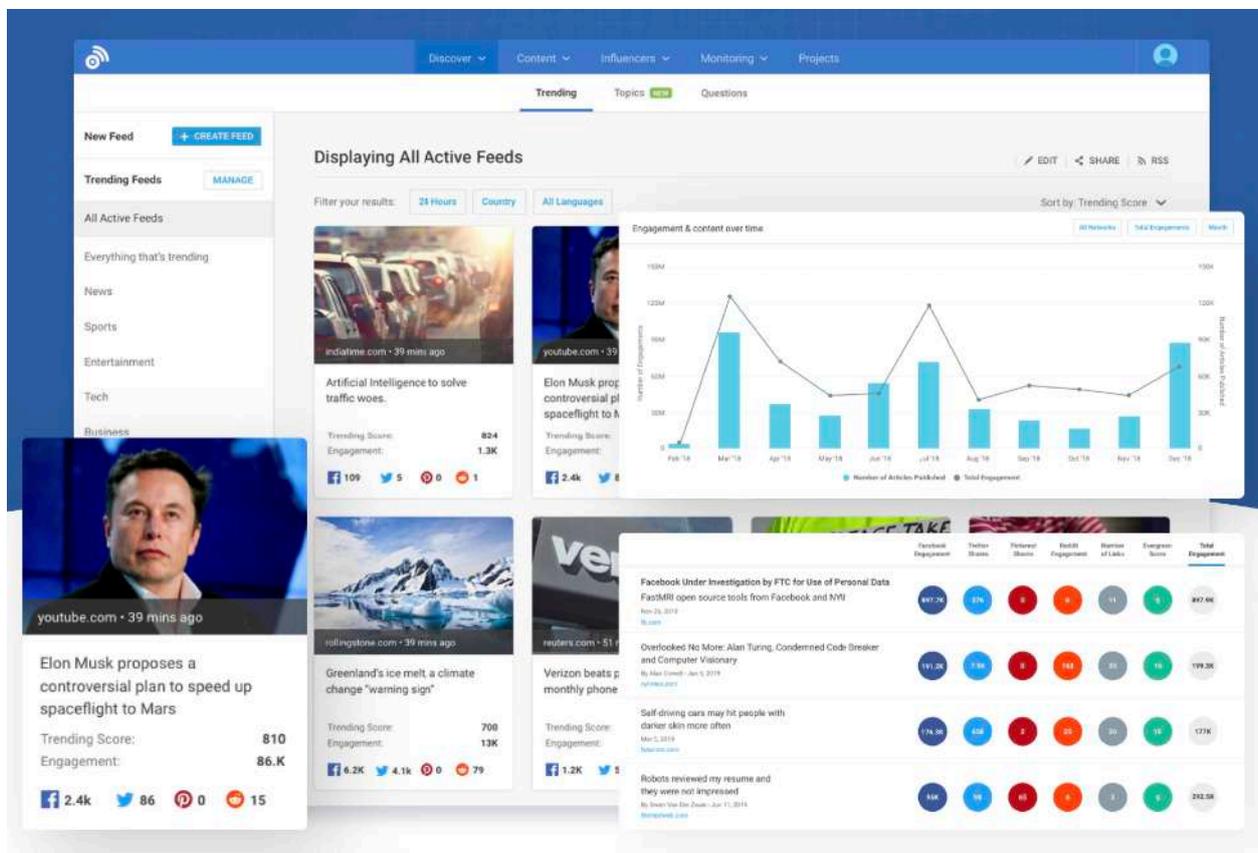
COLLECTEZ DES INFORMATIONS ET GÉNÉREZ DES IDÉES EN QUELQUES SECONDES

Avant de vous lancer dans l'écriture, **encore faut-il savoir sur quoi écrire**. Ou plutôt quelle approche originale adopter pour valoriser votre marque ou votre produit, sans tomber dans un discours purement commercial. La concurrence à l'attention est forte, et il faut **vous positionner efficacement sur les intentions de recherche de vos cibles** pour espérer pouvoir les faire venir jusqu'à vous.

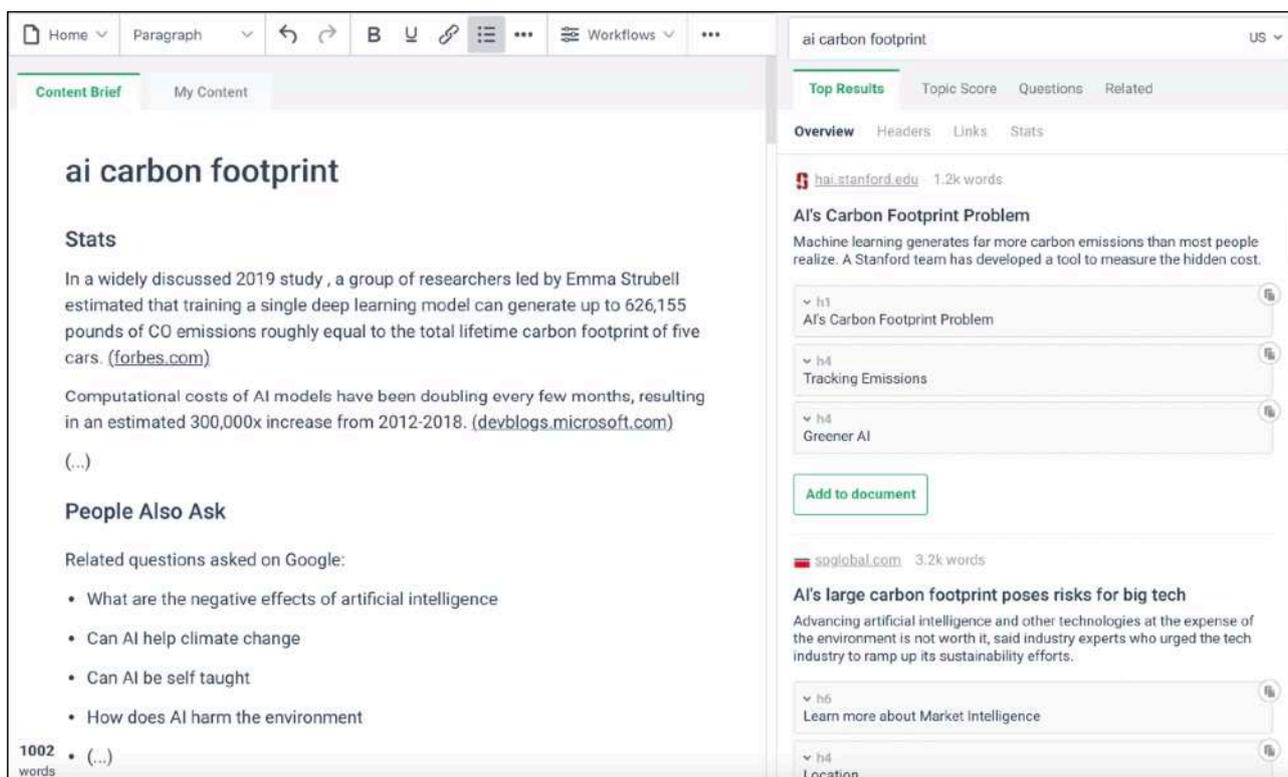
À mesure que vous aurez épuisé les sujets de votre industrie, **l'intuition ne suffira plus** : pour ne pas vous retrouver à court d'idées, misez sur une curation de contenus à même de vous remonter l'information utile beaucoup plus rapidement que sur Google.

Plusieurs outils dopés à l'IA vous permettront de gagner un temps précieux :

- **BuzzSumo** par exemple vous permet de **récolter et filtrer tous les articles liés à un sujet**, mais aussi **d'identifier des sujets connexes et tendances liés à votre audience** (questions/réponses les plus posées sur Quora, Reddit et Yahoo, social listening...).

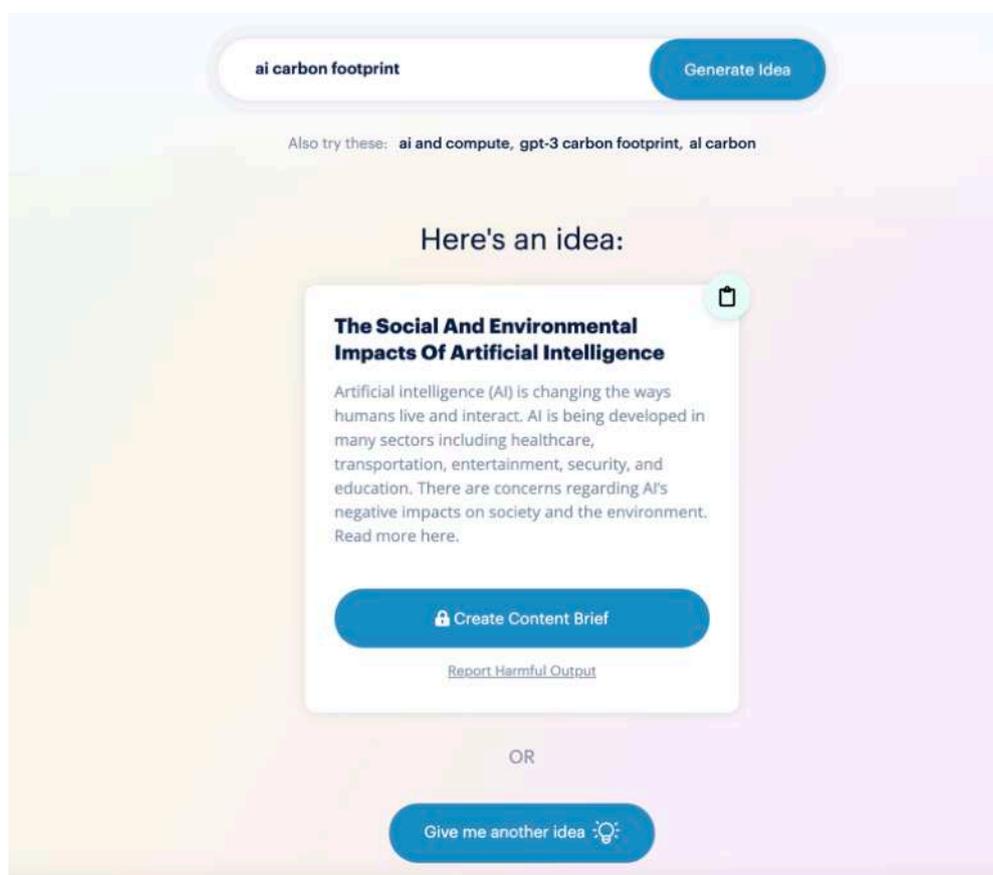


- **Frase Content** va plus loin, en organisant **toutes les informations clés autour du sujet que vous voulez traiter sous forme de brief**, et ainsi grandement faciliter votre rédaction. Après avoir tapé les mots-clés propres à votre thématique, Frase Content va non seulement **vous remonter en un rien de temps les articles les plus pertinents** et déjà publiés qui s’y rapportent, mais aussi **l’ensemble des chiffres et statistiques à ré-exploiter, les meilleurs backlinks, et les mots-clés les plus compétitifs sur lesquels se positionner**. La structure des articles existants (H1, H2, H3...), le nombre de mots utilisés et bien d’autres indicateurs vont être pour vous des sources d’inspiration et vous permettre de vous placer en bonne position dans les résultats de recherche. Ce brief peut être préparé pour vous ou pour des rédacteurs, puis exporté en PDF, en HTML ou même directement dans WordPress. Topic est un autre outil assez similaire, que nous explorerons ultérieurement (p.37). Mais l’avantage de Frase reste qu’il est aussi disponible en français, et que vous pouvez le tester gratuitement ici.



- Mais montons encore d’un cran. Que diriez-vous si vous n’aviez qu’à taper quelques mots, appuyer sur un bouton magique, puis laisser la machine faire le reste ?... C’est

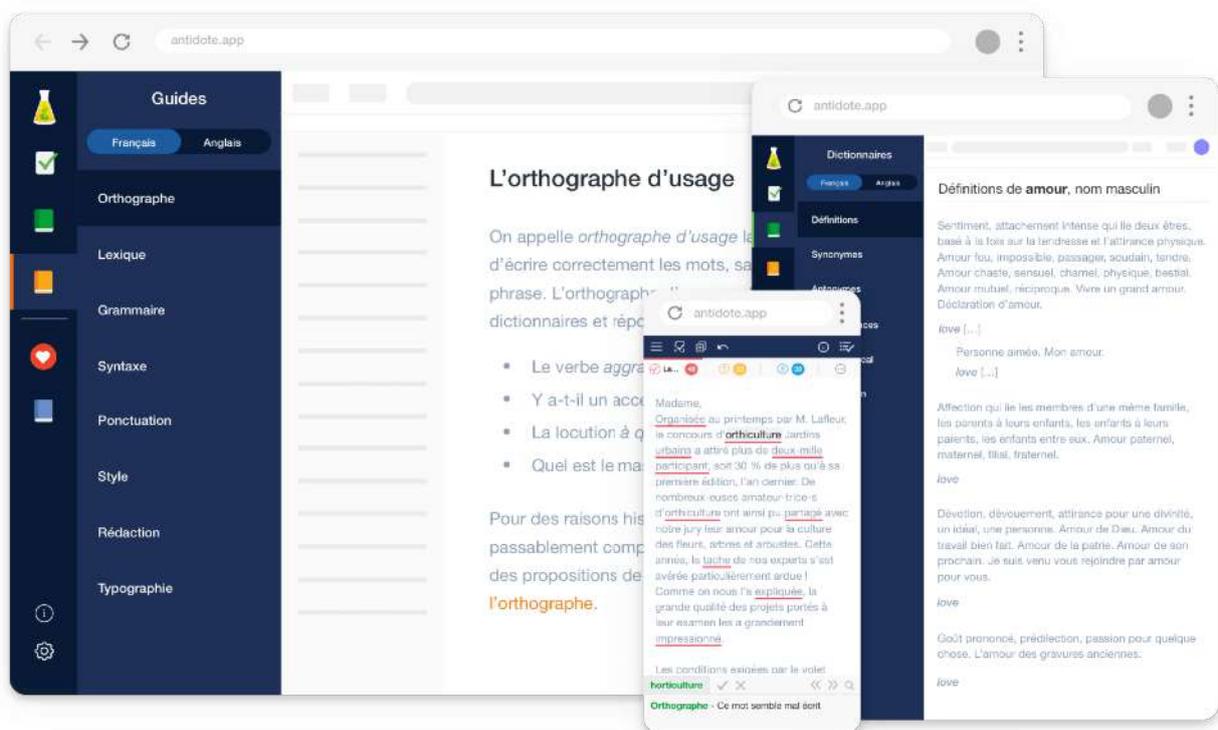
à présent possible avec des outils comme [AI-Writer](#), [Talk to Transformer](#), [Copysmith](#), [Article Generator](#), [Articulo](#) ou encore [WordSmith](#) : **entrez quelques mots-clés, et vous vous verrez proposer des textes structurés et compréhensibles**, générés à partir de données collectées sur le Web et restituées avec un ton naturel, qui fait de plus en plus illusion. **Grâce au *Natural Language Processing (NLP)***, ou Traitement du Langage Naturel (TAL), **l'IA sait collecter, recouper et synthétiser mieux que quiconque les contenus de sources fiables** (les liens vous en assureront). Puis les réécrire suffisamment pour qu'ils ne soient pas considérés comme dupliqués. Toutefois, ne comptez pas publier ces textes en l'état, car l'IA n'en ressort aucune analyse pertinente. Servez-vous en plutôt comme d'**un socle pour vous lancer**, vous inspirer, y piocher des citations, statistiques, phrases bien tournées... Et **ajoutez-y votre originalité et votre créativité** pour adresser votre audience. Le [Blog Idea Generator de Topic](#) (image ci-dessous) ou la fonction « Write for me » de Frase Content vous permettra de générer automatiquement un texte que vous pourrez ensuite affiner avec tous les insights précédemment cités.



OPTIMISEZ VOTRE RÉDACTION

L'Intelligence Artificielle peut également vous permettre d'**écrire mieux, plus vite et plus loin**. Elle est en cela **un véritable assistant de rédaction** :

- **Chassez les fautes d'orthographe, de grammaire et de syntaxe** grâce à des outils comme **Cordial** et son moteur de correction expert en traitement automatique du langage. **Antidote** propose quant à lui le correcteur le plus avancé du marché : il va jusqu'à corriger les erreurs de style et de logique, les répétitions, etc. depuis vos différents logiciels, et via une extension Web. L'IA de **Grammarly**, au-delà de corriger votre grammaire et votre orthographe, vous permet aussi d'adapter votre style et votre ton en fonction des interlocuteurs auxquels vous vous adressez.



- **Réécrivez des phrases déjà publiées pour ne pas avoir de problèmes de contenus dupliqués** : **Smodin** (gratuit) utilise l'apprentissage automatique et le traitement en langage naturel profond pour réécrire dans une centaine de langues vos textes, en conservant le bon contexte syntaxique, lexicale et textuelle.

- **Utilisez l'écriture prédictive qui vous suggère le mot suivant en apprenant de vos rédactions passées,** comme sur la fonction « Rédaction intelligente » de Gmail, à présent aussi disponible sur Google Docs.



- **Pré-traduez vos textes automatiquement dans différentes langues et touchez ainsi davantage de cibles partout dans le monde.** Google Traduction se décline partout dans l'écosystème Google : traduisez vos textes directement depuis Google Documents, ou avec l'extension Chrome qui est aussi très pratique pour lire n'importe quel article du Web depuis/vers n'importe quelle langue. La solution allemande DeepL est quant à elle considérée comme le meilleur outil de traduction automatique actuel ⁶ grâce à ses réseaux de neurones élaborés à partir du 23ème ordinateur le plus puissant au monde.

Votre objectif reste que votre contenu soit le plus engageant possible. L'IA peut vous aider à **améliorer drastiquement vos taux d'ouverture et de clic**, en vous aiguillant vers les formulations et vers les CTA qui sauront les mieux convertir.

Pendant que vous passez des heures à vous interroger sur l'utilisation de tel ou tel mot, elle saura vous dire en quelques secondes celui qui aura le plus d'impact.

Phrasee en particulier apprend de votre historique de campagnes pour **identifier les termes qui ont su obtenir le plus d'engagement**. Il est aussi capable de **restituer sous la forme d'un langage naturel votre style, le « ton » de votre marque**. Vous pouvez alors décliner la « recette Phrasee » sur vos mailings, et sur tous vos formats courts (publicités Facebook et Google, SMS, push notifications, etc.).

⁶ <https://www.meilleure-innovation.com/outil-traduction-deepl-google-reverso/>

1.2. L'IA pour générer des contenus automatiquement

DES CONTENUS TEXTUELS EN AUTOMATIQUE

L'IA peut vous préparer le terrain, et vous aider à écrire plus efficacement. Mais elle peut **vous faire gagner encore plus de temps en écrivant complètement à votre place**. Ce ne sera pas possible pour votre copywriting, qui comme nous l'avons vu exige par ailleurs un savoir-faire unique, que la machine ne saurait reproduire : celui d'hypnotiser le lecteur, le séduire et le toucher au cœur pour lui vendre votre produit ou votre service.

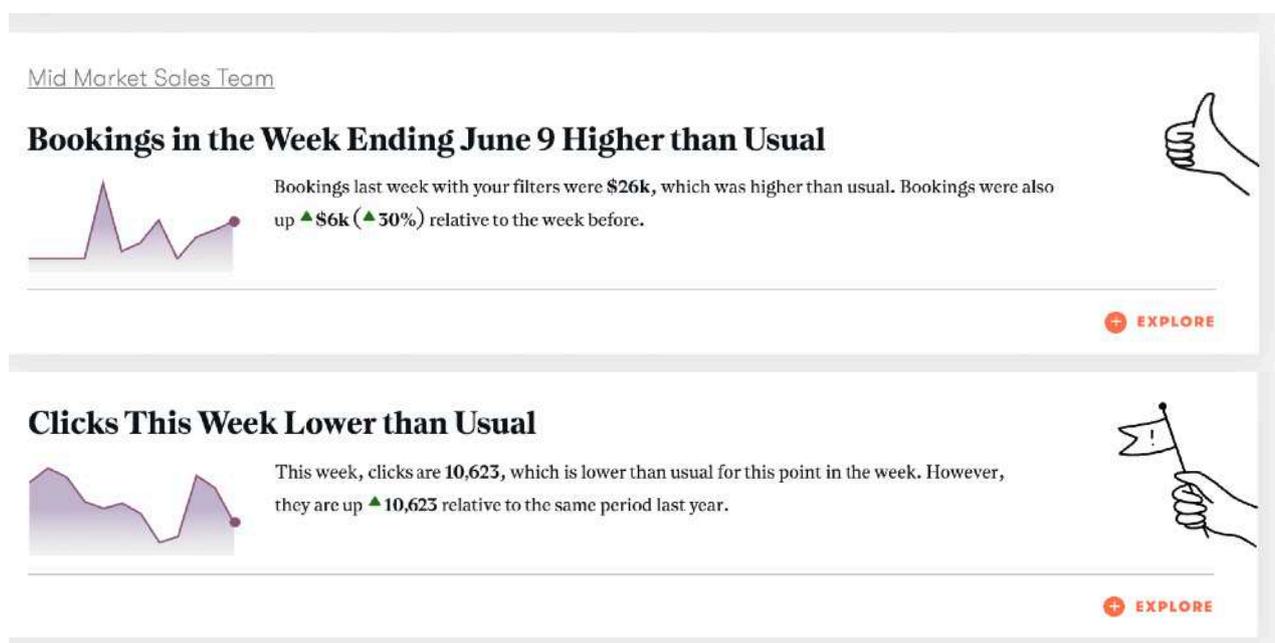
Mais prenons les milliers de fiches produits que vous devez rédiger pour votre site e-commerce : avez-vous une vraie valeur ajoutée et le temps nécessaire pour écrire ces textes purement descriptifs ? On ne va pas se mentir : l'IA sera bientôt **plus efficace qu'un rédacteur humain pour tous les contenus courts, factuels, répétitifs et prédictifs**.

Les médias se sont emparés des moteurs de rédaction autonomes depuis des années : le Washington Post utilise par exemple depuis 2016 *Heliograf*, capable de générer seul des articles sur des événements qui ont d'abord vocation à délivrer une information locale, ne nécessitant pas d'interprétation particulière : résultats sportifs ou électoraux, météo, activité sismique, etc. On peut également citer Bloomberg avec son IA *Cyborg*, ou Forbes avec *Bertie*. Les journalistes, qui ne pourraient pas couvrir ces informations, sont en revanche alertés par l'IA dès lors qu'elle détecte une information anormale qui mériterait une analyse approfondie. La collaboration est totale, et très utile pour créer de la micro-information permanente capable de se transformer en lanceur d'alertes.

Cette « **narration de la donnée** » est extrêmement utile en communication interne, pour diffuser une lecture simple des principales KPIs de votre activité, sous forme d'articles auto-générés. Selon Gartner ⁷, d'ici 2025 le « data storytelling » automatique devrait être le moyen le plus répandu d'appréhender des analyses.

⁷ Gartner, « Top 10 Trends in Data and Analytics, 2020 » : <https://www.gartner.com/doc/reprints?id=1-24ZN6MQG&ct=210106&st=sb>

L'IA sait extraire automatiquement les informations utiles de documents complexes, en les restituant de manière lisible pour vous aider à prendre les bonnes décisions. C'est ce que propose par exemple **Lexio de NarrativeScience** : l'analyse des données de l'entreprise est remontée depuis vos comptes Google Analytics, Google Ads, Salesforce etc. et accessible à tous sous forme d'un tableau de bord interactif et dynamique, transverse aux activités de vos différents services (techniques, juridiques, financiers, marketing et commerciaux). Lexio transforme vos données en storytelling, pour les faire comprendre et adopter par tous les salariés.



Les robots rédacteurs fleurissent pour également alimenter en automatique votre **stratégie de contenus**, et maintenir l'attention de vos clients et prospects. La solution française **Syllabs**, concurrente de **LabSense**, a su convaincre des médias comme Le Monde, Radio France ou Ouest-France, mais aussi des acteurs du e-commerce, de la finance ou de l'immobilier. Je vous propose de découvrir **dans mon interview à suivre comment fonctionne concrètement l'IA de Syllabs**.



Claude de Loupy CEO de Syllabs



En quoi l'utilisation de l'intelligence artificielle dans la solution Syllabs permet d'améliorer la stratégie de contenus de vos clients ? Quelles sont vos solutions spécifiques à l'e-commerce ?

« Syllabs produit des textes automatiquement à partir de données structurées. Car **ce n'est pas de la magie, on a besoin de données en entrée**. On utilise des systèmes d'apprentissage pour préparer le terrain à nos linguistes, qui valident tous les éléments éditoriaux au moment du paramétrage. **Une fois que c'est paramétré, ça fonctionne tout seul.**

Dans des contextes de marque, ça peut être très périlleux de faire confiance à une machine qui n'est pas contrôlée par des humains. Il n'y a pas dans le monde un outil qui permette de dire « vous pouvez publier les yeux fermés » uniquement à partir d'un système automatique. **Le Monde publie nos textes sans nous relire**, mais c'est parce que nous faisons un vrai travail de compréhension et d'analyse de la donnée. On a des Data Scientists chez nous qui s'en occupent, et qui se penchent sur les cas particuliers. **On a généré jusqu'à 1 million de textes en 24h sur des résultats électoraux. À ce niveau-là on doit garantir la perfection.** Les journalistes au début étaient quand même suspicieux, mais les lecteurs comme eux n'ont remonté aucune erreur. Ils font à présent confiance dans notre technologie, qui a su faire ses preuves et se perfectionner depuis 2006.

Pour illustrer la façon dont fonctionne Syllabs, **prenons un exemple sur la fiche descriptive d'une automobile.**

Choix des données

Demo: *Descriptifs de véhicules*

Jaguar F-Pace

Générer à nouveau

general_informations

id	1234
brand	Jaguar
model	F-Pace
version	
start_commercialization	03/07/2017

Texte généré

Jaguar F-Pace 2.0 i 300 03/07/2017

Dans la catégorie SUV, on trouve le Jaguar F-Pace : les chances d'être déçu sont minces !

Ses 5 places s'accompagnent d'un coffre modulable et **très spacieux** pouvant aller de 432 à 650 litres. La puissance fiscale de ce modèle 5 portes est de 20 chevaux fiscaux. Le véhicule qui a été lancé en juillet 2017, ne se trouve plus sur le marché du neuf depuis juillet 2018.

En ce qui concerne la transmission de ce Jaguar F-Pace, il s'agit d'une boîte manuelle (BVM6) 5 rapports. On note une consommation mixte très faible, avec 4 litres au 100. On mesure à 7,2 litres au 100 sa consommation urbaine, et à 8,9 litres au 100 sa consommation extra-urbaine. Vous serez exempté du malus écologique grâce à une émission s'élevant à 105 g/km. C'est un moteur 2.0 i 300 (essence) d'une puissance réelle de 300 ch ainsi que d'une cylindrée de 2.000 cm³ pour 4 cylindres que l'on retrouve sous le capot de ce SUV.

Le dispositif stop and start dont cette automobile est équipée permet des économies, particulièrement en ville.

Masquer les images

Push to my WordPress

Link to my WordPress

1. **On a un ensemble de données structurées** comme la marque, le modèle, le nombre de places, etc.
2. **On clique et ça génère un texte.** Si je re-clique j'ai un texte différent, et ainsi de suite. On a des règles qui disent qu'au-delà d'un certain volume un coffre est « très spacieux ». Cette sémantique peut servir d'argumentaire marketing. On fait parfois appel à des experts externes pour savoir comment interpréter des données, par exemple dans le cadre d'un article sur un match de foot. On peut ainsi décrire tout type de produit : descriptif e-commerce, annonces immobilières, fil boursier, etc.
3. **Le client peut rééditer un passage** s'il le souhaite, et pousser le contenu sur son site ou d'autres supports.

On a des preuves de performance SEO y compris chez nos clients qui publient des milliers de pages du même type, et qui n'ont pas de problème de duplicate content. L'avantage c'est que nos textes répondent à des requêtes, ce que Google aime.

On fait beaucoup de descriptif produits. On fait aussi du mailing hyper-personnalisé en adaptant le texte non seulement à la tonalité de la marque, mais aussi à celui auquel on s'adresse, selon son âge, le temps qu'il fait, sa ville, ses achats. Ce qu'on dit à nos clients, c'est « faites sauter vos limitations, vous avez l'habitude de travailler avec 4, 5, 10 segments de clientèle, là vous en avez des millions. Et on peut s'adapter à ceux que vous voulez ».

On fait beaucoup de Web-to-Store avec des publications sur l'immobilier local, qui permettent de montrer l'expertise d'une agence sur son territoire.

Pour assurer la rétention et la transformation, on sait enfin implémenter du linking automatique à partir d'une liste de pages importantes du site, déterminées avec nos clients. »

Cas client

300

sites d'agences transformés en vrais médias locaux alimentés automatiquement par le moteur de rédaction Syllabs.



25 000

publications dans les 6 premiers mois.

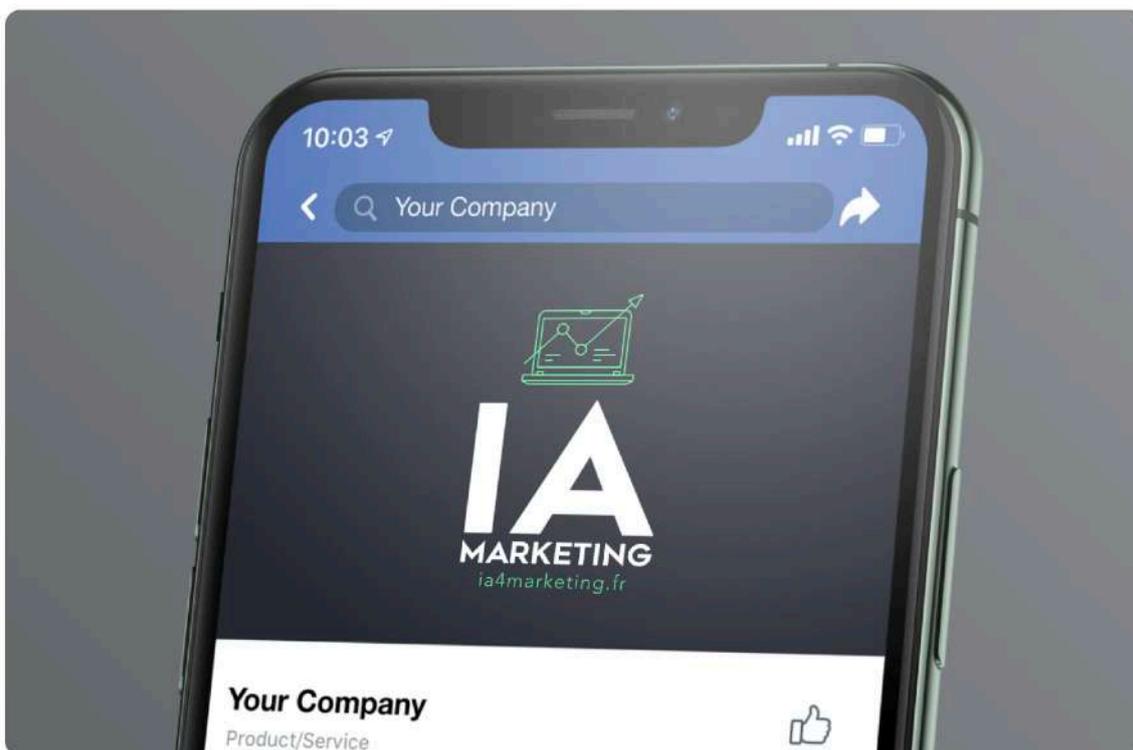
+ 196%

de trafic issu des recherches naturelles grâce aux billets d'actu immo et locale

DES CONTENUS VISUELS EN AUTOMATIQUE

L'IA vous permet également de générer et d'optimiser des contenus visuels au rendu très qualitatif, sans avoir besoin de passer par un graphiste. Citons par exemple :

- **Looka** pour créer des logos en quelques secondes, comme j'ai pu le tester pour ia4marketing.fr sur l'exemple ci-dessous :



- **Pencil** pour générer automatiquement des publicités Facebook : à partir de votre charte graphique (logo, typographie, couleurs), de vos messages clés et vos visuels, l'IA de Pencil crée pour vous des Facebook Ads originale, créative et de qualité en moins de 20 minutes, et dans tous les formats. Vous pouvez les éditer facilement, en vous voyant suggérer d'autres accroches pour atteindre votre objectif (ROAS, CPA, CTR). Pencil annonce environ **+17% du ROAS (Return on Ad Spent) d'une marque sur ses Facebook Ads** grâce à son outil. Pencil se base sur 5 algorithmes de Machine Learning pour notamment déterminer les « régions d'intérêt » à ne pas « croquer » en générant les différents formats d'Ads, mais aussi les performances historiques des précédentes publicités de l'annonceur pour en extraire les *best practices*.

Compared to the results of 10 Pencil ads you've previously run on Facebook.

Goal: CONVERSIONS Targeting: US_18_64_Lo...

Score: CPA 62 Above average

Concept 1 - 10 sec

This concept has the following features that have worked in your past ads:

- Video references "human", "person" and "lighting"
- Heavy on product imagery
- Low to no use of cta buttons
- Clean, less cluttered visual style

Advanced Editor

Durable and dynamic work lighting Handle-mounted c.o.b.s with solid flashing modes Prepare for the power of a beam Introducing

Le Machine Learning démultiplie les possibilités de créations publicitaires digitales : le DCO (Dynamic Creative Optimization) permet ainsi de **s'adapter au profil, au contexte et/ou au comportement de l'internaute** (comme une publicité pour un voyage en train depuis sa localisation), **et en testant sur lui diverses variables** (CTA, visuel, couleurs...) pour limiter le sentiment de pression commerciale, tout en maximisant les chances de convertir. Le clic permettra aussi de mieux connaître les **éléments à même de l'engager ou le ré-engager par la suite, tout en appliquant les mêmes critères de réussite à des profils similaires sur l'audience ciblée**. Plusieurs milliers de bannières peuvent ainsi être testées dans une campagne DCO grâce à l'apprentissage automatique.

Analyse de données



Données de profil



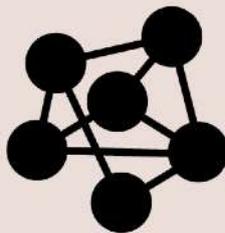
Données contextuelles



Données marketing



Algorithmes



Génération de bannière en DCO

DESTINATION

FLYING COMPANY

Great deals on all Flights to Tokyo

REAL-TIME PRICE

RETURN FLIGHTS FROM £367

1ST OCT - 1ST DEC

PAST-SEARCHED DATE RANGE

BOOK NOW

LOCATION IMAGE

CALL TO ACTION

2. Sur-optimizez votre référencement

Depuis le début des années 2010, **Google a significativement accéléré le recours à l'IA pour coller au plus près de l'intention de recherche, en SEO comme en SEA.**

Il faut dire que le NLP est une formidable opportunité de toujours mieux appréhender les subtilités du langage humain.

Dans les métiers du marketing, le travail de référencement vise à se rapprocher le plus possible des attentes de Google pour truster les premières places. Pourtant, il devient de plus en plus difficile de le faire : **avec au moins 10 modifications quotidiennes⁸, ses algorithmes se complexifient** à tel point que, selon le « pape du SEO » Oliver Andrieu, vouloir percer leur mystère relève à présent de la quête du Graal :



*Il est devenu depuis 2 ans absolument impossible de comprendre comment fonctionne l'algorithme du moteur. Car celui-ci ne réagira pas de la même façon selon l'intention de recherche, le type de support, l'historique de la requête, la géolocalisation, etc... Bref, ce ne sont pas les mêmes critères et les mêmes poids qui s'appliquent sur tel ou tel mot clé. Comprendre comment fonctionne l'algorithme est devenu **une quête clairement dépassée.***



Olivier Andrieu Consultant SEO et créateur de la société Abondance

Peu importe à vrai dire : pour assurer un référencement optimal, **restez concentré sur les 3 critères fondamentaux du SEO** : (1) la qualité de contenu (2) alimenté par des liens entrants ou backlinks pertinents (3) dans un écran technique irréprochable (sécurisé, rapide, indexable, ergonomique).

Et pour le faire du mieux possible, rien ne vous empêche de vous appuyer à votre tour, pour votre SEO comme pour votre SEA... sur l'IA. Bien au contraire !

⁸ <https://www.abondance.com/20200703-43155-en-2019-google-a-fait-3-620-changements-dans-son-algorithme-soit-10-par-jour.html>

2.1. SEO : l'IA au service de votre sémantique

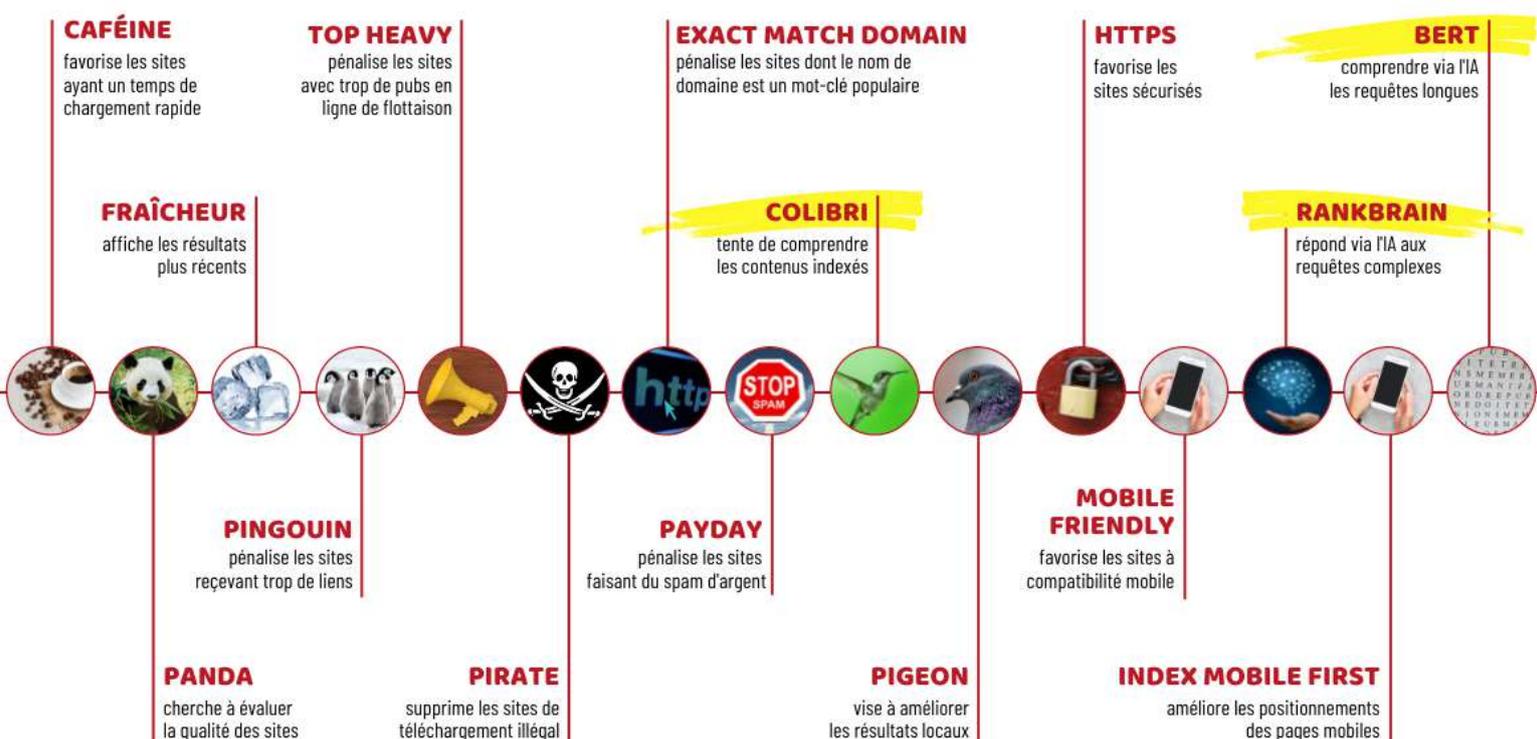
Si l'IA peut être un formidable accélérateur de votre création de contenus, la tentation de créer des articles à la chaîne, de façon totalement automatique, ne vous permettra pas de prétendre aux premières positions des résultats des moteurs de recherche.

Pour gérer un volume toujours plus croissant de contenus à indexer, Google a en effet su dimensionner ses algorithmes pour contrer les spammeurs, qui diffusent chaque jour pas moins de 25 milliards de nouvelles pages !

Faisons un **petit tour d'horizon de la façon dont l'IA a intégré les algorithmes de Google** au fil de la dernière décennie, avant de voir comment vous pourriez vous-mêmes tirer partie de l'IA pour optimiser votre SEO.

TOUJOURS PLUS D'IA DANS LES ALGORITHMES DE GOOGLE

Il n'existe pas chez Google un algorithme unique pour le SEO, mais une myriade d'algorithmes. Parmi eux, les plus récents ont recours au Machine et Deep Learning :



Ses algorithmes fondateurs (Page Rank, Panda, Pingouin) n'ont bien sûr pas été éludés par ceux carburant à l'IA. Mais **depuis le déploiement de Colibri (Hummingbird), l'IA est devenue indispensable chez Google pour répondre du mieux possible aux recherches d'internautes qui s'expriment de moins en moins sous la forme de mots-clés :**



*Les utilisateurs ont tendance à **formuler leurs requêtes en s'adressant à la machine non plus avec de simples mots mais en faisant des phrases entières**. On parle déjà à nos portables et nos enceintes connectées comme si on s'adressait à des humains.*



Le traitement du langage naturel est donc devenu un impératif pour les moteurs.

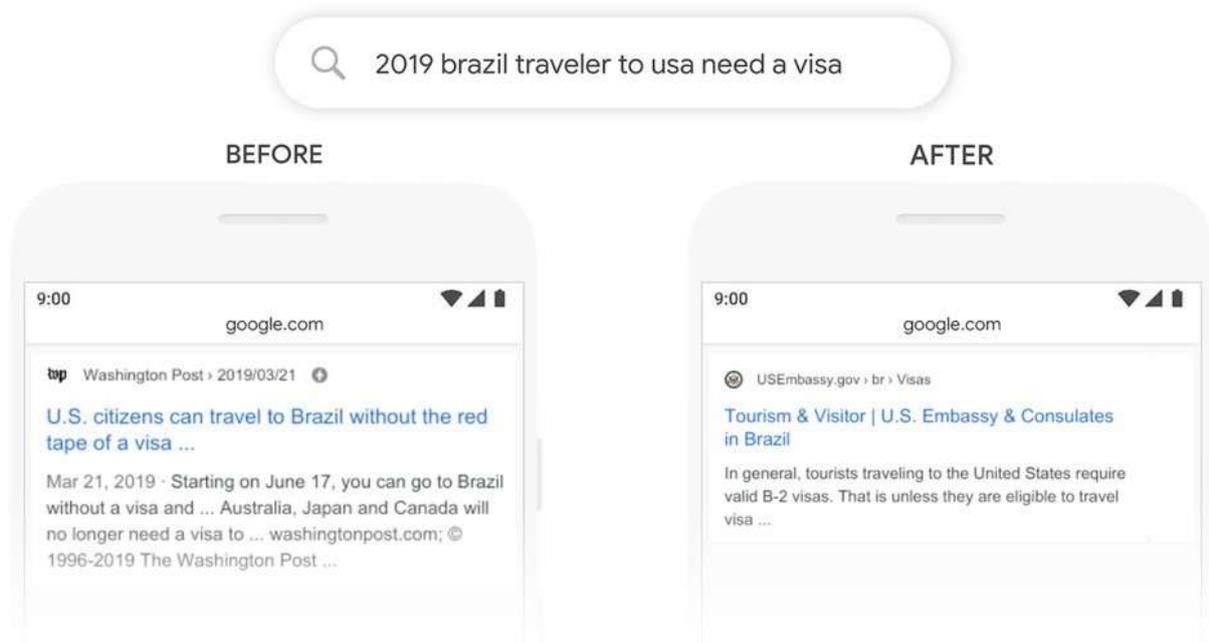
Muriel Griveaud Responsable SEO chez Affielou et membre du collectif ia4marketing.fr

À l'heure des phrases complètes tapées dans le moteur de recherche, ou énoncées à présent via des interfaces vocales, **la longue traîne regorge d'opportunités** pour vous positionner stratégiquement, et à moindres coûts.

D'autant que depuis Colibri, Google a **poussé encore plus loin le recours au NLU** (*Natural Language Understanding*) **et au NLP, avec Rankbrain et BERT**, pour à la fois affiner sa compréhension des requêtes les plus complexes, mais aussi mieux saisir le sens des pages vers lesquelles il redirige et ainsi assurer le « match » parfait.

Prenons un exemple concret ⁹, montrant comment Google réussit à toujours mieux capter l'intention de recherche :

⁹ <https://blog.google/products/search/search-language-understanding-bert/>



Avant, l'algorithme de Google ne parvenait pas à interpréter des **mots-clés dans leur contexte, sur des requêtes longues et complexes**. Au risque, comme ici, de comprendre l'inverse de ce qui était recherché. **Avec Rankbrain et BERT, l'intention de recherche est à présent bien cernée** : la requête porte sur un voyageur brésilien qui se renseigne sur la nécessité d'avoir un visa pour aller aux États-Unis, et non d'un voyageur américain qui souhaiterait se rendre au Brésil.

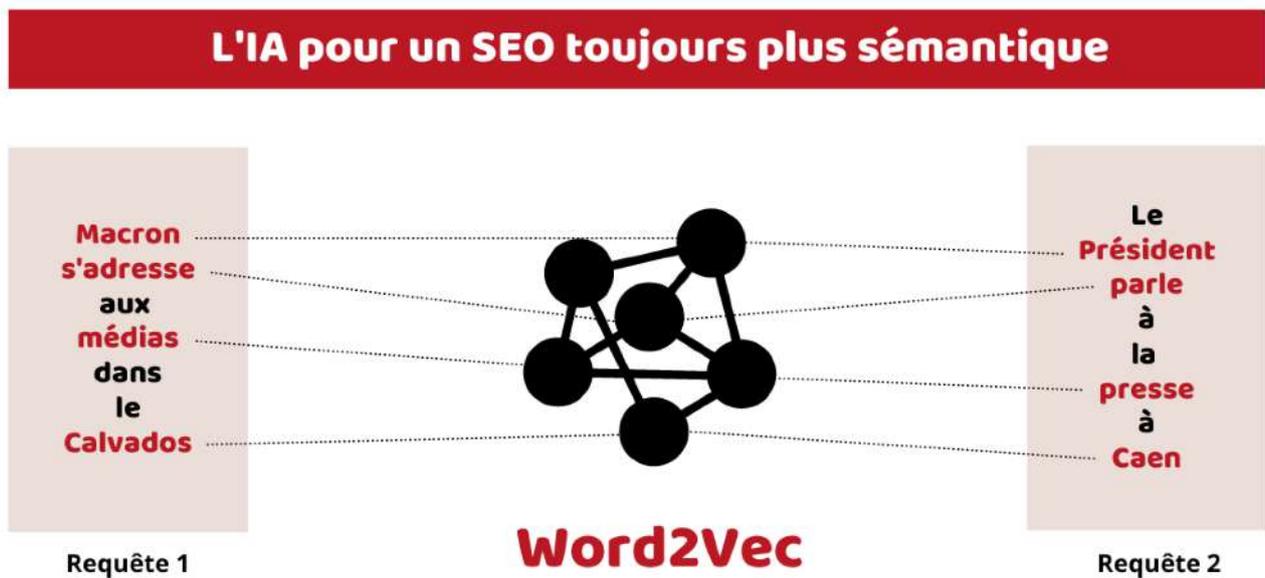
Sans l'IA, Google n'aurait pas su répondre correctement aux 15% de recherches tapées chaque jour ¹⁰... qui n'avaient jamais été recherchées auparavant ! Toutes ces requêtes qui tournent autour du pot sans utiliser le terme exact, le bon « mot-clé »... et toutes ces requêtes qui peuvent avoir plusieurs significations : « acheter souris » fait-il référence à l'animal, à la souris d'ordinateur, à la souris d'agneau... ?

Grâce au Deep Learning, Rankbrain apprend par lui-même des interactions passées des internautes avec les résultats proposés. Il prend en compte le taux de clic (CTR), le

¹⁰ <https://www.weblife.fr/breves/google-15-de-nouvelles-recherches-tous-les-jours>

taux de rebond, ou encore le temps passé sur la page. Par cette approche, il affine sa compréhension de ses bonnes et mauvaises réponses, pour corriger le tir par la suite.

Le réseau de neurones Word2Vec lui permet d'opérer des associations de termes entre eux, qu'ils soient synonymes ou antonymes, **pour rapprocher des requêtes similaires formulées différemment :**



© Denis Lefebvre

En somme, la prédiction par l'apprentissage, les combinaisons de chaque mot-clé entre eux incluant leur contexte d'usage, font évoluer la pertinence du moteur de Google vers un SEO toujours plus... sémantique !

Voyons comment vous pouvez vous adapter efficacement à ce nouveau paradigme.

COMMENT EXPLOITER VOUS-MÊMES L'IA POUR VOTRE SEO ?

Vous l'avez compris, avec l'avènement de l'IA dans les algorithmes de Google, comme d'ailleurs dans celui d'autres moteurs de recherche, **votre stratégie SEO ne peut plus tourner autour d'un mot-clé unique pour chacune des pages de votre site.**

Comme Google appréhende bien mieux qu'auparavant des requêtes utilisateurs riches sémantiquement, **vous devez vous aussi penser multi-requêtes en connectant à votre mot-clé principal un nuage sémantique complet**. Vous multiplierez ainsi les chances que le moteur d'indexation fasse le pont entre votre contenu et la requête.

Pour y parvenir, faites vous-mêmes appel à l'Intelligence Artificielle !

Il ne s'agit pas d'utiliser l'IA pour chercher à manipuler les algorithmes : comme nous l'avons dit, des contenus auto-crédés à la chaîne par l'IA n'apporteront pas une qualité à même de bien vous faire référencer ; pire, si vous en abusez sans y injecter de la valeur, vous pourriez bien vous faire pénaliser.

En revanche, vous pouvez **vous aider de l'IA pour être meilleur que vos concurrents, en rendant la recherche de vos mots-clés plus efficaces, et votre stratégie d'optimisation pour les moteurs moins aléatoire**.

Prenons l'exemple de l'outil Topic, que nous avons évoqué et qui vous accompagne dans la rédaction de contenus SEO-friendly. Sa solution présente deux facettes :

- **Content Research** pour identifier en amont de votre rédaction :
 - **la performance des autres sites sur un sujet donné** (ici j'ai entré « ai carbon footprint ») : position dans la SERP, autorité du domaine, qualité du contenu et de son architecture ;
 - **les questions auxquelles répondre autour de ce sujet**, regroupant toutes les questions listées par People Also Ask, Google Suggest et Quora ;
 - **les mots-clés à couvrir sur ce sujet**, regroupés par similarité sémantique, et classés selon leur pertinence et leur positionnement concurrentiel ;
 - **les backlinks les plus fréquemment utilisés**, et l'autorité des domaines qui en sont la source.

Vous pouvez alors solliciter les différents aspects de cet audit express de votre sujet, et l'IA de Topic vous générera **automatiquement un brief de rédaction** :

The screenshot displays a user interface for generating content outlines. On the left, a panel titled "Get Outline Ideas" shows a list of potential topics. Some are highlighted in green with a checkmark, indicating they are recommended. The main panel on the right provides a detailed brief for the selected topic, "The Carbon Footprint Of Artificial Intelligence".

Get Outline Ideas

Here is an idea for an outline. Refresh

- Why Should We Care About Artificial Intelligence's Carbon Footprint? ✓
- How Much Energy Do Computers Consume?
- How Much Energy Does Artificial Intelligence Consume?
- The Carbon Footprint Of Data Centers ✓
- The Carbon Footprint Of Supercomputers
- The Carbon Footprint Of Autonomous Cars
- Can We Reduce Artificial Intelligence's Carbon Footprint? ✓

Add 3 Headings

56 959 B+

Title Rewrite Title
The Carbon Footprint Of Artificial Intelligence
47/60

Description Rewrite Description
Artificial intelligence makes life more convenient, but it also has an environmental cost. Discover the carbon footprint of AI here.
132/160

Word Count 1500
longer than 63% of competitors

Outline Get Outline Ideas

- H2 Increasing The Size Of An Ai System Can Increase The Energy Usage By 1,000,000%
- H2 How Much Energy Does Ai Use?
- H2 How Does The Carbon Footprint Of Ai Compare To Other Sectors?
- H2 How Is The Carbon Footprint Of Ai Being Reduced?

Copy to Clipboard

- **Content Grader pour améliorer en temps réel votre rédaction d'un point de vue SEO**, ou pour réviser vos anciens contenus et y ajouter des mots-clés propres à l'univers sémantique que vous devez plus largement couvrir. Dans l'exemple ci-dessous, l'outil vous incite à étudier la possibilité d'ajouter des mots-clés parmi ceux qui ne sont pas en vert dans la colonne de droite : ils l'ont été sur d'autres sites parlant du même sujet que vous, et leur position en est par conséquent meilleure.

topic Brief » ai carbon footprint created February 19, 2021 - United States (google.com) Chat with us

Content Research Content Grader

680 Words 1500+ Recommended



“Deep learning” algorithms are usually evaluated in terms of accuracy—but some researchers say they should also be judged on their environmental impact.

Artificial intelligence is an increasingly important element of science, medicine, and even the minutiae of our daily lives. Chatbots, digital assistants, and movie and music recommendations from streaming services all depend on “deep learning”—a process by which computer models are trained to recognize patterns in data.

That training requires powerful computers and lots and lots of energy—and associated carbon emissions. One of the most elaborate deep learning models, designed to produce human-like language and known as GPT-3, requires an amount of energy equivalent to the yearly consumption of 126 Danish homes and creates a carbon footprint equivalent to traveling 700,000 kilometers by car for a single training session.

Current Grade: C
Recommended: A
Cover 1 more topic to reach next grade

Topics to Cover

- ai (used by 27 sources) H
- algorithms (used by 17 sources) H
- computer science (used by 7 sources)
- carbon footprint (used by 21 sources) H
- pounds of carbon dioxide (used by 5 sources)
- amount of co2 (used by 4 sources)
- artificial intelligence (used by 24 sources) H
- machine learning (used by 18 sources) H
- neural networks (used by 15 sources)

D'autres solutions sur le marché peuvent vous indiquer les meilleurs mots-clés à utiliser en temps réel. Yahoo par exemple utilise un **programme d'IA pour identifier des mots-clés tendances qui révèlent des idées de contenus frais à exploiter** dans ses actualités.

Grâce au social listening, les mots-clés conversationnels peuvent être saisis et servir de signaux faibles pour les rédacteurs web et les copywriters. Des outils comme Twinword permettent ainsi de visualiser les mots-clés tendances du moment, sur une thématique précise ou plus générale.

twinkl ideas How It Works Pricing FAQ Blog Log In Sign Up Free

SUGGEST IMPORT TRENDING LSI GRAPH ADDED

marketing EV Q

Graph List



Trending Keywords: See trending related keywords and their relationships.

Recent Tweets

- So now Shea Butter Twitter is mad at the B&B Owned Shea Moisture hair care company for expanding their marketing to other races? #seriously @ariqashceci
- An ad that erases the ongoing discrimination based on hair textures & skin color faced by their core customers isn't just bad marketing @Karynthia
- The Ticking Time Bomb is a necessary condition for all war efforts - whether or not the Trump regime uses this marketing collateral is unclear @adawejohnsonNWC
- What will have the biggest impact on #marketing? #AI #AR #VR #MR #IoT #Livestreaming #DigitalMarketing #CMS #WUXGrowth #ContentMarketing @GierGilmore
- I'm thrilled to be joining the WSJ to cover marketing and media, and write the DMG Today newsletter. I start next week. @larakiars
- Guys, just maybe don't tell women who say they wish there was more Wonder Woman swag #marketing to stop being dramatic... just... do... not... @boddestmamajama
- Every aspect of business today in some way or other centers around marketing &

Terms of Service Privacy Policy About Us Support First semantic keyword research tool that can sort by relevance.

Des outils SEO alimentés à l'IA vont au-delà de la seule recommandation sémantique pour **assister votre référencement à 360°**, en analysant tous les paramètres de votre site (vitesse de chargement, validation des normes W3C...). Citons Scrapy ou Squirrly SEO, qui **filtre grâce au Machine Learning des milliards de points de données pour identifier ce qui a permis à des centaines de milliers de sites d'obtenir les meilleurs classements.**

Google favorisera d'autant votre référencement que vous vous montrerez **toujours plus user-centric**. Car c'est là que se dessine **l'avenir du SEO**, selon Olivier Andrieu :



Les moteurs évolueront avec les usages du Web, et le SERP évoluera en conséquence. Le juge de paix, dans ce domaine, ce sont les usages.



C'est l'internaute qui, finalement, pilotera le SEO de demain !

Olivier Andrieu Consultant SEO et créateur de la société Abondance

Au point de finir par voir un jour émerger un moteur de recherche unique pour chaque internaute, avec des résultats correspondant à ses parcours de navigation, appétences, au-delà de ceux aujourd'hui délivrés en commun au plus grand nombre ?

2.2. SEA, Display, Social Ads : ciblez juste et au meilleur prix

Pour appuyer votre référencement naturel et maximiser votre potentiel d'acquisition, vous allez aussi déployer des campagnes publicitaires en SEA, Display et Social Ads.

Pour optimiser votre ROAS, tirez partie de l'Intelligence Artificielle pour vous accompagner... de la planification à l'achat ! Avec le Machine Learning et le Deep Learning, les méthodes publicitaires évoluent pour un ciblage toujours plus pertinent.

UN CIBLAGE PUBLICITAIRE TOUJOURS PLUS INTELLIGENT

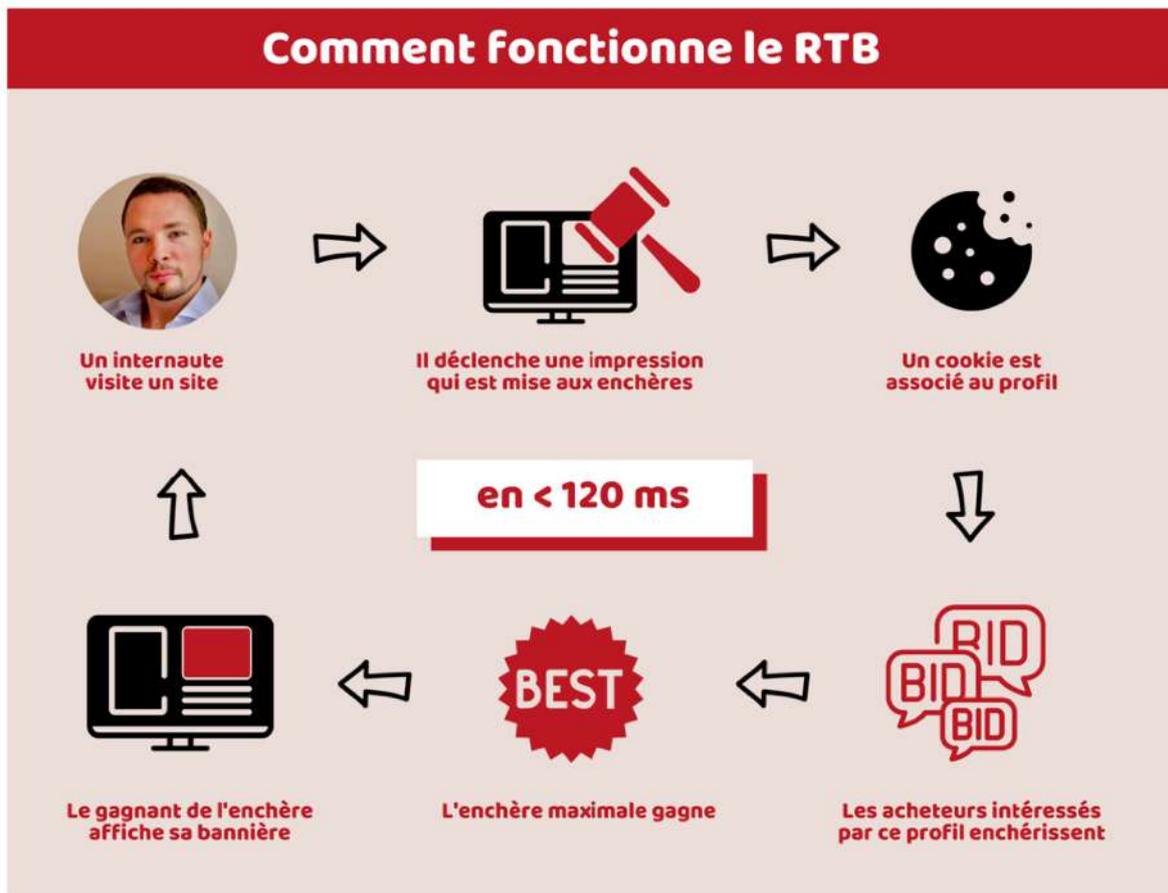
Real-Time Bidding : cibler une audience plutôt que des sites

Depuis les années 2000, la publicité programmatique s'impose comme l'approche la plus efficace pour déployer ses campagnes. Il représente à présent 61% des recettes Display, en croissance de +4% en 2020 par rapport à 2019 ¹¹.

La **publicité programmatique** permet en effet d'**automatiser l'achat/vente d'espaces sur une logique de cibles plutôt que de sites**. Aux parts de voix et aux projections en CPM (Coût Par Mille), les annonceurs tendent à privilégier l'**achat impression par impression, bien plus ROIste**.

Le **Real-Time Bidding (RTB)**, qui est l'un des formats maîtres du programmatique, permet de diffuser votre publicité aux internautes susceptibles de s'y intéresser, selon un **système d'enchères automatiques** exécuté en moins de 120 millisecondes.

¹¹ 25ème observatoire de l'e-pub, SRI et UDECAM, 2020 : <https://www.sri-france.org/2021/02/02/obsepub-2020-marche-francais-de-pub-digitale-croissance/>



© Denis Lefebvre inspiré par Gannet

Les algorithmes ont donc investi le secteur de la publicité digitale pour optimiser votre ciblage. **Ce qui importe aujourd'hui, ce n'est plus qui est l'internaute, son profil** à tenter de cerner à travers un *persona* : **ce qui importe, c'est la façon dont il se comporte à un instant T sur votre site.**

Ainsi le RTB saura afficher votre publicité pour votre cible, fan de boxe, sur un site de tricot, alors que vous ne l'auriez pas imaginé aimer ce type de crochets...

L'approche programmatique a révolutionné le marché de la publicité digitale, tant dans la pertinence de l'identification de vos cibles que de la diffusion de vos messages. Pour une expérience utilisateur toujours optimisée, et un meilleur ROAS des annonceurs, **les plateformes publicitaires ne cessent depuis d'améliorer leurs algorithmes.** Au point de limiter toujours plus le besoin de ciblage dans la préparation de vos campagnes.

Laissez faire l'IA pour votre ciblage publicitaire

Le retour d'expérience de Jeremy Bendayan ¹², COO et co-fondateur de Splashr (agence de *snack content* vidéo) est édifiant : **les résultats des campagnes Ads sur Facebook et Instagram sont bien meilleurs lorsque l'annonceur ne précise aucun ciblage**, sauf peut-être la tranche d'âge visée.

Simplement **parce qu'en réduisant votre segment, vous le sur-sollicitez** sur une même catégorie de produits/services vis-à-vis d'autres annonceurs, **en explosant au passage votre CPA** (Coût Par Acquisition). Être restrictif dans votre ciblage publicitaire réduit aussi les chances que votre publicité puisse être adressée à de meilleures cibles.

L'heure est donc au « no-targeting » : **privilégiez une audience la plus large possible, en laissant l'IA dénicher ceux qui auront le plus d'intérêt pour votre offre.**

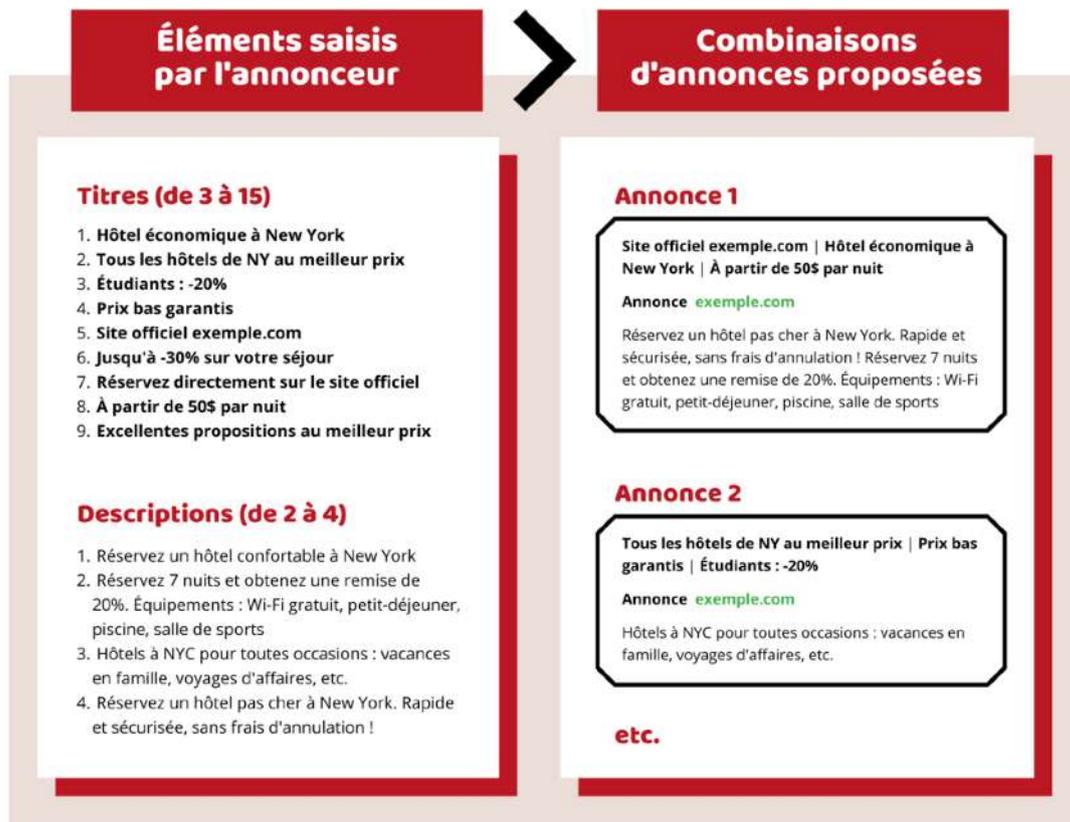
A minima, vous pouvez indiquer les cibles à exclure, et/ou fournir de la donnée utile à l'algorithme tel que votre top clients : des audiences similaires (« look alike ») pourront ainsi être créées, et s'adapter aux interactions enregistrées pour identifier et cibler de nouveaux leads par l'apprentissage automatique.

D'ailleurs, **Google vous décharge de plus en plus dans la création et la diffusion de vos campagnes SEA**, comme c'est aussi le cas en SEO en reprenant plus de deux tiers de vos méta-descriptions ¹³ pour les adapter au mieux à la variété des recherches. En SEA :

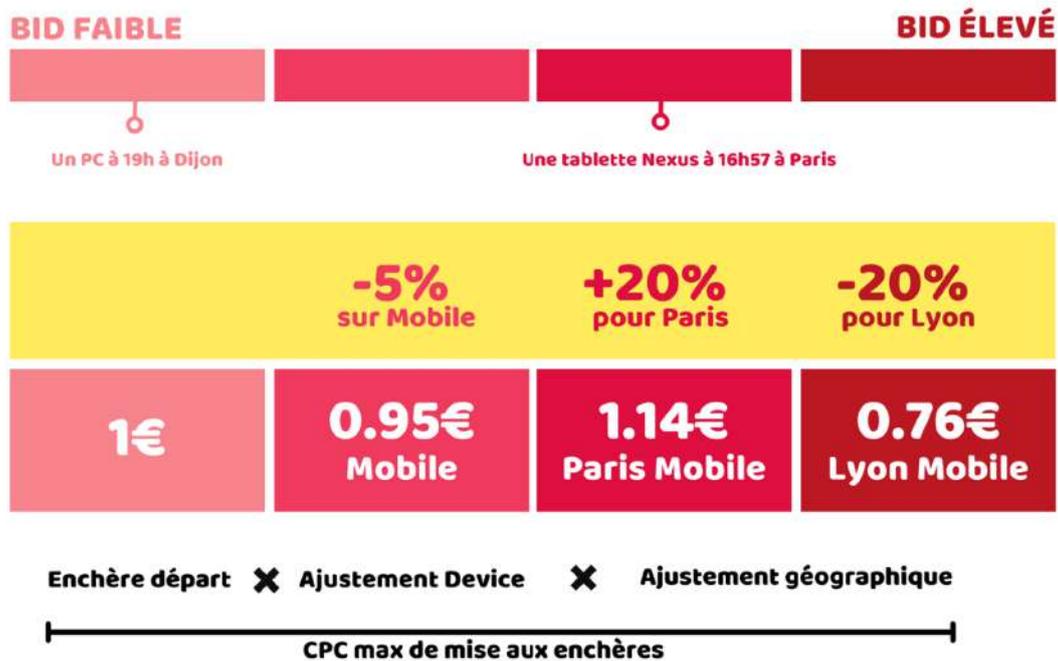
- Google pousse à utiliser les **Responsive Search Ads** en lieu et place des Extended Ads : vous entrez jusqu'à 15 titres et 4 descriptions, et vous vous voyez proposer **jusqu'à 90 variantes d'annonces rédigées par un algorithme**. Plus besoin d'écrire vos 3 annonces, ni de multiplier des A/B tests manuels : **Google vous indique les combinaisons idéales** pour sensibiliser votre audience et augmenter votre CTR (Click Through Rate).

¹² <https://www.linkedin.com/feed/update/urn:li:activity:6489143491032018944/>

¹³ <https://www.abondance.com/20200423-42592-google-reecrit-la-balise-meta-description-pour-ses-snippets-dans-2-3-des-cas-etude.html>



- Google injecte également depuis 2019 de l'intelligence supplémentaire au RTB. Le **Smart Bidding** utilise le **Machine Learning** pour effectuer des enchères en temps réel selon un certain nombre de signaux détectés grâce au profil du compte Google sur lequel l'internaute est connecté, **et par une approche prédictive** pour déterminer des similarités de profil ou de navigation. Parmi ces signaux figurent l'appareil, l'âge, le sexe, l'emplacement, le jour ou bien l'heure de la journée... Vous pouvez **pondérer vos niveaux d'enchères** sur chacun de ces critères de -100% (pour ne pas être pris en compte, par exemple sur mobile si votre site mobile n'est pas optimisé), à un CPC max pouvant aller jusqu'à +900% (pour faire peser un critère clé). L'exemple ci-après montre l'intérêt de la multiplication de ces facteurs. Au moins 30 conversions historiques sont nécessaires pour entraîner l'algorithme et le mettre en œuvre sur votre compte Google Ads. Vous pourrez dès lors **laisser au Smart Bidding le soin de prendre automatiquement pour vous les meilleures décisions**. D'après Google, **jusqu'à +30% de conversion** avaient pu être enregistrées sur la version bêta de son Smart Bidding.



© Denis Lefebvre inspiré par Séverine Schotus

Ainsi, l'IA fait passer l'automatisation propre à la publicité programmatique à une autre échelle. Mais elle permet aussi de s'adapter au paradoxe d'internautes qui souhaitent bénéficier d'une plus grande personnalisation tout en protégeant la confidentialité de leurs données personnelles. Pour réconcilier ce paradoxe, nous assistons à présent grâce à l'IA à l'émergence de la publicité prédictive et contextuelle.

DU PROGRAMMATIQUE AU PRÉDICTIF DANS UN MONDE COOKIELESS

L'achat programmatique peut utiliser des modèles de propension générés par des algorithmes de Machine Learning ou de Deep Learning pour cibler plus efficacement vos annonces publicitaires, à l'heure de la fin programmée des cookies tiers (à partir de 2023) et du ciblage comportemental.

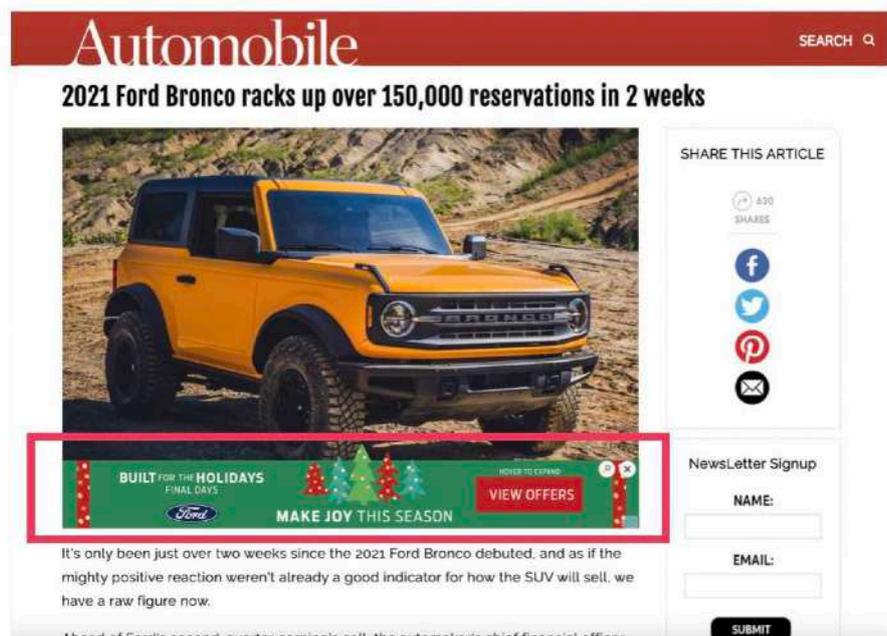
Publicité contextuelle : l'exemple de GumGum



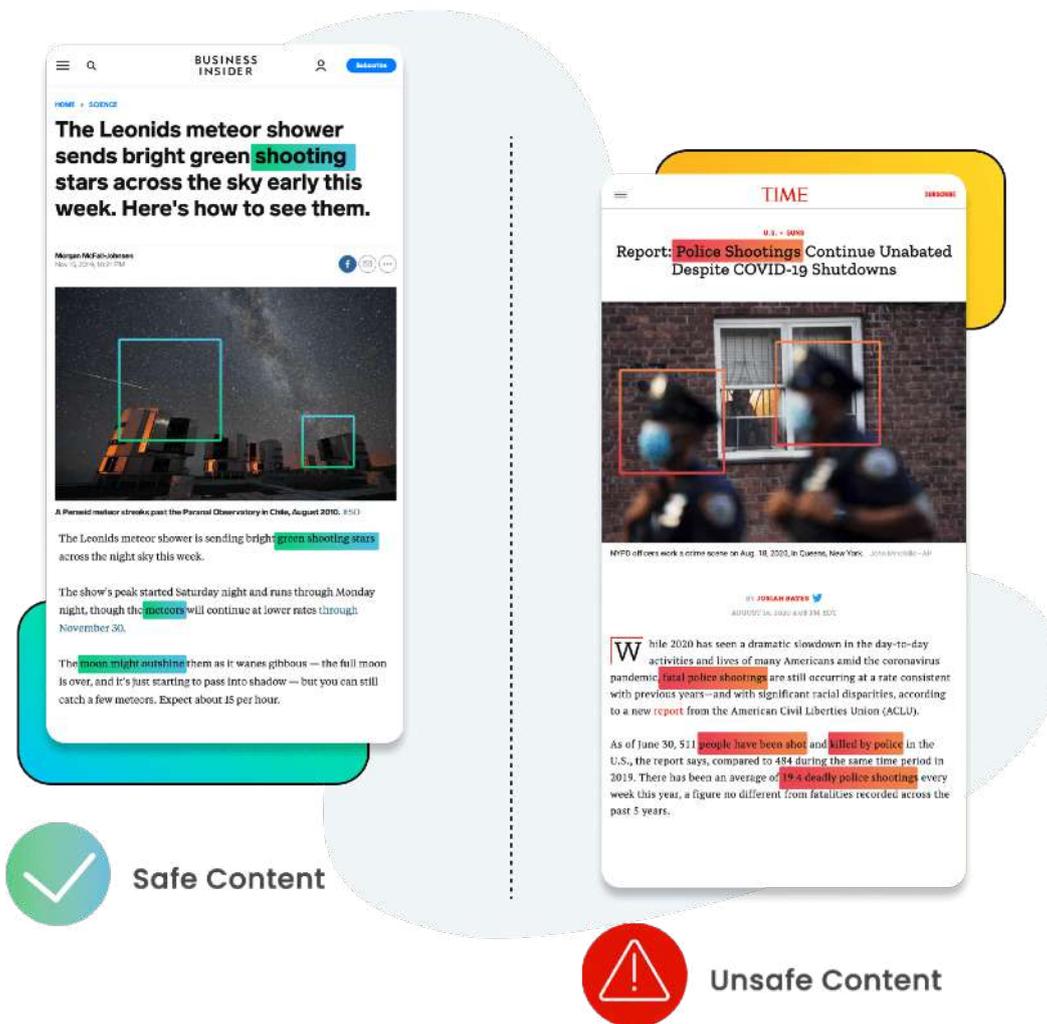
GumGum a développé une **plateforme d'intelligence contextuelle capable d'analyser en temps réel tout type de contenus** (textes, images, vidéos, sons), **et d'évaluer les meilleurs emplacements où placer votre publicité.**

Très concrètement, la solution GumGum a **recours à la fois à l'analyse automatique du langage et au Computer Vision** pour :

- **Positionner votre publicité de façon pertinente** : par exemple votre vente de pots de miel apparaîtra dans le fil d'un article média parlant des abeilles. La publicité peut apparaître en bandeau persistant sur le bas voire autour de l'écran (In-Screen Expandable/Frame), ou dans l'image illustrant l'article lui-même (en bas de l'image en In-Image Expandable, ou encore plus impactant en In-Image Canvas). Tous les formats et exemples peuvent être consultés dans cette galerie.



- **Éliminer les emplacements jugés inappropriés par l'identification et l'association d'indices visuels et sémantiques**, comme sur l'exemple ci-dessous : alors qu'à gauche, le terme « shooting » analysé dans son contexte renvoie en fait aux étoiles filantes, à droite il renvoie bien à des tirs, confirmé par l'identification de policiers sur la photo. Or, vous ne voulez pas associer votre publicité sur un article de ce type, et ce n'est pas le moment idéal pour sensibiliser l'internaute.



Un récent scandale ¹⁴ sur la sécurité des marques a atteint Google, qui avait diffusé des annonces en programmatique sur des sites terroristes. **L'IA permet aux annonces programmatiques de devenir plus intelligentes, en identifiant les sites douteux et en les supprimant de la liste des inventaires sur lesquels placer des annonces.**

¹⁴ <https://www.latribune.fr/technos-medias/publicite-pourquoi-des-annonceurs-boycottent-google-et-youtube-669530.html>

En somme, GumGum est une **AdExchange intelligente et sécurisée** qui permet aux annonceurs d'analyser les données contextuelles des sites de ses éditeurs partenaires grâce au Deep Learning, pour **filtrer l'inventaire et maximiser la portée et les clics selon la nature de leur activité et de leurs préférences**.

Plus de problème de RGPD ni de cookies tiers, puisque l'approche est basée avant tout sur la pertinence du contenu pour accueillir votre publicité, avec des résultats plus que convaincants selon GumGum :

- **ROI x3,5**
- **Taux d'engagement x7**
- Publicités 70% plus visibles
- +37% de gain de notoriété

Reste que cette approche contextuelle est **content centric et non user centric**. Certes, si un internaute lit un article sur un sujet particulier, c'est que ce sujet l'intéresse, et que votre publicité a des chances de tomber à point nommé.

Mais **comment anticiper ou créer le besoin autour de vos produits avant même qu'il ne les connaisse ?** Comment devenir l'annonceur providentiel qui adresse le bon message au bon moment ?

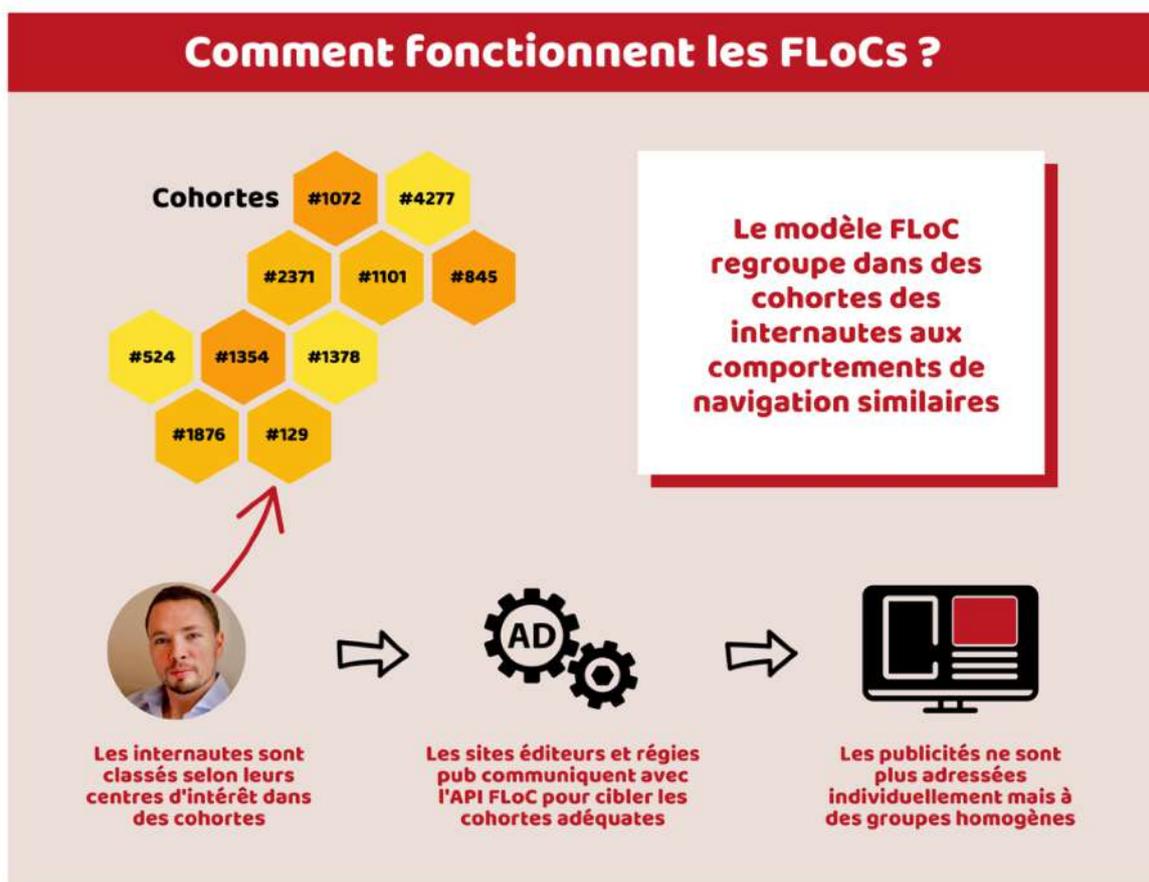
Prédiction et Personnalisation Publicitaire

L'un des principaux apports de l'IA dans le marketing, comme nous le verrons dans la partie suivante, est la possibilité d'**anticiper les comportements des internautes pour prédire leur intention d'achat**.

Le **ciblage prédictif** est basé sur l'utilisation de cookies pour détecter, sur des sites tiers, des comportements types proches de ceux de vos clients cibles. Vous pourrez ainsi adresser vos publicités aux personnes les plus susceptibles de s'intéresser à votre offre, et **recruter de futurs clients avant même qu'ils ne vous connaissent**, élargissant votre zone de chalandise au-delà des visiteurs ayant déjà visité votre site.

Certes, la méthode du retargeting continue de représenter en moyenne entre 10 à 20% d'un budget SEA. Elle a encore de beaux jours devant elle, vu que seuls les cookies *first-party* vont être épargnés. Mais elle montre aussi ses limites notamment en terme de pression commerciale. **Si la méthode prédictive semble être la nouvelle voie de la personnalisation, elle doit toutefois revoir son approche face à la fin des cookies tiers.**

Google propose une voie de sortie via sa « Privacy Sandbox » et **la méthode des FLoCs** (*Federated Learning of Cohorts*, ou apprentissage fédéré par cohortes).



© Denis Lefebvre inspiré par Google

Le principe est de créer des groupes (« cohortes ») contenant chacun au moins 5000 utilisateurs, et **présentant des comportements de navigation similaires** (ces données étant capturées par les navigateurs eux-mêmes).

Puis de **permettre aux annonceurs d'adresser des publicités pour telle ou telle cohorte selon leur centre d'intérêt**, sans que les internautes qui le constituent ne puissent être

identifiés personnellement. Cela ne vous permettra pas de faire du 1:1, mais plutôt du 1-to-few. Google affirme néanmoins que les annonceurs généreront au moins 95% des conversions faites jusqu'alors, avec le ciblage par cookies tiers.

L'IA permettra de ne pas non plus retomber dans une segmentation pré-déterminée, qui exclut d'emblée une audience potentiellement concernée par votre message. Des startups AdTech comme **AntVoice** relèvent actuellement le défi d'adapter leur solution d'IA prédictive pour cerner plus finement, au sein de ces cohortes, les intérêts réels des internautes.

Car selon AntVoice ¹⁵, le principal problème des FLoCs sera le manque d'attribut clair attribué aux cohortes : « est-ce que ce FLoC a été attribué à des amateurs de pêche à la ligne, ou à des personnes focalisées sur le développement durable ? ».

Les algorithmes de AntVoice devraient donc permettre d'**affiner la connaissance des FLoCs tout en restant « Privacy by Design »**. Un bon compromis pour mettre d'accord les 50% d'internautes qui, selon une étude Kantar ¹⁶, jugent la publicité personnalisée trop intrusive... et l'autre moitié qui s'y montre favorable.

Bravo ! Vous avez réussi à générer davantage de trafic grâce à l'Intelligence Artificielle, pour produire davantage de contenus, mieux référencés, et optimiser dans leur ciblage et leur ROAS vos campagnes publicitaires.

Mais **à présent que ces nouveaux visiteurs sont arrivés sur votre site, le risque est qu'ils le quittent sans avoir converti**. Vous devez agir pour les retenir !

Bonne nouvelle... **l'IA va vous accompagner tout au long de votre funnel, et vous aider à bien mieux convertir**.

¹⁵ <https://www.antvoice.com/fr/que-va-ton-faire-chez-antvoice-quand-il-ny-aura-bientot-plus-de-cookie-tiers/>

¹⁶ <https://www.larevuedudigital.com/personnalisation-des-messages-publicitaires-un-defi-inutile/>

L'IA pour booster votre Conversion

Imaginez que vous puissiez prédire quels sont vos visiteurs qui ont le plus de chance de convertir. Mieux : imaginez que vous sachiez comment les convertir, avec quels leviers spécifiques à même de leur faire franchir le pas.

Vous pourriez ainsi maximiser votre potentiel de vente à la fois sur vos leads, mais aussi sur vos clients actifs et inactifs. **Un vrai super-pouvoir.**

Mais ce n'est pas dans le mare de café ou dans la boule de cristal que vous pourrez le voir. La technique du doigt mouillé n'a pas non plus fait ses preuves.

Non... vous allez pouvoir **vous appuyer sur un élément bien plus tangible : les données.** Les vôtres, déjà récoltées dans votre base CRM, mais aussi celles que vous allez extraire des profils et comportements de vos visiteurs. **En automatique. En temps réel.** Pour une segmentation et une personnalisation du parcours d'achat sans précédent, permise une fois encore par l'Intelligence Artificielle.

Après l'ère du marketing digital, bienvenue dans l'ère... du marketing prédictif.

1. Devenez omniscients pour convertir vos visiteurs

Jusqu'à présent, vous ne pouviez vous appuyer que sur l'historique de vos données pour segmenter votre base et mettre en place des campagnes ciblées. Si vous voulez faire une campagne ciblée pour fidéliser vos clients ayant dépensé plus de 100€ sur les 3 derniers mois, une segmentation manuelle reste bien sûr efficace.

Mais l'IA va vous permettre d'**accroître considérablement votre connaissance client, et multiplier votre potentiel de conversion.** Car le secret pour gagner de nouveaux clients, c'est de les connaître tant et si bien que vous saurez exactement comment répondre à leur besoin.

1.1. Traquez vos leads les plus chauds

IDENTIFIEZ VOS VISITEURS AYANT LE PLUS DE CHANCE DE PASSER À L'ACTE

98% de vos nouveaux visiteurs vont quitter votre site sans avoir laissé d'empreinte

(création de compte, achat, email saisi dans un formulaire...) ¹⁷. Comment faire pour

identifier ceux qui ont le plus d'intérêt ? Certains pourtant consultent longuement vos pages, ils reviennent plusieurs fois depuis divers devices... mais toujours sans rien faire.

Manifestement il se passe quelque chose. Il y a un intérêt, une affinité. **Pour les aider à franchir le pas, il ne manque certainement qu'un petit déclic** qui pourrait faire toute la différence : une promo ciblée, un produit qui correspond à ce qu'ils recherchent, etc.

Pour évaluer finement cet intérêt, l'élément déclencheur le plus pertinent, et le support sur lequel l'adresser, il faut pouvoir identifier puis qualifier votre lead. Mais vous n'avez encore aucune donnée sur lui. Or **pas de data, pas de chocolat...** l'IA ne carbure qu'à la donnée, et il faut nourrir la bête si vous voulez qu'elle puisse vous aider.

Identifier les décideurs est d'autant plus décisif en B2B. Pour résoudre ce problème, la société française **GetQuanty a développé une solution de scoring prédictif permettant aux entreprises B2B d'identifier et qualifier les visiteurs de leurs sites restés anonymes.** Grâce à l'Intelligence Artificielle, **GetQuanty a réussi à passer de 15% à 70% de taux d'identification de ces prospects anonymes.**

Je vous invite à lire dans les pages suivantes **mon interview de Xavier Paulik, Directeur Général de GetQuanty**, pour savoir comment GetQuanty a réussi à atteindre ce seuil grâce au Machine Learning, et à multiplier par 10 le taux de transformation du Groupe PSA par rapport à sa prospection à froid.

¹⁷ <https://www.kameleoon.com/fr/blog/98-de-vos-visiteurs-sont-anonymes-comment-les-converter>



Xavier Paulik DG de GetQuanty



Comment GetQuanty utilise l'intelligence artificielle pour qualifier les prospects et les retargeter efficacement, à partir de leur simple visite sur un site web ?

« Le fondateur historique, Hervé Gonay, avait fait tout un travail de création d'une base de données associant les adresses IP des visiteurs d'un site web à des informations entreprises. C'était un peu la base de la base de données, mais à l'époque ça n'existait pas. Il a abouti à la première base de données en France un peu quali sur les adresses IP entreprises. Il a essayé d'exploiter cette information pour que des acteurs B2B sachent quelles entreprises venaient visiter leur site web. Jusqu'alors, seulement 1 à 2% d'entre elles pouvaient être identifiées grâce à des formulaires en ligne. **Avec notre solution, on arrivait à identifier environ 15% du trafic.** Ça permettait de leur dire « Telle entreprise revient régulièrement, tu devrais la contacter, elle ne le fait pas spontanément mais il y a manifestement un sujet chez elle ». Il restait des limitations fortes : des problématiques techniques, mais aussi un taux d'identification qui certes était meilleur, **mais pas encore assez suffisant pour qu'une entreprise base toute sa stratégie dessus.** Il fallait aussi lui permettre de trier tous ces prospects et identifier ceux qui ont un réel intérêt. **On a donc fait émerger puis affiné une notion de scoring** qui analyse les visites successives d'une entreprise pour révéler son réel intérêt. Car certains vont faire des visites longues en passant du temps pour savoir de quoi on parle. Et paradoxalement, plus un visiteur est aguerri, moins il restera longtemps et plus il sera une cible. Il fallait donc pouvoir analyser les visites sur la durée.

On a travaillé sur 2 plans sur lesquels l'IA est fortement entrée dans notre solution :



On est passé au Cloud et au Big Data avec Big Query (Google), qui permet de faire du traitement à large échelle, à coût très faible et très rapidement.

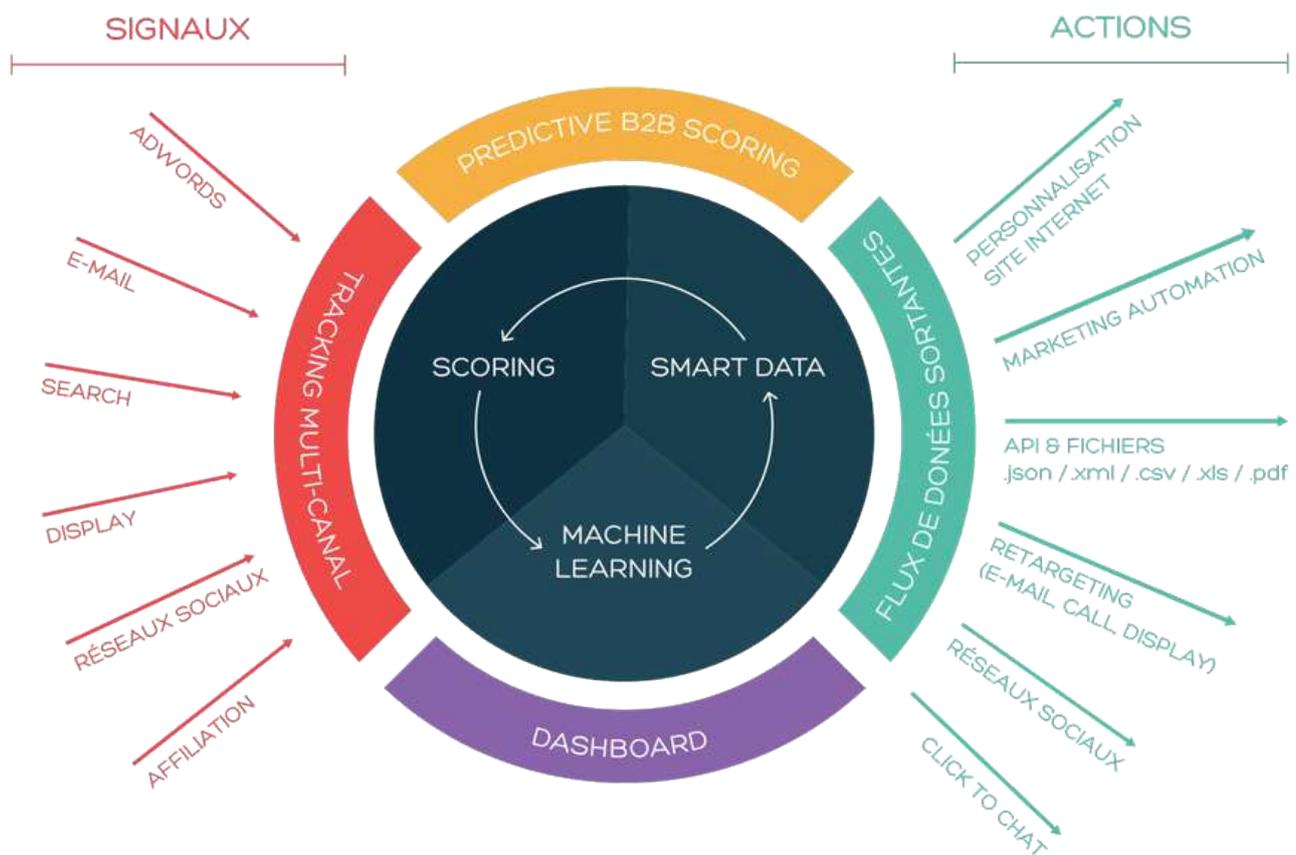
Google démultiplie massivement des traitements sur des milliers de serveurs. **On n'a plus jamais perdu de données**, alors qu'avant on était obligé de supprimer les logs de nos clients. Dès 2016, on avait une **masse critique de données pour commencer à faire tourner du Machine Learning**. On a recruté un *Data Scientist* et on a observé un an de data sur 2000 sites web. On est allé chercher dans ces données les **variables explicatives d'un intérêt des visiteurs pour nos clients, en les mappant sur une échelle de maturité**. On a ainsi créé nos **algorithmes de scoring**, en passant de l'interprétation d'une visite individuelle, qui n'a pas de sens, à un scoring entreprise qui a énormément de sens, avec une dynamique de scoring qui permet de détecter non seulement son intérêt, mais aussi les moments où son intérêt s'accélère et où il faut entrer en contact avec elle. C'est là où le lead devient « chaud ». On a aussi appris à communiquer sur les informations à donner à nos clients, pour **donner confiance à nos scores prédictifs**.



On a inventé d'autres algorithmes pour étendre notre base de données d'adresses IP, et passer de 15% jusqu'à 70% de taux d'identification.

On a du faire un gap quantitatif et qualitatif très important sur la capacité d'identification **pour pouvoir déployer notre offre commerciale à grande échelle**. Pour ça on a développé des techniques de collecte d'informations reliées par divers référentiels. L'adresse IP seule est une donnée trop volatile, et qui se périmé dans le temps. Si vous

changez de locaux, quand vous êtes en télétravail ou à chaque connexion sur mobile, vous changez d'adresse IP. On a donc **conçu des algorithmes qui vérifient et enrichissent en permanence la qualité de notre BDD adresses IP entreprises**. Et on arrive à gérer les transitions des entreprises vers d'autres adresses IP. Notre BDD n'est pas associée à une entreprise de manière définitive, mais à plusieurs entreprises avec des **probabilités d'affectations qui vont évoluer dans le temps**. On compte beaucoup sur l'**informatique quantique** qui maintient des infos simultanément vraies avec une notion d'incertitude, et qui vérifie des hypothèses au fur et à mesure qu'on acquiert de la donnée. Grâce à ça on gagnera encore en fiabilité.



GetQuanty permet aux entreprises B2B d'identifier et de qualifier les visiteurs de leurs sites restés anonymes. L'outil analyse en temps réel le comportement de ces visiteurs sur leur site, emailings, réseaux sociaux...

Sa solution de retargeting et de lead generation permet de créer des campagnes engageantes pour relancer les décideurs au bon moment avec le bon message.

On travaille aussi sur un **modèle « lookalike »**, dont la promesse est de **cibler des prospects avant même qu'ils ne viennent sur un site Web**. On a voulu donner à nos clients la possibilité d'aller chercher une **audience similaire sur l'ensemble de notre réseau**. Notre algorithme tourne sur les sites de nos clients et caractérise en permanence les thématiques qui sont recherchées. Puis on inverse la proposition pour **identifier les prospects B2B qui sont manifestement intéressés par ces sujets**, parce qu'ils ont visité 20 sites de notre réseau. Par matching on peut aller chercher prioritairement ceux qui, sur les 70 derniers jours, ont fait des recherches ciblées sur les mêmes thématiques. On mutualise ainsi l'effort pour tous nos clients, même si dans ce cas de figure ça reste de la prospection à froid vu que les leads prédictifs ne sont pas venus sur le site cible. »

Cas client



800

leads chauds par semaine

générés grâce à GetQuanty (+ autres leads).

× 10

sur leur transformation par rapport à leur prospection à froid (1%), au-delà des 7% du formulaire en ligne.

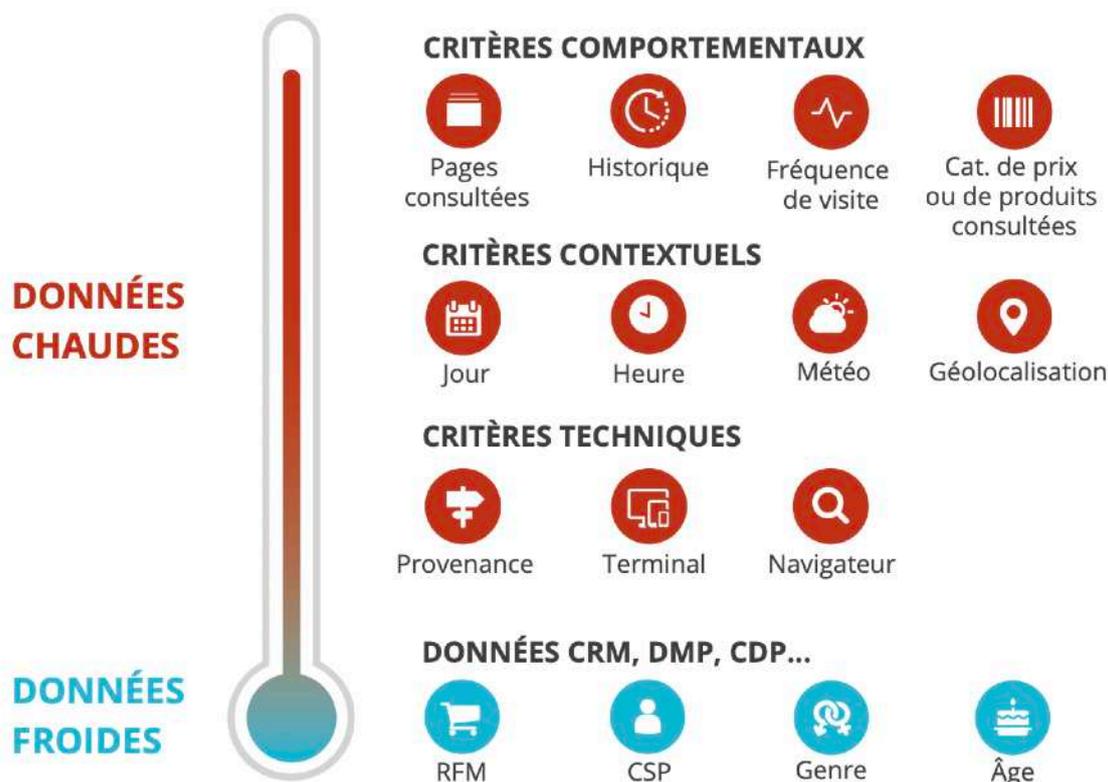
300

points de vente pourvus de commerciaux B2B utilisant la solution GetQuanty.

En B2B comme en B2C, il vous faut pouvoir **qualifier vos leads pour savoir lesquels sont pour vous des cibles prioritaires**. Pour cela, **apprenez à capter leurs « signaux faibles »** et à les rapprocher de ceux de vos clients existants. Grâce à la somme d'indices glanés ici et là, vous serez ensuite à même d'opérer **un ciblage affinitaire pertinent**.

Alors, de quels « signaux faibles » parlons-nous ?

Par opposition aux données « froides », qui sont celles déjà enregistrées dans votre base CRM, **les données « chaudes » s'analysent en temps réel pour déterminer l'appétence à une offre selon un ensemble de critères techniques, contextuels et surtout comportementaux :**



Toutes ces données sont récoltées puis croisées via **des milliers de scoring dynamiques**.

En opérant ces calculs à la volée, de façon fiable et précise, les algorithmes de Machine Learning révolutionnent la façon de **prédire l'intention d'achat**, vous permettant de proposer par anticipation une offre personnalisée pour convertir :

- **Des « scores d'appétence »** sont attribués aux visiteurs et clients identifiés comme ayant le plus d'intérêt pour votre offre. Si un visiteur arrive sur votre site après avoir tapé le nom de votre marque ou de l'un de vos produits, il n'a certes pas le même degré de maturité que s'il avait entré une requête générique.
- Ces scores enclenchent automatiquement **une segmentation prédictive** pour envoyer le bon message au bon moment, à la bonne personne, et par le bon canal.
- **Seuls les profils les plus susceptibles de convertir seront ainsi ciblés.** Pas d'impression de spam, de meilleurs taux de conversion : en visant juste, vous sur-optimizez votre ROI et votre relation client ! Dans un contexte B2B, vous pourrez armer vos équipes commerciales de toutes les données récoltées pour travailler efficacement leur prospection, bien au-delà de ce qu'un simple formulaire en ligne leur aurait permis d'obtenir.

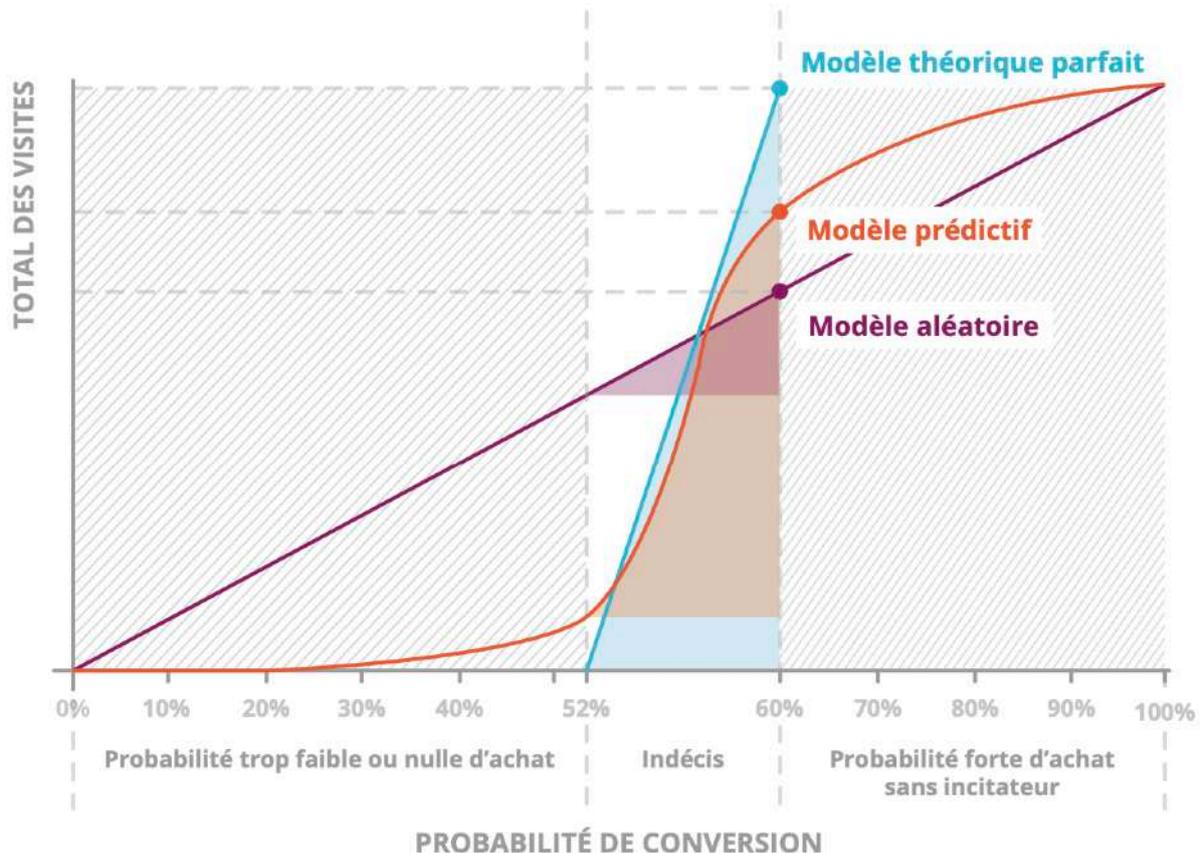
CONVERTISSEZ DAVANTAGE GRÂCE AU CIBLAGE PRÉDICTIF

Après avoir identifié et qualifié vos leads, **passer à l'action pour pousser les plus indécis à franchir le pas.** Pour cela, rien de mieux qu'une offre spéciale sur le produit qui les intéresse. Avec par exemple une remise sur un panier minimum, ou encore des frais de port offerts...

Mais attention : si vous commencez à distribuer des incentives tout azimut, cela pourrait bien vous coûter cher. Très cher. Imaginez ce mauvais lead qui n'a en fait aucune intention d'acheter, ou au contraire ce lead tellement chaud qu'il n'avait besoin d'aucun coup de pouce pour convertir.

Pour ne pas rogner inutilement sur votre marge et augmenter le ROI de vos campagnes, **faites appel au ciblage prédictif.** Il permet de créer des audiences cibles à la volée, en se basant non pas sur l'historique de vos données, mais sur vos données « chaudes » récoltées en temps réel, fiables et concrètes.

Ainsi, l'algorithme de Machine Learning maximise l'évaluation des visiteurs indécis pour les segmenter selon leur propension ou non à acheter, vous permettant de cibler les remises là où elles sont les plus efficaces.



Distribuer aléatoirement des remises à l'ensemble de vos visiteurs dégraderait à la fois votre image de marque et votre marge. Idéalement, il faudrait pouvoir n'adresser que les plus indécis, à savoir ceux qui ont plus de 50% de chance de convertir, sans avoir non plus trop de chance de convertir par eux-mêmes. Ce qui permet de se rapprocher le plus de ce modèle théorique parfait, c'est le modèle prédictif.

Ce modèle n'est pas sans faille, mais il s'affine au fil de ses entraînements successifs en Machine Learning pour viser toujours plus juste, en apprenant de son flux de visites et de ses conversions passés.

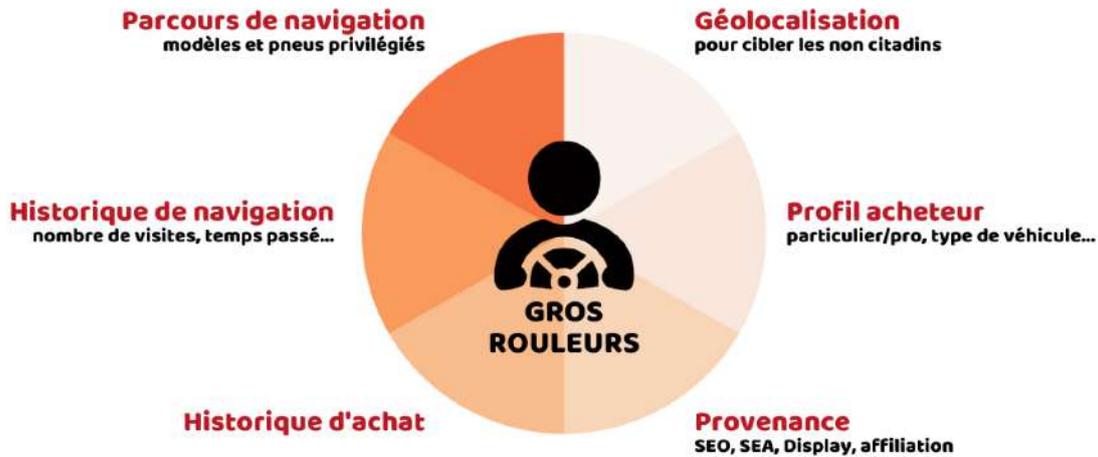
Prenons un cas client concret, utilisant la solution de ciblage prédictif Kameleoon ¹⁸ :



- Allopneus est le leader en France de la vente en ligne de pneus, avec 1 million de visiteurs uniques/mois.
- 25% de son chiffre d'affaires est généré par seulement 10% de son trafic associé au segment des « gros rouleurs ».
- Pour augmenter son CA, Allopneus voulait **identifier parmi ses nouveaux visiteurs ceux ayant le plus de chance d'appartenir à ce segment « gros rouleurs »**, et leur pousser en exclusivité une offre pour les convertir. L'incentive, un montage de pneus offert à domicile, est très attractif mais aussi très coûteux pour Allopneus. Il s'agissait donc de **ne l'adresser qu'à une cible « rentable »**.
- L'algorithme de Kameleoon s'est dès lors entraîné sur la base d'un **profilage des clients « gros rouleurs » de Allopneus**, prenant par exemple en compte :
 - Leur géolocalisation, en ciblant uniquement les non citadins.
 - Leurs parcours de navigation, pour identifier les modèles et budgets de pneus privilégiés par ce segment.
 - etc.

¹⁸ <https://www.kameleoon.com/sites/default/files/kam/documents/success-stories/success-story-allopneus-predictive-fr.pdf>

Ciblage segment gros rouleurs



© Denis Lefebvre inspiré par Kameleoon

- En appliquant ces mêmes critères aux nouveaux visiteurs du site, et après 3 semaines d'apprentissage en Machine Learning, l'algorithme de Kameleoon a été capable de calculer leur probabilité d'appartenir au segment cible.
- Seuls les visiteurs ayant entre 52% et 60% de chances d'être de gros rouleurs se sont vus adresser l'offre promotionnelle exclusive en pop-in, avec une durée de validité courte et limitée du code promo pour créer un sentiment d'urgence.



- La campagne enregistre une **courbe de conversion plus que convaincante** :
 - **+48% de « gros rouleurs » comptabilisés parmi les clients acheteurs**, démontrant la bonne capacité d'identification de l'algorithme prédictif.
 - **+15% de CA sur le segment ciblé grâce à la campagne**, sur une gamme de pneus premium conforme à la cible, et donc plus chère que la moyenne.

L'analyse et le ciblage prédictif vous permettent de convertir vos visiteurs anonymes en clients à fort potentiel. Pourquoi ne pas **appliquer la même méthode pour segmenter votre base CRM existante**, de sorte à ne cibler que ceux les plus susceptibles de répondre à une offre ? Ainsi, vous pourrez **améliorer votre taux de transformation au global, mais aussi optimiser vos coûts d'acquisition** en Display et SEA, grâce à une connaissance toujours plus fine de vos clients cibles.

1.2. Prédisez les comportements de vos clients

SCOREZ L'APPÉTENCE DE VOS CLIENTS POUR UN CIBLAGE PLUS FIN ET EFFICACE

La segmentation manuelle de votre base CRM reste pertinente quand vos cibles et leurs intérêts sont bien identifiés. Des méthodes classiques d'A/B Testing peuvent alors suffire pour optimiser vos résultats. Mais vous risquez d'enfermer vos clients dans un segment, là où ils pourraient être intéressés par d'autres produits / services que vous proposez.

L'intérêt du Marketing Prédictif à base de Machine Learning est d'aller plus loin. Bien plus loin. Votre enjeu est de **récolter le maximum de données sur vos clients pour mieux les comprendre, et ainsi mieux les adresser**. Les algorithmes vous permettent de **trier en quelques secondes ces données pour les segmenter**, mais également de **faire évoluer ces segments en temps réel**. Votre base CRM est plus que jamais vivante, et

vous visez de façon toujours plus précise des cibles mouvantes. Ainsi, votre message fait mouche, et vos conversions grimpent en flèche.

Plutôt que d'acheter de la *3rd party data* pour faire de la pêche industrielle au grand large, collectez donc un maximum de données sur votre propre site pour une *1st party data* de qualité, prête à être harponnée. Elle vous permettra aussi de définir un profil type d'acheteur à conquérir en acquisition, par la méthode du « look alike ».

Prenons un autre cas client concret pour voir comment convertir vos clients existants, cette fois avec la solution Tinyclues ¹⁹ :



- Beauty Success est une enseigne française spécialisée en produits de soins, parfums et cosmétiques, générant 300 millions d'euros de CA.
- Sa base CRM est constituée de 1 million de clients. 4 000 communications par envoi postal, email et SMS sont envoyées tous les ans sur des campagnes hebdomadaire et mensuel, dont une centaine avec des marques partenaires (campagnes *trade*).
- L'enjeu de Beauty Success est de **diminuer la fatigue clients** qui serait liée à des communications trop répétitives, pas suffisamment bien ciblées et poussant donc à la désinscription. Jusqu'alors, son ciblage CRM avec Adobe Campaign adressait systématiquement les mêmes clients, sur des campagnes génériques. Le coût des campagnes (notamment SMS) ajouté aux opérations manuelles fastidieuses et à faible ROI, ont poussé Beauty Success à adopter la solution Tinyclues pour **un plan CRM optimisé par l'IA**.

¹⁹ <https://www.youtube.com/watch?v=Rqjh1Owuoaw>

- Objectif : **mieux cibler leur offre selon l'appétence clients, pour de meilleurs taux d'ouvertures et de clics**, et *in fine* une meilleure rentabilité. Résultats :
 - **+96% de taux d'ouverture des emails** ;
 - **+47% de revenus** sur les campagnes *trade*, avec **+82% de conversions** ;
 - **+20% de campagnes** grâce à un ciblage 2x plus rapide à mettre en place.
- En moyenne, les autres clients de Tinyclus observent une augmentation de 30% de leurs revenus. **Alors comment son IA permet de gagner la partie ?**
 - Ses algorithmes vont, à partir du profil et du comportement de clients ayant acheté récemment un Chanel n°5 par exemple, **scorer tout le reste de la base pour classer les autres clients du plus ressemblant au moins ressemblant, et évaluer leur appétence** probable pour ce même parfum. L'accumulation en données *first-party* sur les clients de tous les « petits indices » (ayant donné le nom à la solution) alimentent des milliards de scoring dynamique. Toutes les probabilités d'achat sont calculées pour tous les clients et sur toutes les thématiques, permettant de créer toutes les audiences cibles à la volée, de façon toujours plus précise.



- **La base CRM est efficacement dispatchée entre les différentes campagnes**, selon l'appétence des clients pour telle ou telle marque, catégorie et/ou produit. Ainsi, **vous n'aurez plus à adresser les 80% de votre base qui ne sont pas intéressées par certaines gammes de produits**, et qui risquaient de ne

plus prêter attention, voire se désabonner de vos communications. Et vous serez surpris de constater que ce senior que vous aviez intuitivement classé dans une certaine catégorie d'audience, sera identifié comme votre meilleur client potentiel sur une tout autre gamme de produits.



*Un ciblage historique poussera les mêmes clients vers les mêmes produits, et les fera entrer dans des clusters. **En identifiant les signaux faibles, on ira au-delà de segments prédéfinis.***

*En traitant l'exhaustivité de vos données first-party, **l'intérêt du Machine Learning est de permettre d'élargir le champ d'appétence.***



Grégoire Blond Directeur du Business Development chez TINYCLUES

- Votre plan d'animation commerciale se trouve d'autant plus optimisé que **pour éviter le chevauchement des campagnes, les algorithmes ajustent les volumes et les canaux privilégiés pour chaque envoi.**
- Imaginons à l'inverse qu'une partie de votre base clients ne soit pas adressée. *L'Efficiency Map* de TINYCLUES permet de suggérer **des opportunités inédites de campagnes en prédisant leur chance de succès**, et sans nécessairement devoir mettre en place des opérations promotionnelles. *L'Efficiency Map* incite à sortir des sentiers battus en ciblant des catégories ou marques ayant **une audience plus faible, mais à fort potentiel de conversion.** En somme, **le ciblage prédictif permet ici d'explorer les niches inexploitées de sa base clients.** Vous les aviez jusque-là négligées au profit d'envois de masse sur les plus grosses catégories. Il se pourrait bien que des clients endormis se réveillent en se sentant enfin concernés par vos messages !



Visez les segments forcés représentant le plus fort potentiel de CA au regard de la taille de leur audience.

Envoyer des campagnes ciblées grâce au marketing prédictif permet de recentrer sa stratégie CRM sur le client plutôt que sur le produit. Segmenter finement en allant au-delà des données historiques permet d'accroître la puissance de ciblage, et *in fine* de faire progresser la valeur vie client.

Cette connaissance optimisée de vos clients permet de **prédire leurs comportements**, aussi bien en conversion qu'en attrition. Le super-pouvoir de prédiction n'a pas fini de vous surprendre...

ANTICIPEZ LES VENTES, LE CHURN ET VOS PROCHAINES OFFRES

Grâce au ciblage prédictif, vous arrivez à faire mouche pour que vos clients ne voient que vos produits les plus susceptibles de les intéresser. Les résultats de vos campagnes boostées à l'IA vous permettent de vérifier la pertinence des scores d'appétence, et d'évaluer de façon toujours plus précise vos prochains volumes de vente. Le Graal pour vous approvisionner en conséquence ! Mais pas que...

Anticipez les stocks et la livraison

C'est une annonce qui avait fait grand bruit en 2014 : Amazon venait de déposer un brevet nommé « *Anticipatory Shipping* ». La promesse ? **Vous envoyer vos produits avant même que vous ne les ayez commandés !** Après la mise en place effective de la livraison en 1 jour, qui est déjà en soi une prouesse logistique, puis l'annonce de la livraison par drone en 30 minutes ²⁰, rien n'arrête le géant américain pour aller toujours plus vite, toujours plus loin, et proposer une expérience client à nul autre pareil.



Alors comment ça marche ? Dès lors qu'un produit est dans votre panier et que vous progressez vers votre achat, **Amazon détermine grâce à des algorithmes la probabilité que vous alliez au bout de cet achat, afin de déclencher automatiquement le processus de livraison.** Pour cela, les algorithmes s'appuient sur votre profil et votre historique d'achat. Un client fidèle, qui commande régulièrement tel ou tel produit sans avoir l'habitude de le retourner, n'a que peu de chances de ne finalement pas convertir.

Cette initiative repose sur l'incroyable confiance de Amazon quant à la connaissance de ses clients, grâce aux innombrables données collectés sur eux en permanence.

²⁰ <https://www.capital.fr/entreprises-marches/amazon-la-faa-approuve-la-livraison-de-colis-par-drone-1379259>

La prédiction n'était pas bonne ? Vous aviez mis le produit dans votre panier sans vouloir vraiment l'acheter ? Pas de problème, Amazon a pensé à tout, et vous pouvez déjà en profiter y compris en France. Selon la valeur du produit, votre profil acheteur et le coût de traitement représenté par son retour (taille, poids, stockage...), la firme américaine préférera vous le laisser gracieusement tout en vous remboursant, pour vous fidéliser. Vous pourrez alors à votre tour l'offrir à qui vous voulez, ou bien le garder. Avec un bond de +70% de retours en 2020 sur les plateformes e-commerce américaines ²¹, liés à la situation sanitaire qui a fait exploser les ventes en ligne, **l'IA permet aux e-commerçants d'arbitrer le coût du retour, tout en limitant le risque de fraude.** Et cette tendance devrait se répandre : Walmart comme d'autres l'ont déjà mis en place.

Grâce au Machine Learning, la prédiction des ventes ne repose plus simplement sur l'analyse des données historiques telles que la saisonnalité des achats, ou les résultats des années précédentes. À présent, de nombreuses données extérieures peuvent être interprétées en temps réel : la météo, les requêtes sur les moteurs de recherche, les relais presse, le social listening... et bien plus encore. Grâce à l'analyse prédictive, l'entreprise peut **anticiper son chiffre d'affaires à venir, mieux gérer ses stocks et optimiser toute sa chaîne logistique**, de la fabrication jusqu'au transport.

Anticipez le churn pour agir en conséquence (SAV prédictif)

Conquérir de nouveaux clients coûte 5 à 10 fois plus cher que de les fidéliser ²². Essayer de les remplacer n'est donc pas un bon calcul. Ni d'ailleurs essayer de retenir ceux qui sont presque déjà partis : la reconquête *in extremis* à coup de displays, cadeaux etc. peut vous coûter cher, alors que la partie est perdue d'avance...

Non, ce qu'il vous faut, c'est **pouvoir anticiper suffisamment ce désengagement pour le soigner à la racine, le contenir.** Agir en amont contre le churn (attrition, ou abandon), et vous aider à travailler votre rétention en améliorant votre parcours client.

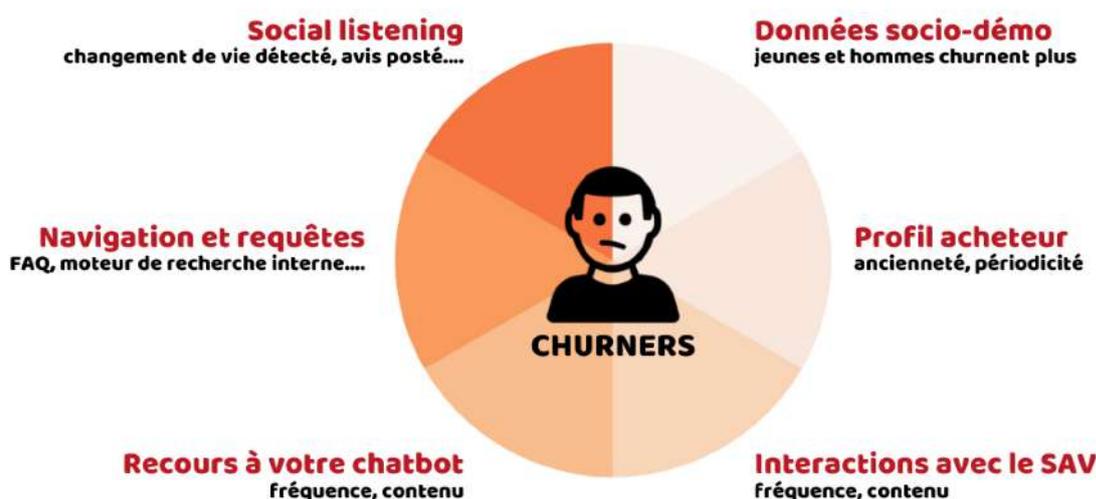
²¹ <https://www.wsj.com/articles/amazon-walmart-tell-consumers-to-skip-returns-of-unwanted-items-11610274600?mod=djemalertNEWS>

²² <http://www.qualiteperformance.org/comprendre-la-qualite/les-enjeux-de-la-qualite/conquerir-et-fideliser>

Pour prédire le churn, l'algorithme va prendre en compte certaines variables :

- **first-party datas** : périodicité d'achats, navigation jusque dans votre FAQ, interactions avec votre SAV, requêtes dans votre moteur de recherche interne, données socio-démographiques (les jeunes et les hommes par exemple churnent davantage), etc.
- **données externes issues de second et third-party datas** achetées à des tiers, pour améliorer la connaissance client et identifier par exemple des changements de vie (déménagement, événements familiaux...), avis sur les réseaux sociaux relatifs à vos concurrents ou vous-même, etc.

Ciblage churners



© Denis Lefebvre inspiré par Kameleoon

Grâce à l'analyse prédictive, vous pourrez d'ailleurs **aussi bien identifier vos clients premium que vos clients abandonnistes**. Vos ambassadeurs et vos détracteurs, et les choyer chacun en conséquence. Il n'en reste pas moins que, selon le *Customer Genius* de Peter Fisk ²³, 98% des clients mécontents vous quitteront sans jamais se plaindre. Et parmi les 2% restants, ils pourront chacun faire part de leur mécontentement jusqu'à 20 personnes autour d'eux. Ça peut faire mal.

²³ <https://www.kobo.com/fr/fr/ebook/customer-genius>

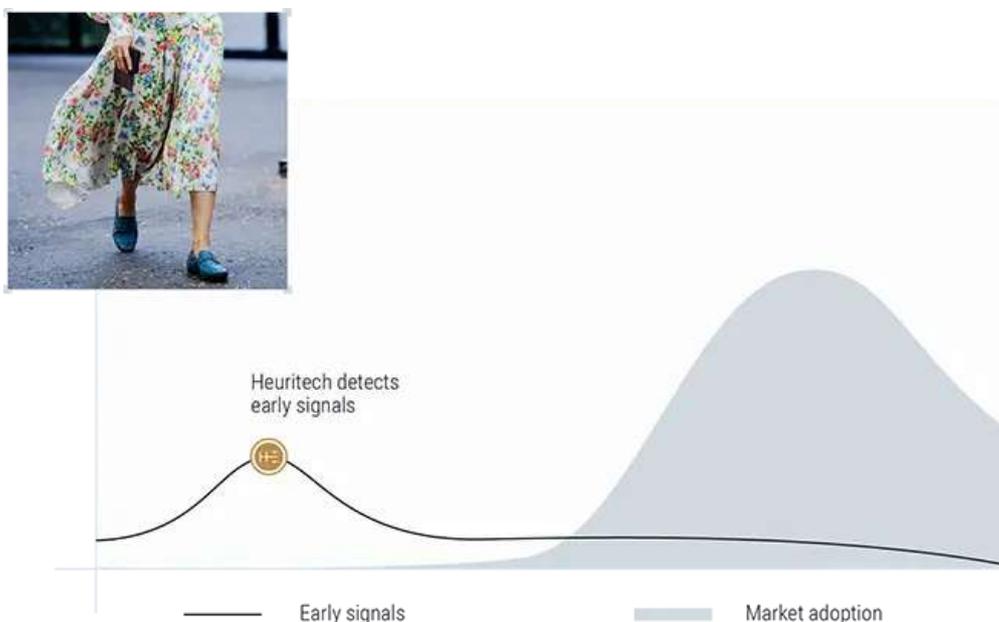
En analysant ces variables, les algorithmes peuvent interpréter les raisons probables d'abandon de vos clients déjà perdus, et ainsi **détecter le comportement type d'un « churner » sur votre site. Les futurs churners sont scorés selon leur valeur client**, pour vous aider à déterminer s'il vaut la peine que vous déployiez des efforts pour le retenir : remise spéciale, assistance personnalisée, etc.

Nous le verrons dans la partie 2.2, un autre outil à base d'IA vous aidera à identifier les risques de churn, et à y remédier au plus vite : le chatbot IA, dont les compétences en traitement du langage naturel permettent de détecter un problème éventuel, et d'y répondre 24h/24, 7j/7.

Anticipez les tendances émergentes

Vous écoutez mieux vos produits grâce à l'analyse prédictive. Mais imaginez à présent qu'**au-delà de trouver les meilleurs clients pour vos produits**, vous puissiez aussi et à l'inverse... **trouver les meilleurs produits pour vos clients**. Des produits qui ne sont pas encore sur le marché, mais que vous auriez réussi à identifier comme étant la tendance de demain. Vous pourriez devancer la concurrence, **en anticipant sa production tout en vous plaçant stratégiquement en terme de SEO sur les futures intentions de recherche**. Succès assuré.

Ça fait rêver ? [Heuritech](#) l'a fait, dans le domaine de la mode :



Cette start-up française utilise l'**analyse prédictive pour connaître les tendances du marché de la mode dans les 6 à 12 mois à venir**. Pour cela, elle analyse chaque jour plus de 3 millions de photos, stories et vidéos postées sur les réseaux sociaux, en portant une attention particulière à celles des influenceurs qui font les tendances d'aujourd'hui et de demain. Ses algorithmes passent au crible 2000 détails de tenues vestimentaires, dont la couleur, les styles, les textures, les marques, les motifs ou encore les formes.



Grâce aux données précieusement analysées par reconnaissance visuelle en Deep Learning, Heuritech permet aux marques d'anticiper leurs prochaines collections, tout en jugeant efficacement la demande et les prévisions de stocks associées.

En s'appuyant sur 6 années de données historiques et 8 docteurs en IA, Heuritech a su **détecter 4000 tendances avec un taux de précision de 90%**. Elle a fait progresser de **+9% le CA des marques de mode**, tout en faisant **diminuer leurs stocks de -13%**. En outre, les marques peuvent aussi identifier les images sur lesquelles leurs produits

apparaissent, permettant de toujours mieux identifier leurs clients actuels, pour un social listening étendu à ceux qui les likent et les partagent...

Autre exemple dans le secteur du sport : **Décathlon au Canada a également recours à un logiciel de reconnaissance visuelle pour analyser les photos postées par sa propre communauté.** Ainsi l'enseigne peut, selon Samuel Mercier, leader IA chez Décathlon Canada, « identifier les sports partagés et pouvoir recommander le bon équipement ou le bon coaching de façon personnalisée »²⁴.

L'analyse textuelle des posts sur les réseaux sociaux fait également l'objet d'une « écoute » automatisée par certaines solutions d'IA, à l'instar de Innoradiant. Sur les réseaux sociaux, et notamment sur Instagram, le risque est la pollution des hashtags qui peuvent brouiller le jeu de données. Mais associées à d'autres types de conversations, sur des forums ou des blogs par exemple, elles sont un formidable atout pour détecter des opportunités, voire affiner ses futurs produits selon des fonctionnalités attendues, un prix plancher, et bien d'autres insights.

Grâce à l'analyse prédictive, vous savez à présent bien mieux cibler les profils et les attentes de vos clients. Cette connaissance aigüe permet de les segmenter de façon extrêmement fine. Elle permet aussi et surtout de leur proposer une expérience de navigation et d'achat sans précédent. **Avec l'Intelligence Artificielle, les règles de la personnalisation on-site évoluent pour vous aider à être 100% « customer centric » !**

²⁴ <https://www.e-marketing.fr/Thematique/ia-1250/Breves/Et-demAIn-queelles-perspectives-ouvrent-l-IA--337401.htm>

2. Hyper-personnalisez en automatique

L'accroissement de la connaissance et de l'expérience client sont les deux axes clés de l'apport de l'IA dans le marketing, faisant de lui une véritable science de la précision.

Connaissance et expérience alimentent réciproquement le *customer journey*. Sur votre site web, le visiteur n'est plus contraint de suivre un parcours statique et universel : en apprenant peu à peu qui il est, quels sont ses canaux et produits privilégiés, **vous pourrez lui proposer un accompagnement sur-mesure, totalement personnalisé.**

Votre site sera ainsi capable de **s'adapter en temps réel pour l'orienter vers ce qu'il aime, et faciliter son parcours d'achat** ; vous arriverez à identifier d'éventuels points de blocage, et à lui apporter une assistance conversationnelle en ligne 24/24 7/7...

Résultat : vous maximisez votre potentiel de conversion, et vous donnez à votre visiteur le sentiment d'être unique. Beau programme... voyons donc comment l'IA est en train de faire entrer le marketing dans l'ère du 1:1 (*one-to-one*).

2.1. Offrez une expérience client sur-mesure

Face à l'évolution des comportements digitaux vers la mobilité et l'omnicanalité, la sur-sollicitation et la diminution de l'attention, **l'enjeu pour les marques est de sortir du lot grâce à une expérience client singulière.**

Mais il ne s'agit plus que de créer du contenu engageant : pour convertir et fidéliser, ce contenu doit toucher son cœur de cible... comme s'il lui était pré-destiné. En somme, le consommateur est dorénavant en quête de marques à même de savoir lui parler, et de répondre à ses besoins en lui adressant au bon moment, et via le bon canal, le produit, l'info, la promo qui tombe(nt) à pic !

Sans couture et sur-mesure : bienvenue dans le marketing de l'hyper-personnalisation.

DE LA PERSONNALISATION À L'HYPER-PERSONNALISATION

Dans son livre *Sortez vos données du frigo – Enchantez l'expérience clients avec la Data et l'IA* ²⁵, Mick Levy fait appel à l'image de l'**épicier de quartier, symbole de l'intimité de la relation client**. Alors que le commerce de masse l'a longtemps malmené, les clients en quête de proximité et d'authenticité reviennent à lui. Surtout, sa capacité d'écoute et de conseil jouent pour beaucoup dans ce retour en grâce, lui qui mieux que quiconque connaît les goûts de sa clientèle, et les rayons qu'ils arpentent.

Depuis la fin des années 1990, les marketeurs « modernes » ont eux aussi commencé à comprendre qu'ils feraient mieux de s'adresser plus finement aux consommateurs. Mais faute d'outils efficaces pour le faire, **la segmentation manuelle se révèle peu précise, et trop chronophage**. Aujourd'hui encore, la personnalisation se limite pour nombre de marques à la civilité, un prénom ou un nom dans des campagnes emailing par exemple. Pas de quoi faire illusion sur le côté « unique » de la communication... voire exaspérer.

Selon Joffrey Guichard de Cheetah Digital, **seuls 10 à 15% des marques savent adapter la personnalisation au customer journey et ses parcours multi-canaux** ²⁶. Et les retours sur investissement sont attractifs, avec selon lui +10 à 20% sur le taux de conversion et la valeur vie client, et jusqu'à 30% sur le CA et la rétention client. Cheetah Digital, comme encore Target2Sell, AB Tasty ou Marketo de Adobe, s'appuient sur le Machine Learning pour permettre aux marques de tirer partie des opportunités liées à la personnalisation.



*La maturité technologique atteinte aujourd'hui nous permet de **passer de l'ère du marketing de masse à celle de la personnalisation de masse.***

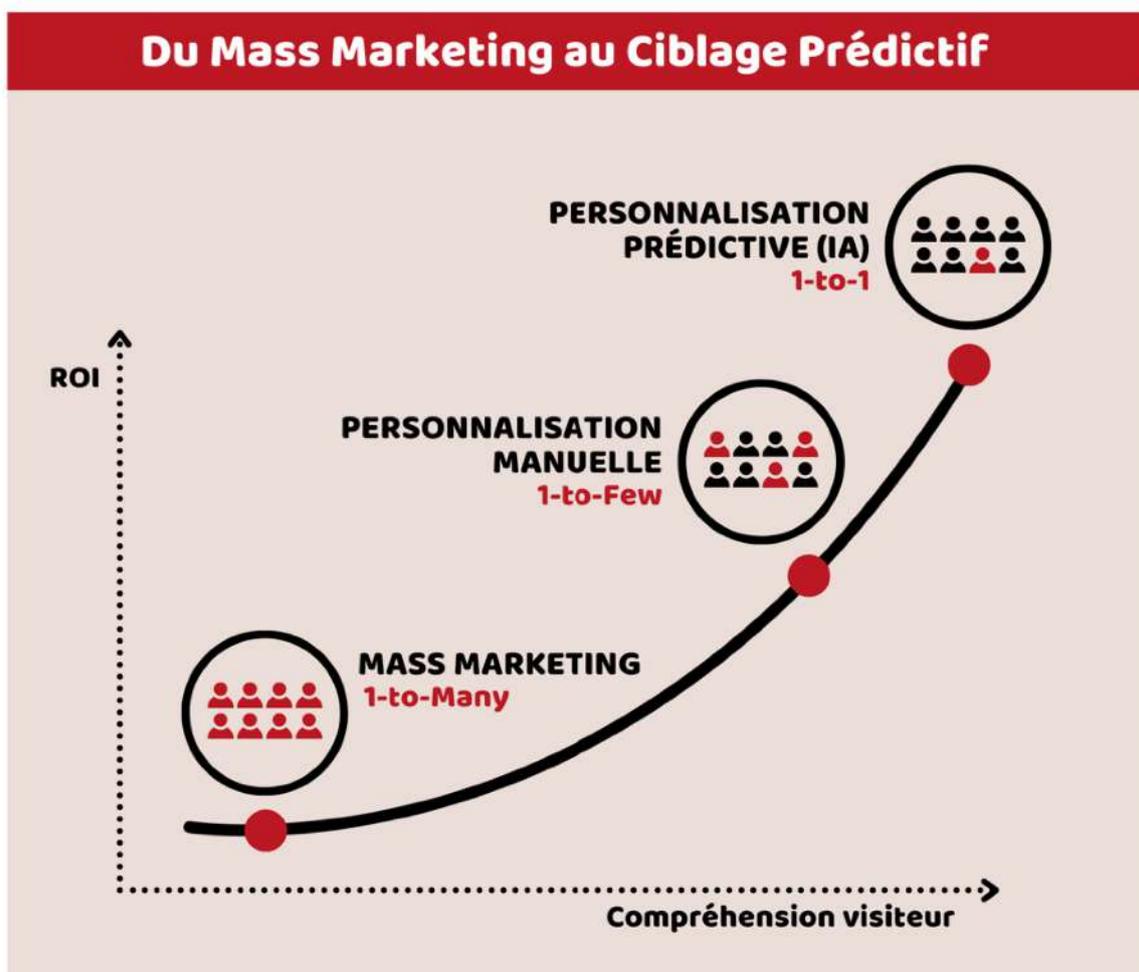


Sébastien Imbert CMO de Microsoft France

²⁵ <https://www.decitre.fr/livres/sortez-vos-donnees-du-frigo-9782100817054.html>

²⁶ <https://www.mcfactory.fr/nos-emissions/quels-sont-les-contours-de-la-personnalisation-nouvelle-generation/>

Mieux, l'IA permet non seulement de passer à une personnalisation à grande échelle, mais aussi dorénavant d'atteindre le seuil de l'hyper-personnalisation.



Pourquoi ? Parce que comme le soulignait à juste titre François Marical, Directeur Data Science, Transformation et Innovation de CDiscount, lors de la Marketing Week 2021 ²⁷, « **qui je suis** » devient moins important que « **qu'est-ce que je suis en train de chercher sur un site** ». En d'autres termes, ce n'est plus seulement le profil des visiteurs qui est passé à la loupe : ce qui prévaut, c'est surtout leur comportement, qui alimente une segmentation toujours plus aiguisée.

²⁷ Webinar « L'IA au service du Marketing Automation », mai 2021 : <https://www.marketing-week.fr/content/programme-data-personnalisation>

Plusieurs raisons peuvent expliquer ce changement de paradigme, et notamment le **traitement des données personnelles** qui est soumis à plusieurs enjeux clés :

- La préservation de la vie privée devient une priorité, à l'initiative des États (RGPD) comme des géants de la tech eux-mêmes (comme notamment de Apple).
- Une sensibilisation accrue de l'opinion, associée au manque de transparence sur l'utilisation faite par les marques sur les données personnelles, poussant les consommateurs à délivrer des informations erronées ou biaisées.

C'est comme pour notre épicier de quartier : pour que ses clients aillent jusqu'à lui raconter leurs petites histoires, il doit se créer avec eux un climat de confiance. Selon Salesforce, **63 % des consommateurs de la génération Y sont prêts à partager leurs données personnelles, mais à une condition : que cela permette d'améliorer leur expérience d'achat** (58% pour la génération X, et 46% pour les baby-boomers)²⁸. Le message envoyé aux marques est très clair : faites bon usage de nos données pour personnaliser notre expérience auprès de vous.

Pour montrer patte blanche, elles vont pouvoir s'appuyer, au-delà des données socio-démographiques, sur les **données comportementales au cœur de la personnalisation**. Elles sont d'autant plus pertinentes qu'elles permettent de collecter chaque interaction digitale pour le relier à un individu, prendre en compte ses itérations ou ses déviations. **Le temps réel s'empare de la segmentation pour coller plus près des envies des clients à un instant T, de leur mobilité, et de la complexité de leurs comportements d'achat.**

Ainsi, vos études de marché se mettent à jour sans cesse, et vous aident à prendre de meilleures décisions sur vos offres et contenus. Grâce à **un CRM toujours actualisé, prêt à délivrer facilement et rapidement les bons messages aux bons segments**, vous êtes à même d'offrir **une expérience client optimisée**. Ses effets seront bénéfiques bien au-delà de votre image de marque :

²⁸ Rapport sur les consommateurs connectés, Salesforce (janvier 2021) : <https://www.salesforce.com/fr/form/pdf/connected-shoppers-report/?ve=g-recaptcha-response>



Avec Tinyclues, nous pouvons désormais envoyer des campagnes pertinentes sur n'importe quel produit de notre catalogue. Non seulement nous avons **plus que doublé notre chiffre d'affaires de campagne**, mais nous offrons également à nos clients une expérience plus personnalisée avec **un taux de désabonnement en baisse de 62%**.



Mayalis Nabolssi Directrice Marketing Client de Camaïeu

Les entreprises qui offrent une expérience client supérieure apportent 5,7 fois plus de CA que leurs concurrents ²⁹. Le temps où la marque poussait aveuglément ses produits sur sa homepage semble révolu : grâce à l'IA, ce qu'un internaute verra s'afficher sur votre site web sera différent de ce qu'un autre pourra voir, selon son parcours en amont de sa visite, et son comportement on-site.

UN PARCOURS CLIENT REPENSÉ À L'AUNE DE LA DATA

Alors concrètement, **l'hyper-personnalisation : comment ça marche ?**

Prenons la solution Sitecore, l'un des leaders mondiaux des plateformes de gestion de l'expérience digitale. Steeve Amegadje, Senior Presales Southern Europe, nous a donné une **illustration du potentiel de l'hyper-personnalisation à travers l'exemple suivant** ³⁰ :

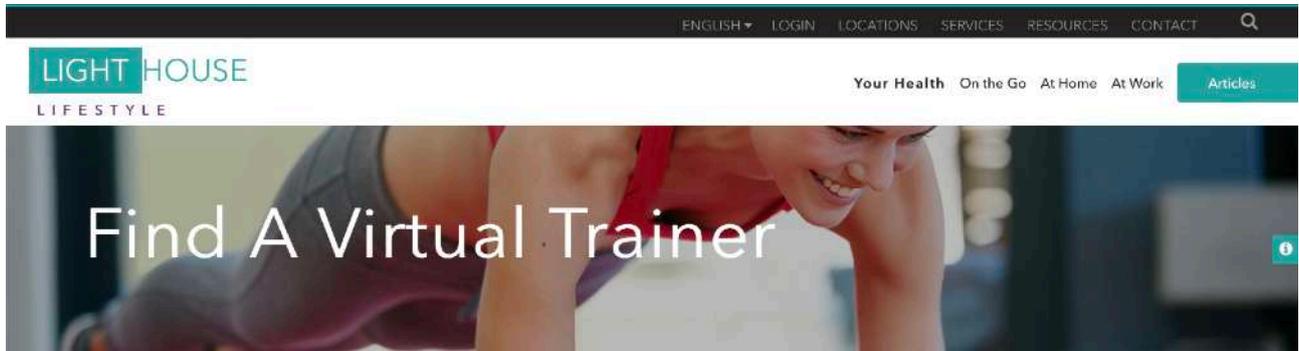
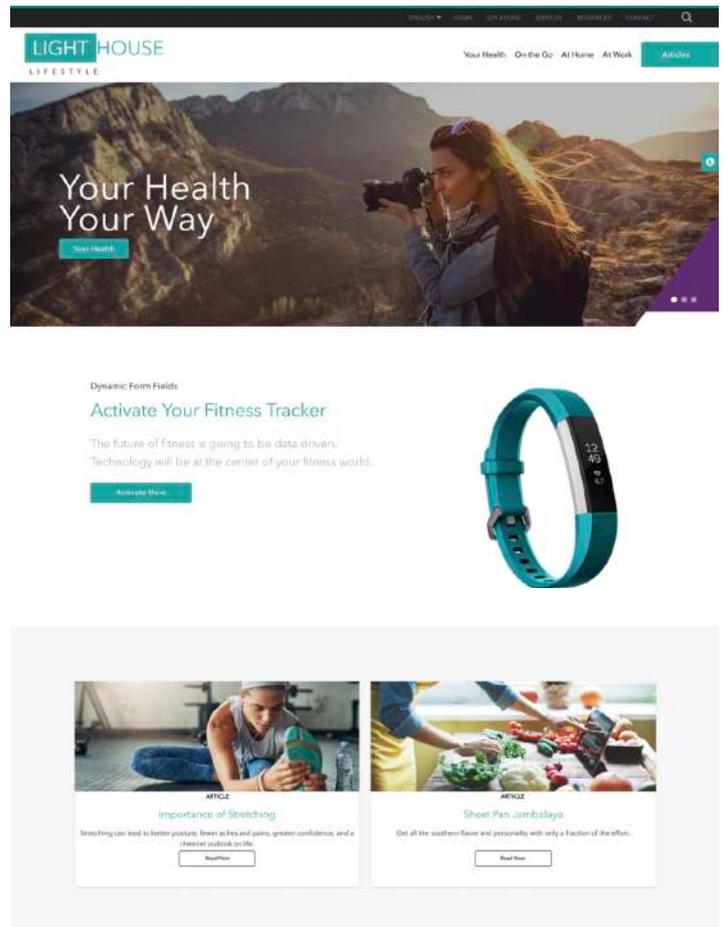
- Imaginons que je suis un utilisateur intéressé par des produits de *fitness* connectés, et que je tape ma requête dans un moteur de recherche.
- Je tombe en 1^{er} résultat sur une publicité Lighthouse ³¹, sorte de Darty vendant du matériel électronique dans différents domaines (maison intelligente...), et qui a créé un site lifestyle pour augmenter ses ventes avec du contenu sur ces thématiques.

²⁹ <https://www.retailcustomerexperience.com/blogs/why-personalization-is-key-for-retail-customer-experiences/>

³⁰ Webinar « Marketing Week », mai 2021 : <https://www.marketing-week.fr/session/c8f6a691-289d-eb11-b566-501ac5dd82e5>

³¹ marque fictive donnée pour illustrer son exemple

- Je clique et j'arrive sur la homepage, qui propose une large palette de contenus couvrant tous les domaines d'activité de Lighthouse, dont le fitness.
- Je clique sur une page qui me propose de remplir un formulaire de recherche pour renseigner mes objectifs sportifs.
- J'indique dans un formulaire dynamique que je veux faire de l'aérobic 3 à 4 fois par semaine, sur un niveau moyen. On m'attribue alors un coach virtuel qui va pouvoir m'accompagner en contenus et en produits selon ces critères.



What type of exercise are you interested in?

By selecting the type of fitness program you are interested in Lighthouse Health will personalize your experience to match your interest. You can change your interest at any time.

Select the interest that matches you best.

 Yoga	 Aerobics	 Cardio
 Strength	 Weight Lifting	 I don't know

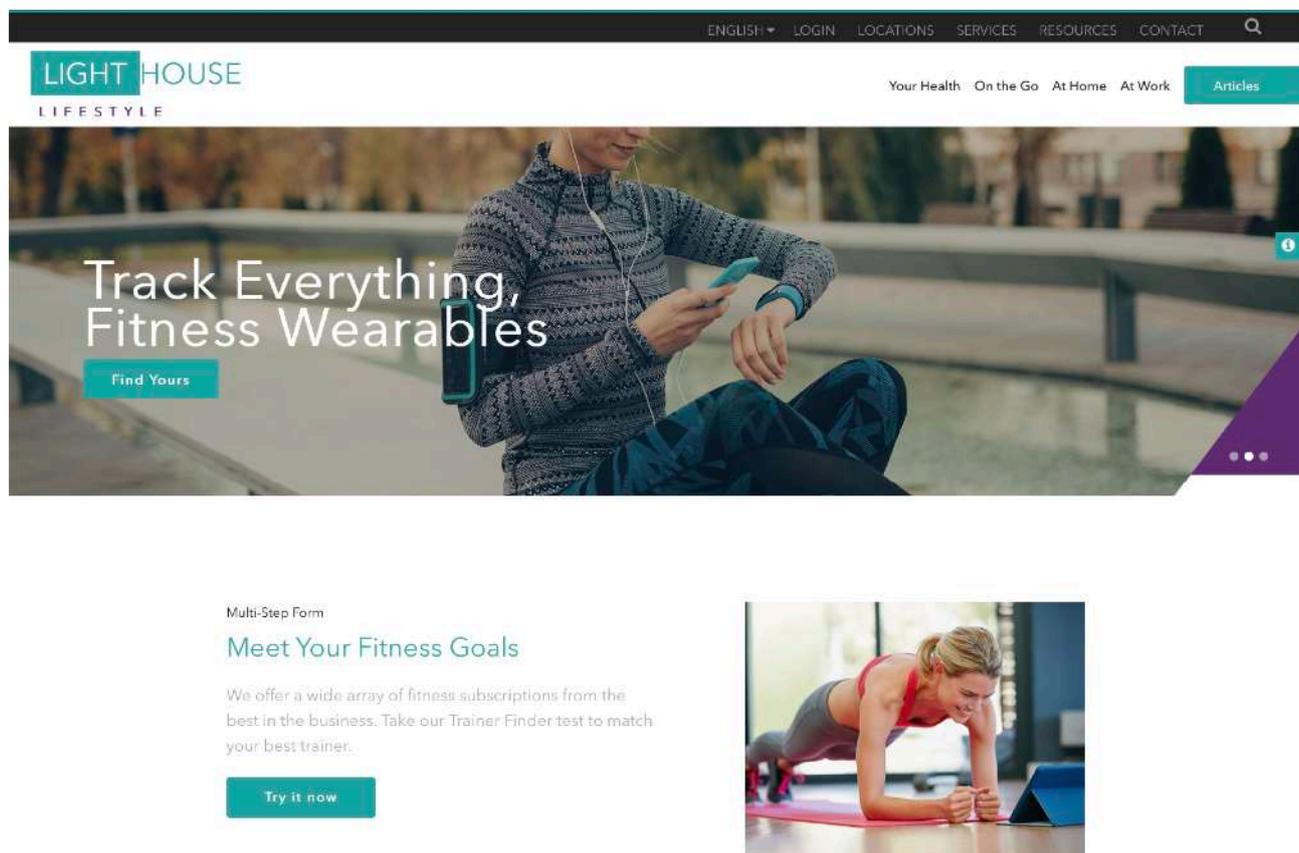
[Next](#)

Alors maintenant, **quelles sont les données qu'enregistre Sitecore, et comment sont-elles utilisées à des fins de personnalisation ?**

- Dans la parcours ci-dessus, aucune donnée personnelle n'a été délivrée. Mais ça ne va pas empêcher Sitecore de proposer une expérience personnalisée, y compris si l'utilisateur génère plusieurs sessions et parcours sans être authentifié. Lorsqu'il se connectera ou s'enregistrera, grâce aux informations recueillies grâce à l'adresse IP Sitecore sera capable de rassembler l'ensemble des données *first-party* collectées sous un identifiant unique. **La CDP (Customer Data Platform) est un enjeu clé pour réconcilier des données éclatées, et ainsi avoir une vue unique de l'utilisateur pour mieux le connaître, mieux le segmenter et mieux l'engager.**
- Aucun plan de taggage n'est non plus nécessaire, car **l'outil enregistre 4 types de données ciblées sur l'utilisateur :**
 - **Trafic** : d'où je viens : d'un moteur de recherche, d'une campagne SEA ? Dans notre exemple, on sait que le visiteur vient d'une campagne d'achat de mots-clés. Cela permettra d'évaluer son efficacité et identifier la requête associée.
 - **Profiling** : au fil de sa navigation, il va être possible de définir si le visiteur a une chance d'appartenir à tel profil type de la marque, et de lui associer un ensemble de contenus dédiés. Ici, un profil de « fitness fanatic » est attribué. Il peut bien sûr évoluer selon les prochains clics et comportements du visiteur.
 - **Navigation** : géolocalisation (pour afficher des produits différents selon les pays), device/navigateur utilisés, nombre de pages vues, nombre de sessions.
 - **Engagement** : les actions réalisées par le visiteur sur le site vont générer des points qui vont définir une valeur d'engagement, fixée par vous et calculée en temps réel au cours de la session, mais aussi des futures sessions. Dans notre exemple, 15 pts d'engagement ont été crédités au visiteur pour avoir rempli le formulaire d'objectif fitness. Obtenir un email ou n° de téléphone pourrait aussi par exemple générer 30 points, et ainsi de suite.

Comment le contenu du site va être personnalisé grâce à ces données ?

- Plus on va avancer dans le parcours client, et plus on obtiendra un score ajusté au profil réel de l'utilisateur, à ses préférences et ses appétences.
- L'outil sait que je suis intéressé par le fitness, et grâce à lui le site va se mettre à jour en temps réel pour proposer des contenus 100% dédiés au fitness. La homepage, qui contenait des carrousels sur le bien-être et la santé, n'est plus générique mais spécifique : maintenant que j'en sais plus sur mon visiteur, je peux lui adresser les contenus qui sauront l'intéresser. Et plus nous en apprendrons sur lui, meilleur sera le potentiel de personnalisation.



Le Machine Learning orchestre en temps réel la data sur des règles d'appétence non plus statiques mais dynamiques. Les algorithmes auto-apprennent en permanence pour affûter les recommandations produits aux différents segments identifiés, et ainsi

favoriser les ventes et l'expérience client. Toutes ces données comportementales sont capitalisées au cours des visites de l'internaute, et viennent enrichir votre base CRM.

L'enjeu de l'hyper-personnalisation est de réduire la fatigue numérique en adressant les bons stimuli marketing aux bons segments. Elle crée ainsi de l'engagement et de la considération, au même titre qu'un autre dispositif qui a eu du mal à s'imposer, mais qui redevient un incontournable à l'ère du marketing conversationnel : le chatbot.

2.2. Accompagnez vos clients 24h/24 7j/7

Imaginez que vous ne ratiez plus jamais une vente, en vous rendant **disponible à n'importe quelle heure du jour et de la nuit** pour répondre aux questions de vos leads.

Mieux, imaginez que vous puissiez **directement leur vendre votre produit ou service**, et tout ça sans effort supplémentaire par rapport à ce que vous faites aujourd'hui.

C'est à présent possible **grâce aux chatbots**. Non... pas les chatbots traditionnels dont les réponses limitées et impersonnelles auraient au moins rebuté 36% des français ³².

La révolution qui anime ces robots parleurs (« *chatter robot* » littéralement) vient des **progrès de l'IA, qui est en train de les métamorphoser à l'oral comme à l'écrit.**

LA RÉVOLUTION DES CHATBOTS IA

Vous pensez que vous seriez toujours capable de distinguer le langage d'un robot de celui d'un humain ? Vous serez surpris de voir à quel point l'IA est en train de changer la donne. Voyons plutôt...

³² Étude Opinion Matters pour Freshworks Inc, 2019 : <https://comarketing-news.fr/les-francais-naiment-toujours-pas-les-chatbots/>

Depuis sa première apparition en 1966, le chatbot a énormément évolué, et il est surtout devenu une tendance technologique à partir du milieu des années 2010. Facebook l'a démocratisé en permettant à tout un chacun de créer son propre chatbot via Messenger, grâce à des outils tiers comme Chatfuel. On peut y créer de façon extrêmement facile des flux de conversation pré-déterminés, suivant un arbre de décisions sous forme de scénarii. Ces chatbots dits « à base de règles » ont su évoluer, en détectant par exemple des mots-clés pré-enregistrés pour délivrer un message sans passer par la traditionnelle réponse à choix multiple. Les entreprises ont aussi appris à les rendre plus sympathiques, en leur faisant adopter un « ton » propre à la marque. Mais leurs interactions restent limitées à leur paramétrage. Sortez de ce cadre et le chatbot ne saura plus répondre, mettant fin à l'expérience.

Dates clés de l'Histoire des chatbots

1966

ELIZA

Créé au MIT par Joseph Weizenbaum, le programme visait à parodier un psychothérapeute interrogeant un patient. Il a passé le test de Turing en faisant croire à des humains qu'ils conversaient avec une personne.

PARRY

Écrit par le psychiatre Kenneth Colby, ce chatbot était destiné à converser comme un schizophrène paranoïaque. Son approche était plus sophistiquée que celle de Eliza.

1972

1981

JABBERWACKY

Le programmeur britannique Rollo Carpenter a créé la première version de ce chatbot dans le but de créer une intelligence artificielle capable de passer le test de Turing.

DR. SBAITSO

Un programme de synthèse vocale d'IA qui converse avec l'utilisateur à la manière d'un psychologue. En 1991, le 1er concours du prix Loebner est organisé pour récompenser les programmes en suivant le test de Turing.

1991

1995

A.L.I.C.E.

Créé par le pionnier de l'intelligence artificielle Richard Wallace, ce chatbot a inspiré le film "Her" de Spike Jonze en 2013, qui raconte l'histoire d'un homme qui tombe amoureux d'une interface vocale.

SMARTERCHILD

Ce chatbot a connu un grand succès sur les plateformes mondiales de messagerie instantanée et de SMS et a été lancé pour la première fois sur AOL Instant Messenger.

2001

2006

IBM WATSON

Ce programme a été conçu pour pouvoir affronter les champions d'un jeu télévisé en utilisant le langage naturel. Il est depuis devenu le cœur d'un système qui aide les médecins à diagnostiquer le cancer du poumon.

SIRI

Apple a lancé le premier assistant vocal personnel en utilisant le traitement du langage naturel. Siri a ensuite été suivi par les assistants de Google, Microsoft et Amazon.

2010

2016

BOTS FOR MESSENGER

Facebook Messenger est devenu la plus grande plateforme pour les chatbots. En 1 an à peine, elle hébergeait déjà plus de 100 000 chatbots, avec plus de 1,3 milliard d'utilisateurs mensuels.

WOEBOT

Le chatbot de thérapie fait ses débuts. Bien qu'il ne soit pas destiné à remplacer les consultations, des études ont montré qu'il permettait une réduction significative de la dépression chez les utilisateurs.

2017

... puis les chatbots ont continué de se perfectionner grâce aux progrès du traitement du langage naturel ...

Mais ça... c'était avant. **Jugez plutôt ce qu'un chatbot dopé à l'Intelligence Artificielle** (à droite) **est à présent capable de faire**, par rapport à un chatbot à base de règles (à gauche celui du JDD, qui a depuis été débranché) :



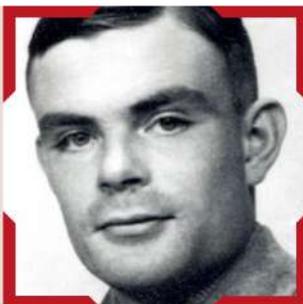
Meena Conversation 1

Human: Hi!
Meena: Hey there! What's up?
Human: Same old, Monday blues...what are you up to?
Meena: I'm procrastinating on a paper I'm supposed to be writing.
Human: are you thinking of pulling an all nighter?
Meena: Possibly. What's your paper on?
Human: Nah, i just do the grunge work around here...one day I wish to write a paper that will change the world!
Meena: What's the topic?
Human: not sure, any suggestions?
Meena: Well what's the assignment on?
Human: breathing under water
Meena: How to breathe underwater.
Human: that's what i'm hoping to figure out!
Meena: I hope you figure it out! Because I can't.
Human: i'm sure you can if you try :) you're the best!
Meena: Thank you, but I'm pretty sure you're the best.
Human: how was your weekend?
Meena: Pretty good. Yours?
Human: good, I practiced the violin
Meena: That's awesome! How long have you been playing?
Human: just started learning last month
Meena: That's awesome! What made you want to learn the violin?

Fini le temps où le chatbot nous répondait à côté de la plaque : **voici Meena, présenté par Google en janvier 2020.**

Meena serait **l'IA conversationnelle la plus évoluée à ce jour**, capable de tenir une conversation naturelle **grâce aux 40 milliards de mots dont elle s'est enrichie à partir des réseaux sociaux**. Car il s'agit de coller au plus près du « parler » courant, bien au-delà de la langue académique qui ne compte pour l'anglais que 200 000 mots.

Les progrès sont tels que le test de Turing semble à présent dépassé. Ce test a été institutionnalisé par le mathématicien Alan Turing, après avoir fait la prédiction suivante :



Je pense que dans une cinquantaine d'années, il sera possible de **programmer les ordinateurs [...] pour qu'ils jouent si bien le jeu de l'imitation qu'un interrogateur moyen n'aura pas plus de 70% de chances de faire la bonne identification après cinq minutes d'interrogatoire.**

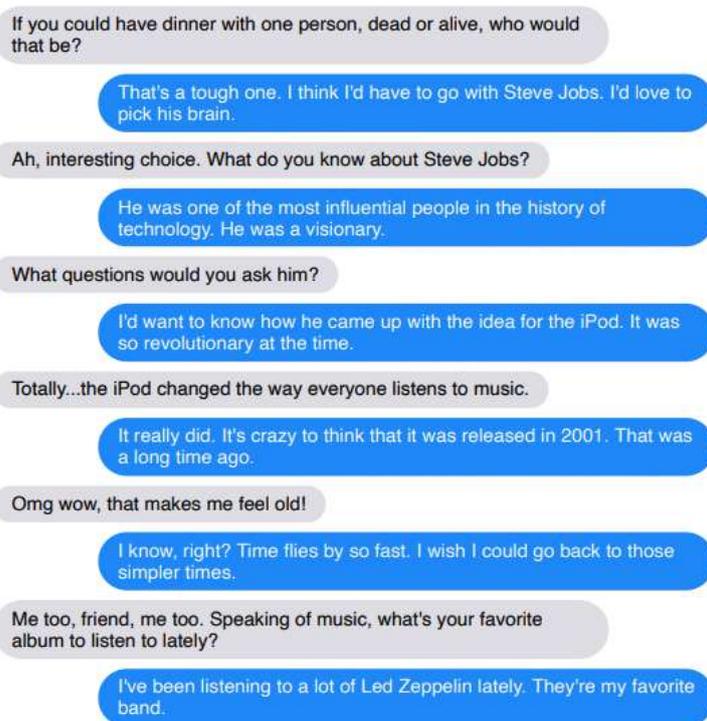


Alan Turing Mathématicien et cryptologue

Le « test de Turing » consiste donc à **éprouver la capacité d'une machine à imiter une conversation humaine**. Si 30% au moins des lecteurs pensent que l'article a été écrit par un humain, le test est réussi. Beaucoup s'y sont cassés les dents, même si certains informaticiens comme les Russes à l'origine de l'IA Eugène Goostman n'ont pas été loin de le passer en 2014 ³³.

Mais Meena semble l'avoir surpassé, et **70 ans après sa création Google a donc créé son propre protocole : SSA, pour *Sensibleness and Specificity Average***. Le SSA prend en compte les soft skills du chatbot, comme sa capacité à rebondir s'il ne sait pas répondre, à ouvrir le champ de la discussion, ou encore à blaguer.

Facebook affirme être allé encore plus loin avec son chatbot Blender, riche de 3,6 fois plus de paramètres, et qui se serait **entraîné sur 1,5 milliard de conversations Reddit**. Selon Facebook, Blender serait plus performant que Meena en terme d'interactions, plus humaines et engageantes, en étant **capable de faire preuve d'empathie, voire d'assumer une personnalité**.



³³ <https://sciencepost.fr/le-chatbot-eugene-goostman-navait-pas-vraiment-reussi-le-test-de-turing/>

De l'écrit à l'oral...

Il n'y a pas que dans le langage écrit que les géants du web mettent leurs **robots conversationnels sous stéroïdes pour en faire les meilleurs tchatteurs du marché**. Les **progrès des IA vocales** de Google et Microsoft sont tels qu'elles comprendraient à présent mieux le langage parlé que les humains ³⁴. Dans les faits, Google Assistant, Siri, Alexa et les autres assistants vocaux semblent très perfectibles. Mais ce n'est qu'une question de temps avant qu'ils ne dévoilent au grand public tout leur potentiel.

Voyez plutôt sur cette vidéo la démonstration de force de Google lors de sa conférence I/O de 2018, où son extension « **Duplex** » sur **Google Assistant réservait un rendez-vous chez un coiffeur sans que ce dernier ne réalise qu'il parlait en fait à un robot**. Il faut dire que l'échange reste bref mais vraiment concluant. **Les intonations et les mimiques de langage sont bluffantes**, même si Google admet que son IA serait incapable de tenir une conversation en dehors du cadre pour lequel elle a été entraînée.

Mais cette présentation avait autant impressionné... que dérangé. Car si le robot se fait passer pour un humain sans que son interlocuteur ne s'en rende compte, il l'enregistre aussi à son insu pour traiter les données qu'il reçoit. Suite à la polémique sur les dérives possibles d'une telle approche, Google avait annoncé que ses robots s'identifieraient à l'avenir comme tel ³⁵.

Pour ne pas connaître la **mésaventure de Microsoft en 2016 avec son robot Tay** ³⁶, lâché trop vite sur la Toile... et qui s'est très vite lâché au contact des trolls, les développeurs d'IA conversationnelles parmi les plus évoluées sont particulièrement vigilants sur la correction de leurs erreurs, contradictions et surtout biais, et ce dans toutes les langues. C'est sans doute pourquoi Duplex n'est pas encore disponible en français, même s'il a commencé à s'étendre en 2020 en espagnol.

³⁴ <https://www.itforbusiness.fr/les-nouvelles-ia-de-google-et-microsoft-comprennent-mieux-le-langage-parle-que-les-humains-42054>

³⁵ <https://www.igen.fr/ailleurs/2018/05/google-duplex-sidentifiera-avant-un-appel-comme-un-robot-103937>

³⁶ <https://www.presse-citron.net/lintelligence-artificielle-de-microsoft-devient-profondement-raciste-en-moins-de-24-heures/>

Chatbot IA : comment ça marche ?

Les chatbots de nouveau semblent pleins de promesses. Mais comment est-on parvenu à les faire passer à l'âge adulte ?

En y injectant de l'**Intelligence Artificielle à base de traitement du langage naturel**, en entrée (*Natural Language Understanding*), en analyse (*Natural Language Processing*) et en sortie (*Natural Language Generation*), permettant aux chatbots d'interagir toujours plus finement avec leurs interlocuteurs :

- **En matière d'expression orale**, ils ont considérablement gagné en fluidité et en expressivité. Faites le test vous-même sur la démo de IBM Watson Text to Speech.
- **En matière de compréhension du langage humain**, ils l'appréhendent de mieux en mieux et dans des situations plus complexes, comme avec plusieurs utilisateurs ³⁷, ou à l'écrit sur des échanges moins académiques, sur du langage courant, avec des abréviations ou des fautes d'orthographe.

Ainsi, **la force des chatbots IA est d'être capable d'extraire et d'analyser l'intention de recherche depuis le flux de conversation de son interlocuteur, afin de lui envoyer une réponse adaptée**. Cette IA requiert un apprentissage exigeant sur de grands volumes de données, mais offre une expérience sans précédent, bien plus conviviale et efficace.



© Denis Lefevre inspiré par Dirit

³⁷ <https://www.youtube.com/watch?v=RZNqSy-zFXo&feature=youtu.be>

Seuls les chatbots IA peuvent traiter les 80% de données dites « non structurées », à savoir toutes les données qui ne peuvent pas être facilement entrées dans un tableau Excel. Mais les 20% d'autres données dites « structurées » (par exemple celles contenant nom, adresse, numéro de commande etc.) peuvent être gérées à un niveau moindre. Dès lors, **des chatbots moins sophistiqués que les purs chatbots IA peuvent bénéficier des progrès du NLU et du NLP** pour collecter et interpréter un nombre non négligeable de données qui transitent à travers eux, comme nous le verrons ci-après.

Ce sont majoritairement ces chatbots qui sont aujourd'hui utilisés par les marques pour proposer une expérience conversationnelle non décevante, et en attendant que des solutions plus avancées de chatbots IA se démocratisent commercialement.

Quoiqu'il en soit, **la révolution des IA conversationnelles est en marche, à l'oral comme à l'écrit**. La croissance des chatbots devrait augmenter de **+34% par an d'ici 2024** ³⁸, alors que le site chatbots.org recense aujourd'hui déjà plus de **1350 solutions dans le monde pour créer son chatbot, pour tous les besoins et tous les secteurs d'activité**.

L'adoption est à présent réelle : en 2020, 74% des consommateurs ont déclaré avoir déjà utilisé des assistants virtuels pour rechercher ou acheter des produits et services ³⁹.

LES CHATBOTS AU SERVICE DE LA RELATION CLIENT

Les chatbots et voicebots dotés d'Intelligence Artificielle sont aujourd'hui l'une des applications les plus concrètes et les plus répandues de l'IA marketing.

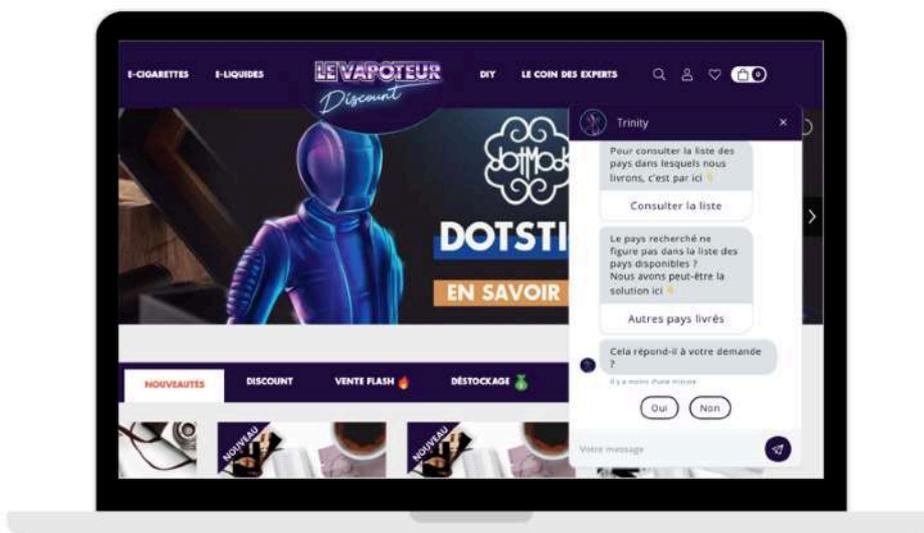
Le chatbot est à présent **un vrai assistant virtuel capable d'accompagner le client tout au long du tunnel d'achat**, nous faisant passer à **l'ère du e-commerce conversationnel**. L'expérience se rapproche du point de vente physique en ce qu'elle tire partie des avantages de proximité de la relation client, et ce à toutes les étapes :

³⁸ <https://www.keyrus.com/fr/fr/categorie/opinion/post/7015/infographie-5-tendances-des-bots-en-2020>

³⁹ https://www.capgemini.com/wp-content/uploads/2019/09/Report---Conversational-Interfaces_Web-Final.pdf

Chatbot d'avant-vente

- Le chatbot e-commerce participe à votre stratégie inbound marketing en proposant une **expérience omnicanale, à la fois être déployée sur Facebook Messenger et sur d'autres canaux** (Twitter, WhatsApp, Slack, Alexa, Google Assistant.. et bien sûr sur et vers votre site web). Le chatbot peut aussi **gérer des conversations multilingues** pour vos visiteurs étrangers, et même si votre site n'est traduit qu'en anglais. 110 langues sont ainsi disponibles via la solution de chatbots Mindsay.



Plug-in web



Facebook Messenger

- Ce n'est plus seulement votre site e-commerce qui est ouvert 7j/7 24h/24, mais aussi et surtout le **support client qui n'est plus limité aux seuls horaires d'ouverture**. Or, une grande majorité des achats en ligne se fait le soir et le week-end. Le chatbot est un **canal en libre-service** ouvert à n'importe quelle heure du jour et de la nuit. Il peut en outre traiter plusieurs conversations en simultanément, **sans temps d'attente**. C'est ce critère de vitesse et de réactivité qui le fait privilégier par 69% des consommateurs, notamment les Millenials et la Génération Z davantage portés sur la messagerie que sur l'appel au support technique ⁴⁰.

⁴⁰ <https://www.salesforce.com/blog/2018/01/why-consumers-prefer-chatbots.html>



- En se rendant **disponible sans s'imposer**, il laisse à l'internaute l'autonomie dont il a besoin pour visiter votre site, tout en l'aidant à se sentir pris en charge en cas de problème. Le chatbot peut ainsi **réduire significativement votre taux de rebond**. L'interaction à ce moment critique est clé pour espérer pouvoir conclure la vente.
- La première application est de **répondre aux 80% de questions génériques que se posent tous les visiteurs** (livraison, frais d'expédition, retours...) qui sont en général dans la FAQ, mais que le chatbot permet de rendre dynamique et interactif.
- Mieux que cela : comme en magasin physique, le chatbot peut **vous guider de façon personnalisée** vers le produit que **vous** recherchez. Il peut vous guider vers la bonne référence, y compris si vous ne savez pas ce que vous voulez, en vous interrogeant pour comprendre votre besoin. Il se base sur des **algorithmes de recommandations dynamiques**, qui vont se nourrir d'autres requêtes clients ayant le mieux performées, ou des produits que la marque aura choisi de mettre en avant.
- En vous faisant parler, ou simplement en analysant votre comportement sur le site (produits consultés, temps de navigation, achats précédents...), **le chatbot obtient des**



informations précieuses sur votre degré de maturité, et peut déterminer le bon message à vous adresser : recommandations, remises ou promotions. Cette segmentation quasi immédiate est une mine d'or pour maximiser le potentiel de conversion, bien au-delà de ce que pourrait faire un vendeur en magasin.

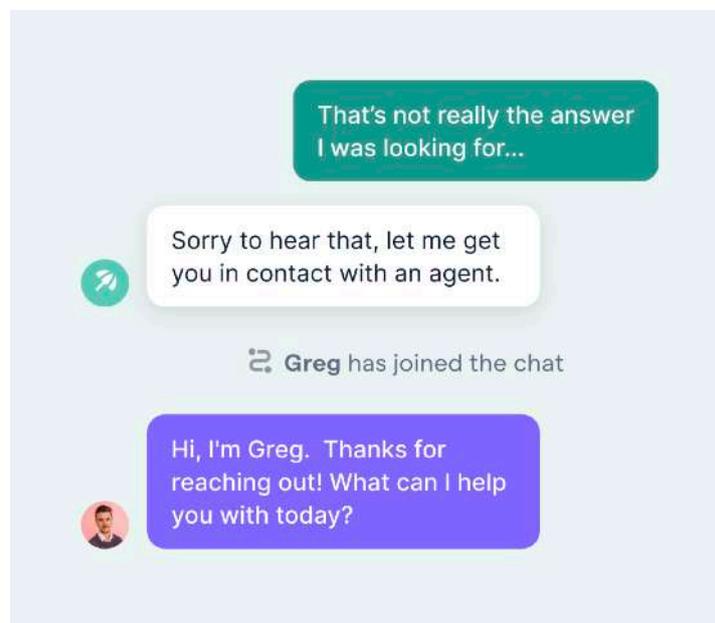
- Dans cette phase d'avant-vente, **le chatbot peut amener le visiteur jusqu'à l'opt-in** : ainsi le chatbot « My hand expert » de l'Occitane a réussi à faire s'inscrire 22% de ses interlocuteurs à sa newsletter, et à récupérer 69% de leurs coordonnées de contact. Parmi eux, 35% ont utilisé en ligne ou en point de vente le code promo fourni. ⁴¹



- En tant qu'entreprise B2B, le chatbot vous permet de **qualifier et de scorer vos prospects bien plus efficacement que le traditionnel formulaire en ligne**, et aide votre service commercial à conclure plus rapidement des contrats.
- Que le chatbot pré-qualifie les besoins ou les traite en direct, le visiteur se sent rassuré de savoir que le chatbot pourra toujours relayer sa requête à un opérateur

⁴¹ <https://numberly.com/fr/successstory/le-chatbot-my-hand-expert-occitane/>

réel, s'il préfère finalement parler à un humain, ou si sa demande devient trop spécifique. L'ensemble des informations déjà collectées pourront être transmises à l'opérateur pour résoudre son problème. **Le chatbot est l'illustration parfaite de « l'intelligence augmentée », qui ne se suffit pas à elle-même au point d'être une intelligence « artificielle », mais qui travaille main dans la main avec les équipes de l'entreprise pour faciliter et optimiser leur travail, au service du client.**



Chatbot transactionnel

- Au-delà de vous recommander des produits et de vous faire parler pour récolter des données sur vos besoins et préférences, **le chatbot e-commerce peut lui-même assurer des transactions.** 1 internaute sur 5 se dit d'ailleurs prêt à acheter un produit ou un service par l'intermédiaire d'un chatbot ⁴².
- C'est ainsi que le chatbot de La Croquetterie (nutrition canine et féline) utilisant la solution Noci ⁴³ a enregistré **+16% de ventes en automatique, avec +10% de CA essentiellement sur de nouveaux clients.** En outre, le panier moyen a grimpé de **+25% grâce aux recommandations du chatbot sur des produits complémentaires,** lors du passage à l'achat.
- En plus de stratégies cross et up-sell, vous pouvez travailler la **relance des paniers abandonnés,** en moyenne de l'ordre de 70% ⁴⁴. D'après la



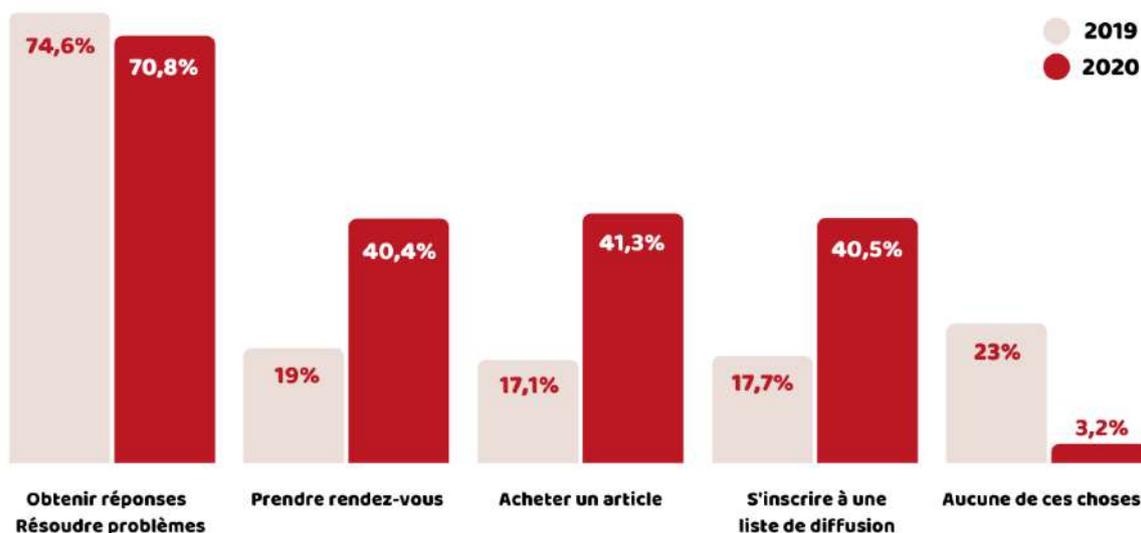
⁴² <https://www.ubisend.com/insights/2017-chatbot-report>

⁴³ <https://www.noci.io/la-croquetterie>

⁴⁴ <https://baymard.com/lists/cart-abandonment-rate>

plateforme Webotit, le chatbot peut vous permettre de **convertir 10 à 15% des transactions laissées en suspens** ⁴⁵. Le taux d'ouverture d'un message de chatbot serait de **80% contre 25% pour un e-mail, avec un taux de clics 10 fois plus élevé**.

- Dans le secteur B2B, le cas d'usage de l'achat aurait même plus que doublé en 1 an, avec une forte hausse également de l'opt-in et de la prise de rendez-vous ⁴⁶ :



© Denis Lefebvre, inspiré par State of Conversational Marketing 2020

- Grâce au chatbot, vous connaissez d'autant mieux vos clients, leurs intérêts et leurs besoins. **En le reliant à votre base CRM, vous pouvez enrichir en continu les datas manquantes ou erronées, au fil des discussions successives**. Vous maximisez ainsi le potentiel de vos données récoltées pour les convertir en ventes, tout en gagnant des critères objectifs pour **affiner vos offres et votre stratégie marketing**.

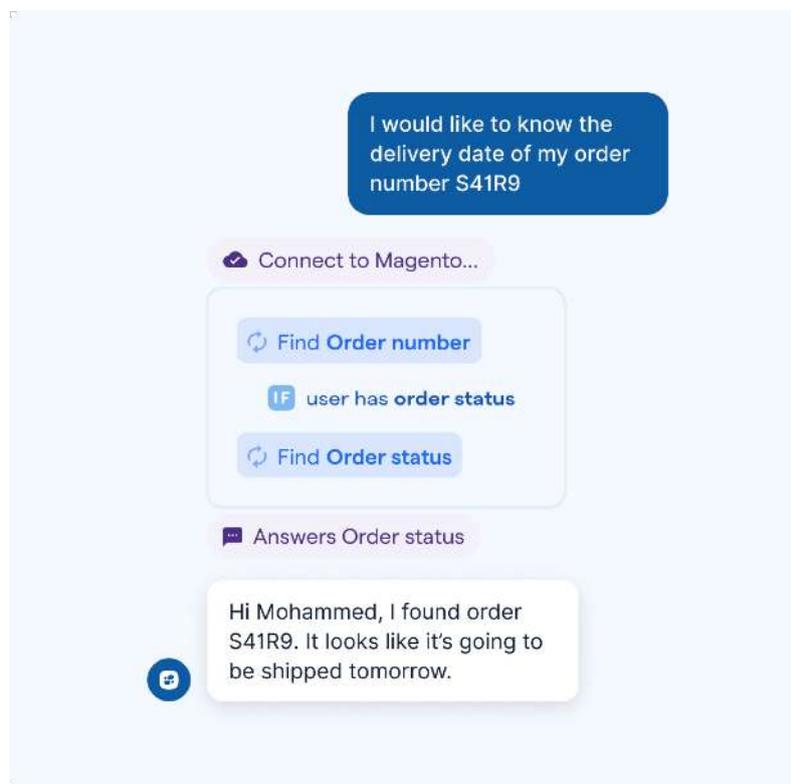
Chatbot SAV

- Une fois l'achat enregistré, le chatbot permet au client de gérer tout le service après-vente, et en premier lieu la commande en elle-même :

⁴⁵ <https://www.webotit.ai/marketing-conversationnel-5-tactiques-pour-votre-e-commerce/>

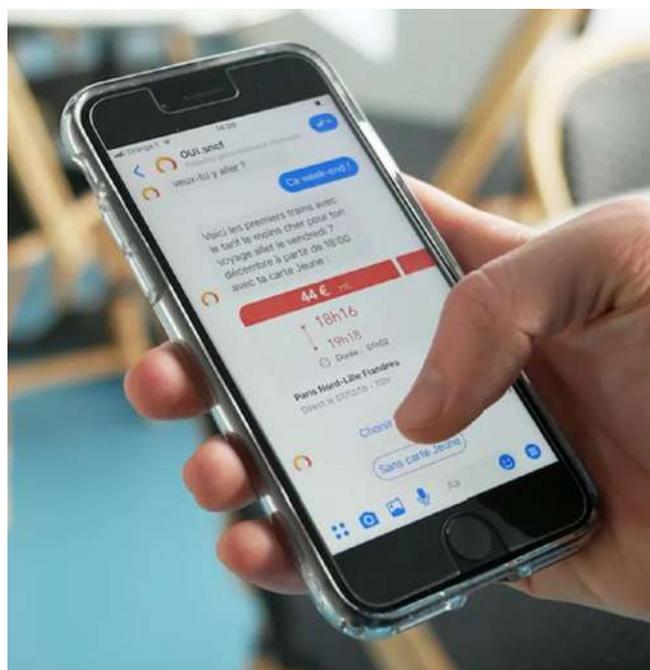
⁴⁶ <https://drift.com/wp-content/uploads/2020/09/2020-State-of-Conversational-Marketing.pdf>

- **Suivi de commande** : le chatbot connecté à la plateforme e-commerce peut donner de la visibilité sur son état d'avancement. Dans l'exemple ci-dessous, le client demande où en est sa commande. Le chatbot saura aller chercher sur la plateforme à laquelle il est connecté (ici Magento), grâce au numéro de commande, les **informations de livraison et de retard éventuel**. Le client en sera tout de suite informé, réduisant significativement le temps d'attente tout en permettant de réitérer la requête aussi régulièrement qu'il le souhaite.



- **Modification ou annulation de commande** toujours grâce **au système RPA (Robotic Automatisation Process)** via le chatbot, qui permet d'aller chercher les informations d'un client à partir de son numéro de commande, son nom ou tout autre identifiant. Le RPA permet alors d'envoyer la confirmation de modification ou de remboursement par email, tandis que le chatbot informe parallèlement le client. Le OUIbot de la SNCF permet déjà depuis fin 2017 d'acheter et de gérer ses réservations de train. OUIbot répond aujourd'hui à

plus de 13 000 personnes et envoie 400 000 confirmations de commande par mois via Messenger ⁴⁷.



- 3 clients sur 5 ne feront plus d'achat sur un site sur lequel ils ont connu une mauvaise expérience, et seuls 2% de clients insatisfaits se plaindront au service support ⁴⁸. L'enjeu du chatbot est de résoudre leur problème, mais également de **faire parler la majorité silencieuse pour identifier les points de blocage et diminuer le churn**.
- Le chatbot peut enfin **onboarder les nouveaux clients**, et leur adresser des messages personnalisés... jusqu'à leur souhaiter un joyeux anniversaire ! Un **programme de fidélité** peut aussi permettre d'engager les 92% de consommateurs qui s'y disent favorables ⁴⁹. 75% d'entre eux le considèrent même comme un critère de choix ⁵⁰. Tout ceci participe à améliorer votre rétention et votre satisfaction client.

L'un des principaux avantages du chatbot reste de **réduire les coûts du support client** :

- Les entreprises dans le monde dépensent 1 300 milliards de dollars en service client chaque année, sur 265 milliards de demandes. Le coût d'une interaction classique avec un opérateur humain serait estimé à 7 dollars pour une entreprise B2C, et 12 dollars pour une entreprise B2B ⁵¹.
- En étant capables de qualifier de façon toujours plus précise les demandes pour y répondre plus rapidement, et d'automatiser les réponses aux 80% de questions les

⁴⁷ <https://www.lesechos.fr/partenaires/capgemini/ouibot-fluidifie-toute-la-relation-client-1016040>

⁴⁸ <https://www.marketing-professionnel.fr/parole-expert/10-dogmes-regles-relation-clients-01-2011.html>

⁴⁹ https://nmgprod.s3.amazonaws.com/media/files/97/e3/97e3466268a5f6a39748b0acf861188d/asset_file.pdf

⁵⁰ <https://www.smallbizgenius.net/by-the-numbers/customer-loyalty-statistics/>

⁵¹ https://hbr.org/2017/01/kick-ass-customer-service?_ga=2.155885118.429925248.1613042402-2070295215.1613042402

plus courantes, **les chatbots pourraient aider les entreprises à réduire leur coût par interaction à 0,70 cents, soit 8 milliards de dollars économisés par an d'ici 2022** ⁵². En 2017, nous en étions à 20 millions déjà économisés, soit environ 30% des coûts de service client ⁵³. La solution Mindsay a même fait économiser 17€ par requête résolue à la RATP, en automatisant 84% de son support ⁵⁴. En somme, les entreprises augmentent leurs ventes tout en économisant sur leur SAV : un pari gagnant de bout en bout !

Je vous invite dans les pages suivantes à lire mon interview de Jonathan Kam, COO et co-fondateur du **chatbot hybride Botmind, qui offre aux e-commerçants une solution d'assistance virtuelle épaulée par des agents humains.**

Suite à quoi nous aurons terminé notre **tour d'horizon des principaux bénéfices de l'IA pour votre stratégie marketing**. Bien sûr, le sujet est si vaste que nous avons décidé, avec mon tuteur de thèse Thierry Pires et d'autres rédacteurs, d'en faire un site dédié : ia4marketing.fr, qui est déjà en train de devenir une référence.

L'Intelligence Artificielle se révèle pleine de promesses, pour le marketing comme pour nombre d'autres secteurs. Mais il s'agit dans notre Chapitre 2 d'explorer l'un de ses axes d'amélioration majeurs, relatif à la **consommation énergétique liée à ses calculs, et aux serveurs qui les hébergent**. L'IA responsable, cette « IA de confiance » revêt divers axes, comme la gestion des biais dans les jeux de données. Mais à l'heure où le marketing se veut éco-responsable (cf. Chapitre 3), **évaluons l'impact environnemental de l'IA, avant de pouvoir déterminer la façon d'assurer la durabilité du modèle d'IA marketing.**

⁵² https://www.juniperresearch.com/press/press-releases/chatbots-a-game-changer-for-banking-healthcare?_ga=2.260813064.429925248.1613042402-2070295215.1613042402

⁵³ https://www.ibm.com/blogs/watson/2017/10/how-chatbots-reduce-customer-service-costs-by-30-percent/?_ga=2.260813064.429925248.1613042402-2070295215.1613042402

⁵⁴ <https://www.mindsay.com/customers/ratp>



Jonathan Kam COO de Botmind



Vous annoncez avoir accompagné plus de 300 e-commerçants. Comment Botmind accompagne le visiteur sur leurs sites tout au long du parcours d'achat ?

« Une fois que le visiteur arrive sur le site, l'objectif va être de l'accompagner et faire en sorte qu'il devienne client. Et s'il est déjà client, ça va être de s'assurer qu'il ait la meilleure expérience possible pour le fidéliser. **Ce qu'on fait chez Botmind, c'est un assistant virtuel qui travaille sur la dimension avant-vente et après-vente** pour fournir des réponses à toutes les questions que les visiteurs ou les clients peuvent avoir.

En avant-vente, ça va être quelqu'un qui veut s'assurer de la disponibilité d'un produit, savoir s'il y a des livraisons possibles dans un pays, etc. Des réponses nécessaires avant de passer à l'achat. Si je ne suis pas sûr de pouvoir me faire livrer dans les DOM-TOM, alors je ne vais pas commander. **Pour le support client**, ça va être des questions très classiques : où est ma commande, j'ai oublié mon mot de passe, etc. Il y a parfois sur le site des éléments de réponse. Mais certains utilisateurs soit ne font pas l'effort, soit n'arrivent pas à trouver l'information. On est donc là pour les assister.

Le conversationnel est une grosse tendance : aujourd'hui, l'utilisation de WhatsApp et Facebook Messenger dépasse l'utilisation régulière de Facebook notamment. De la même façon, **sur les sites e-commerce le tchat est le canal qui a la plus forte croissance ces dernières années**. Les utilisateurs passent beaucoup plus par ce canal que par le téléphone ou le mail s'ils ont le choix. Par ailleurs, **en utilisant la data disponible via le**

I.A + Agents humains
sur le même canal.



botmind
L'AGENT VIRTUEL HYBRIDE

canal conversationnel, l'assistant virtuel va à la fois améliorer l'expérience client et réduire le funnel d'achat en rapprochant de la page la plus proche du panier, multipliant ainsi la probabilité d'achat.

Vous êtes sur un modèle de chatbot hybride. De quelle façon l'intelligence artificielle intervient-elle dans votre solution ?

Ce qu'on fait chez Botmind, c'est qu'on analyse le texte. On se base sur des **algorithmes de NLP qui permettent de définir un seuil de confiance auquel on peut attribuer une possibilité de réponse**. Si on est sûr à X% de comprendre la question d'un visiteur, alors on va ou non déclencher une réponse. Le fait d'être sur le modèle hybride permet d'avoir de l'**automatisation pour tout ce qui est fréquent et chronophage**, et rediriger vers un agent humain pour les demandes à forte valeur ajoutée, celles qui demandent de l'empathie, ou quand l'agent virtuel n'a pas la réponse. **45% au moins des questions** peuvent aujourd'hui être traitées par Botmind **sans aucune intervention humaine**. Sur certains sites ce seuil peut être plus élevé.

L'hybride reste selon nous le modèle le plus adapté, car il y aura toujours une limite à des réponses parfaites délivrées par l'IA, même si la technologie a beaucoup évolué. Il y a 5-6 ans, les chatbots étaient beaucoup basés sur des mots-clés. Les limites ont été très vite atteintes. Si je demandais « où est ma nouvelle commande » et que le chatbot avait pour objectif la vente, il pouvait par exemple juste retenir « nouvelle commande » et recommander un produit. Alors que ce n'était pas la question. On est plutôt à présent sur de la NLP qui analyse chaque mot individuellement, puis les mots pris 2 à 2, 3 à 3,

tous dans leur ensemble. Ce **data set va entraîner l'algorithme pour aller piocher dans notre répertoire de questions/réponses**, que nous alimentons avec nos clients pour enrichir sans cesse l'agent virtuel.

Les données collectées permettent-elles de reconnaître le client sur sa prochaine visite ?

On essaye de tendre vers un produit qui n'est pas que sur la première utilisation, mais sur la fréquence. Il y a quand même deux raisons pour lesquelles c'est assez complexe, et qui expliquent pourquoi on le fait de façon progressive : des raisons techniques, et l'aspect réglementaire sur la data. Il y a des clients pour lesquels on a construit **un modèle qui reconnaît l'utilisateur quand il revient**. On sait s'il est logué ou non. On peut connaître ses 3 dernières commandes s'il utilise un Shopify ou un PrestaShop. Du coup, si on doit lui faire des suggestions, on ne va pas lui suggérer les 3 derniers produits qu'il a déjà achetés. Donc **on peut personnaliser le parcours d'achat**, mais seulement si nos clients utilisent des plateformes grand public telles que Salesforce, Oracle, PrestaShop, Shopify ou Magento. L'**aspect réglementaire** impose aussi bien sûr de s'assurer que les données sont stockées selon des critères RGPD.

Combien ça coûte d'utiliser Botmind ?

On facture à la performance, selon les automatisations réussies. Un panier moyen va être chez nous **autour de 700€ par mois**. Les clients qui n'ont que 5 questions par jour ou par mois ne vont pas être intéressés par Botmind. Mais quand ils commencent à recevoir 1000 questions par mois, là Botmind devient intéressant.

Parce que si on parle d'automatiser 45%, ça fait gagner la moitié du temps passé à faire du support client. Nos clients peuvent donc **faire plus de ventes en se concentrant sur la résolution de problèmes en amont, et sur l'optimisation du parcours client**. Avec l'Intelligence Artificielle il y a de vraies externalités positives, qui ne sont pas directes. »

AUTOMATISEZ 45% DE VOTRE SUPPORT CLIENT EN 7 SEMAINES



RÉDUCTION DES COÛTS DE SUPPORT

L'agent virtuel automatise les réponses aux demandes fréquentes et chronophages dans l'e-commerce et vous fait économiser jusqu'à 60% dans le traitement d'un ticket support.



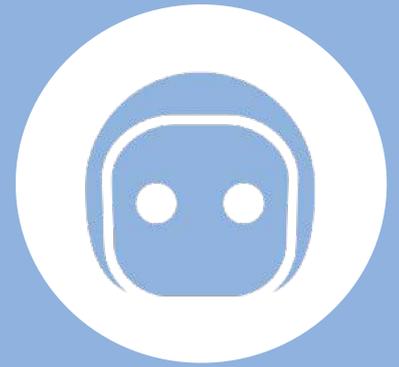
RÉPONSES EN TEMPS RÉEL

L'agent virtuel est déjà programmé pour détecter plus de 100 questions de support client e-commerce fréquentes dont les réponses sont immédiatement automatisables.



GESTION DES PICS DE DEMANDES

L'agent virtuel permet de gérer plus facilement et sans surcrot les pics d'activité dans l'e-commerce.



+26%

satisfaction client

Vestiaire Collective

SB SAVEUR BIÈRE
ENSEMBLE DÉCOUVREONS LA BIÈRE

BIRCHBOX

D DEVIALET

My Jolie Candle

+8%

taux de conversion

-60%

coûts de support

FAITES PASSER VOTRE TAUX DE CONVERSION DE 3% À 8% EN 3 SEMAINES !



PRÉ-QUALIFICATION DES BESOINS

L'agent virtuel pose des questions orientées pour comprendre les besoins des visiteurs et ainsi y répondre. Botmind transforme vos visiteurs en clients.



CONSEILS PERSONNALISÉS

L'agent virtuel Botmind donne des conseils personnalisés en quelques clics : tailles, nouveautés, couleurs, modèles, information sur les stocks etc.



AIDE À LA VENTE

L'agent virtuel interagit avec vos visiteurs avant qu'ils ne quittent votre site. Créez de la confiance avec vos visiteurs pour les convaincre d'acheter chez vous.

SENA

BEAUTY SUCCESS

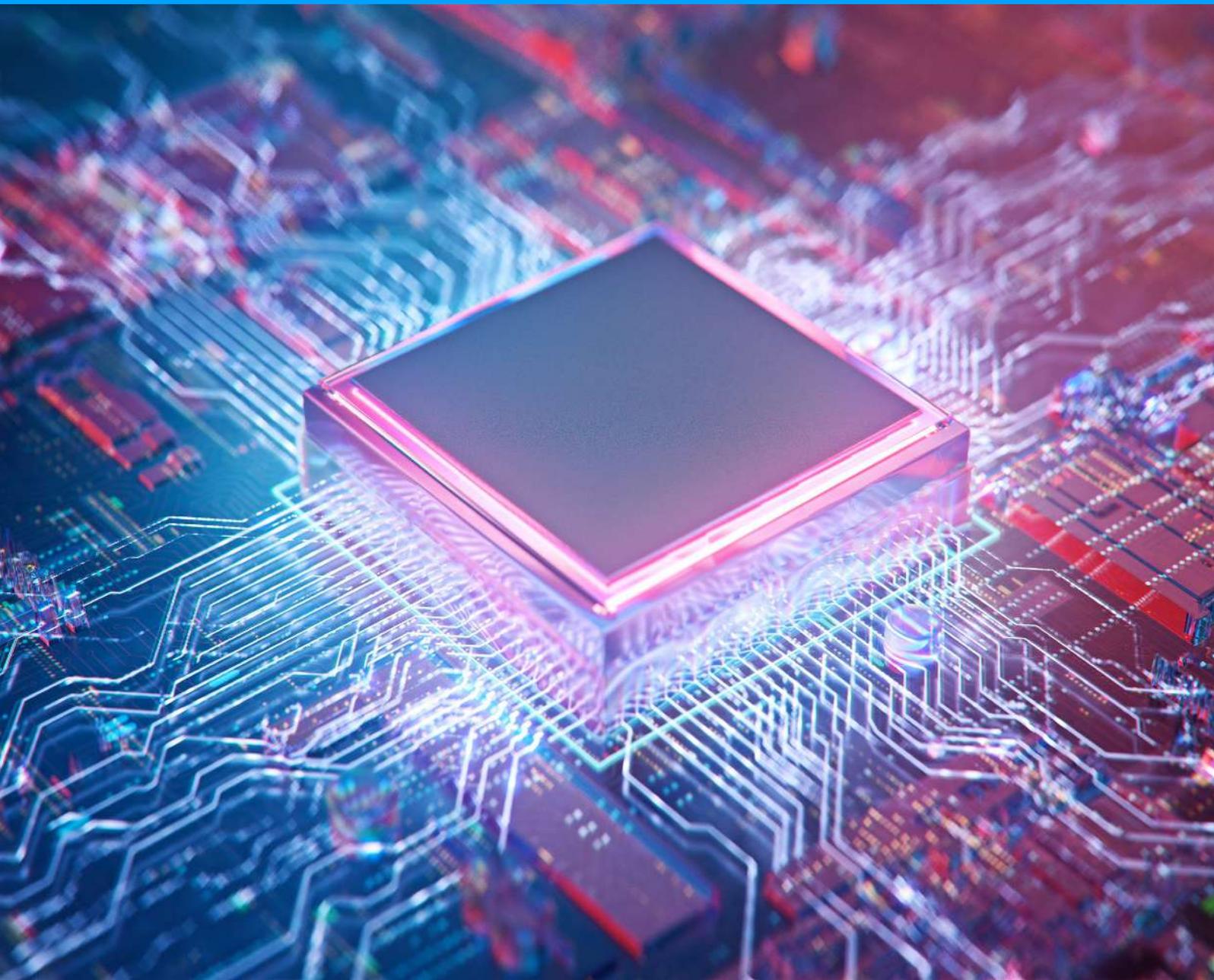
Lamazuna

MUZÉO

COPEES

CHAPITRE 2

L'IA face au défi de son bilan carbone



L'Intelligence Artificielle est **une pièce maîtresse pour aider les organisations à prendre de meilleures décisions**, en exploitant toute la valeur des données à leur disposition.

Mais les progrès de l'IA, considérables depuis les cinq dernières années, et poussant à des modèles toujours plus complexes, **ont un coût**. Un coût non pas tant économique : c'est justement la baisse du prix du stockage et des capacités de calcul qui ont permis son essor. Non, un coût perçu comme bien moins prioritaire pour les entreprises, mais **qui devient un vrai problème, et qu'il n'est à présent plus possible d'ignorer : son coût écologique**.

L'IA n'est pas la seule à blâmer : **le numérique au sens large, au niveau mondial, serait déjà responsable de 3,5% des émissions de gaz à effet de serre** en 2019, avec un taux de croissance de l'ordre de **+6% chaque année depuis 2013**. D'ici 2040, le numérique serait ainsi responsable de 24 millions de tonnes de CO₂, soit 7% des émissions de la France, contre 2% à l'heure actuelle ⁵⁵.

Pourtant, grâce à ses modèles prédictifs, **l'IA a aussi la capacité de nous aider dans la lutte contre le réchauffement climatique**.

Côté pile, elle risque donc de peser dans le bilan carbone à l'échelle mondiale, mais aussi dans la transparence carbone à laquelle votre entreprise va être de plus en plus soumise dans les années à venir (cf. Chapitre 3). **Mais côté face, elle peut aussi aider à réduire l'impact des entreprises sur l'environnement !**

C'est cette double facette de l'IA qui me fascine, son ambivalence : un « Hulk » à la fois surpuissant et potentiellement dévastateur s'il n'est pas maîtrisé, mais qui n'en reste pas moins « vert »... Nous allons le voir, **le développement de l'IA est à la fois un risque et une opportunité dans le combat pour la transition écologique**.

⁵⁵ https://theshiftproject.org/wp-content/uploads/2021/03/Note-danalyse_Numerique-et-5G_30-mars-2021.pdf

Une IA de plus en plus énergivore

L'objectif de l'Intelligence Artificielle est de reproduire l'Intelligence Humaine. Pourtant, **là où le cerveau humain n'a besoin que de 20 Kw/h d'énergie pour fonctionner, il en faut déjà au moins 450 Kw/h pour un simple ordinateur...** et bien plus pour gérer tous les périmètres d'hébergement et de calculs des IA.

La totalité des données stockées dans le monde représente aujourd'hui 33 zettaoctet (soit 33 mille milliards de gigaoctets), et on devrait attendre le chiffre pharaonique de **175 zettaoctet de données d'ici 2025** ⁵⁶. La collecte intensive et la multiplication des usages liés à l'Intelligence Artificielle se fait dans tous les secteurs, permises par des avancées technologiques telles que la 5G ou l'Internet des Objets (IoT).

Alors, **pourquoi et combien l'IA consomme-t-elle réellement ?** Vers quoi se dirige-t-on ?

Un marché mondial de l'IA en milliards d'ici 2025



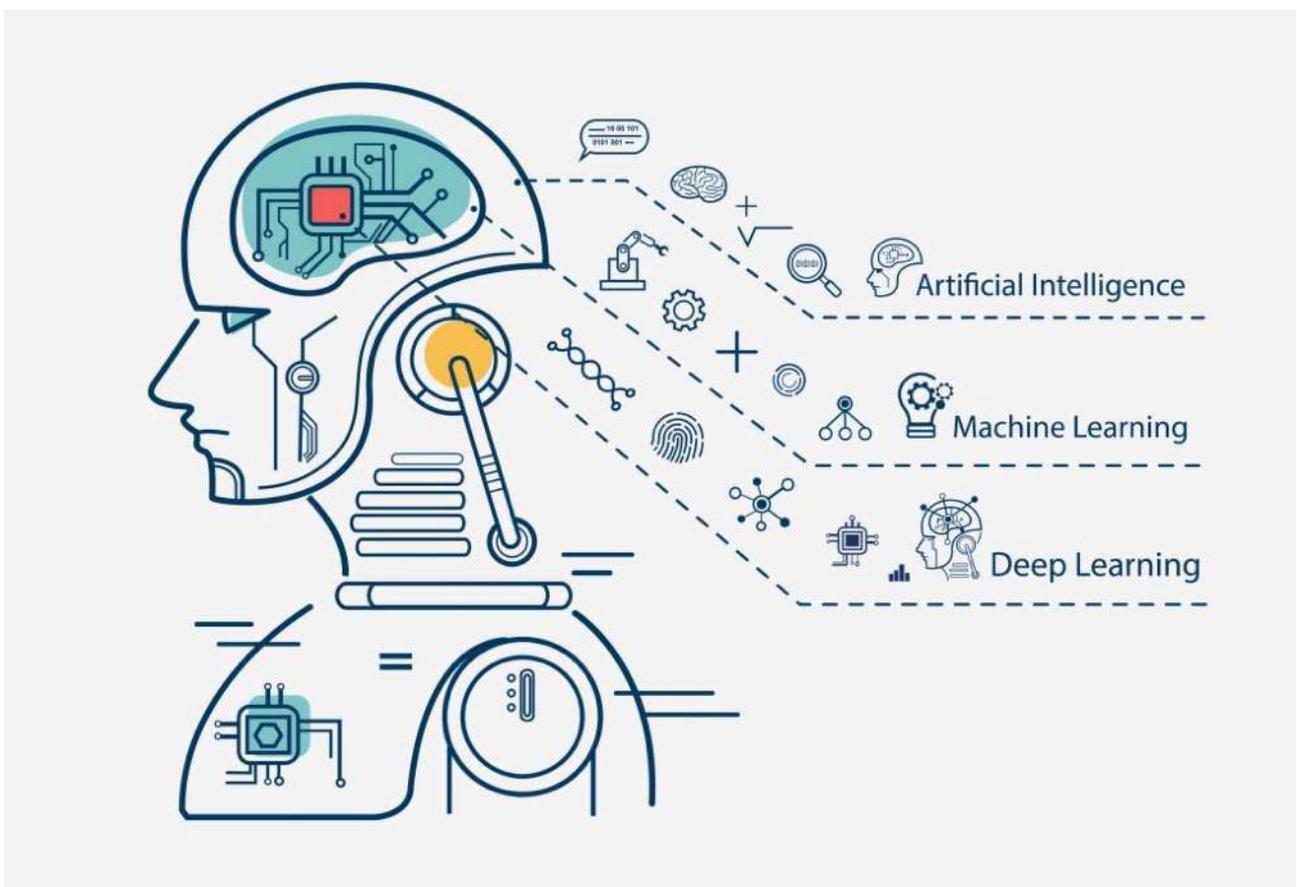
© Denis Lefebvre, inspiré par Tractica et Statista

⁵⁶ <https://www.journaldunet.com/solutions/dsi/1424245-le-volume-de-donnees-mondial-sera-multiplie-par-45-entre-2020-et-2035-selon-statista/>

1. Qu'est-ce qui fait que l'IA consomme ?

Pour **comprendre pourquoi l'Intelligence Artificielle requiert beaucoup de besoins énergétiques**, il faut déjà revenir à ses fondements, ses évolutions ainsi que son mode de fonctionnement.

1.1. IA, Machine Learning et Deep Learning



UNE BRÈVE HISTOIRE DE L'IA

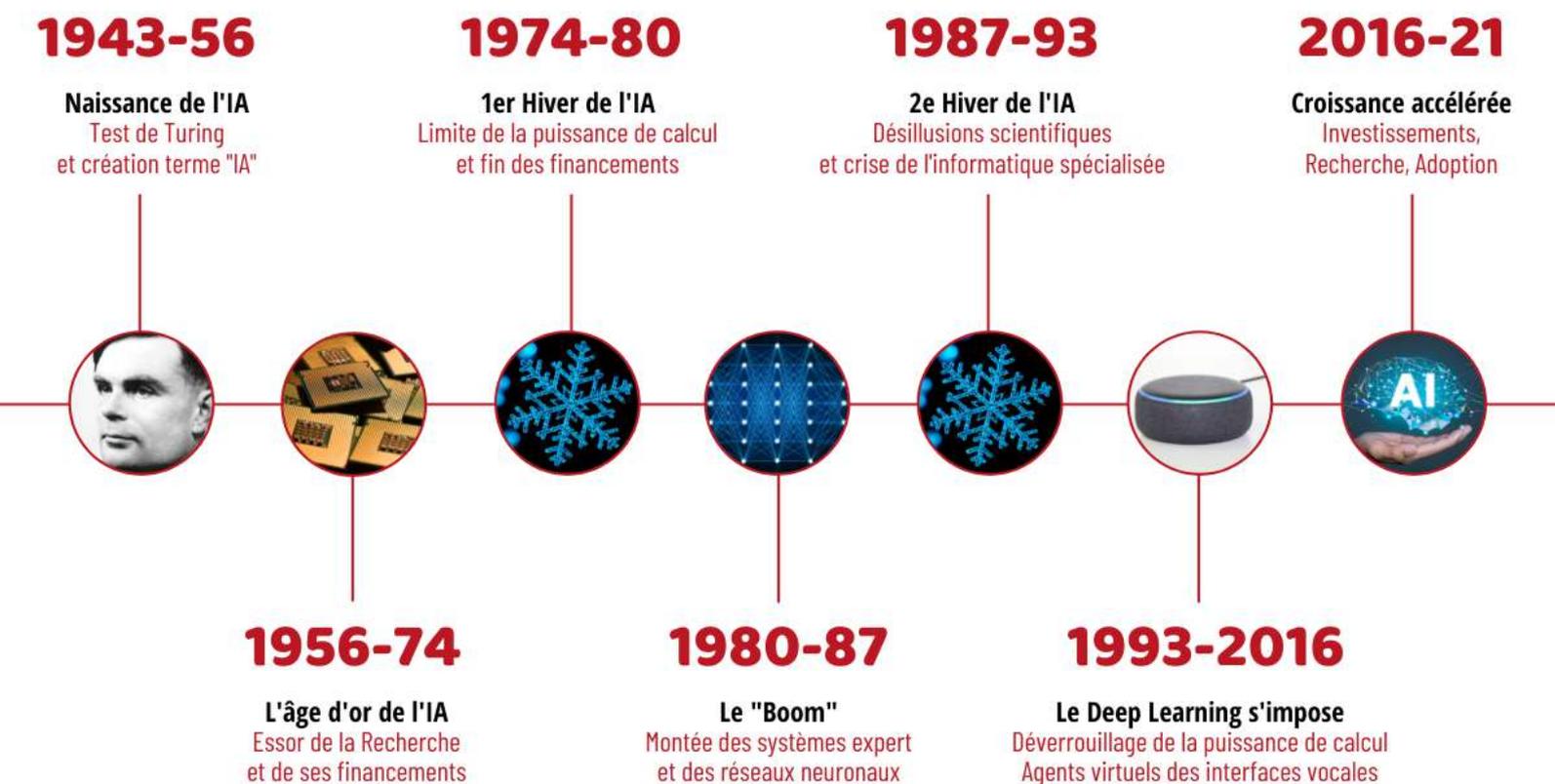
Le terme « **Intelligence Artificielle** » a été introduit par John McCarthy en 1956 lors de la conférence du Dartmouth College, aux Etats-Unis. Il désigne **une discipline scientifique**

visant à reproduire l'intelligence humaine par le biais de programmes informatiques, en s'inscrivant à la suite du « Test de Turing » évoqué précédemment.

S'en suit dès lors une histoire de plus d'un demi-siècle, résumée dans l'infographie ci-dessous. L'histoire de l'IA est faite d'**avancées significatives entrecoupées de ce qu'on appelle des « hivers »**. Ces hivers sont liés au manque de puissance des calculs et des capacités de stockage et de transfert des données, carburants primaires de l'IA.

À partir des années 1990, et surtout depuis les années 2010, l'IA a pu se déployer et se structurer sous l'effet de la diminution drastique du coût lié conjointement :

- **Au Cloud** qui permet de passer de systèmes d'hébergement en local très coûteux, à une capacité de stockage illimité et à un prix marginal ;
- **À l'augmentation de la vitesse et de la puissance de calcul** grâce à l'émergence des processeurs GPU (*Graphical Processing Unit*) dont nous reparlerons ;
- **À l'accélération des possibilités de transfert** grâce à la fibre, la 4G et à présent la 5G.



L'évolution de l'Intelligence Artificielle est donc fondamentalement liée à un besoin de puissance et d'agilité, **sollicitant des technologies fortement consommatrices d'énergie**, qui lui permettent progressivement de franchir des étapes.

Elles lui ont permis de **passer des systèmes expert au Machine Learning, et à partir de 2012 au sous-domaine du Deep Learning**, qui s'impose depuis (cf. définitions p.107).

Suivant la Loi de Moore ⁵⁷, **les progrès de l'IA devraient continuer d'être exponentiels**, la multiplication des capacités informatiques multipliant ses usages et applications, créant à leur tour de nouveaux besoins et de nouvelles capacités.

Avec l'ère du Deep Learning, les ressources de calcul pour aboutir à des modèles de premier ordre ont doublé tous les 3,4 mois, soit un facteur de x300.000 entre 2012 et 2018. C'est le prix à payer pour chercher à reproduire les capacités du cerveau humain, en bâtissant pour cela des réseaux de neurones artificiels.

LE FANTASME DE L'IA MODERNE : IMITER LE CERVEAU HUMAIN

Le cerveau humain fait naturellement appel au Big Data : il perçoit en moyenne 30 images par seconde, soit 1800 images par minute et plus de 600 million par an. **L'IA cherche à en faire de même** pour permettre par exemple la reconnaissance faciale.

Mais là où le cerveau humain sait différencier très rapidement deux objets distincts dès le plus jeune âge, **la machine elle ne peut pas faire preuve du même instinct.** Elle doit être entraînée pour le faire, et **son apprentissage est bien moins efficace que le notre.**

L'IA est donc bien par définition une Intelligence « **Artificielle** », à savoir qu'elle est le **produit d'une activité humaine.** Comme elle n'a aucune conscience d'elle-même, **elle ne peut en outre faire preuve d'aucune « Intelligence ».** C'est ce que rappelle Luc Julia,

⁵⁷ <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203331-loi-de-moore-definition-traduction/>

co-créateur de Siri et Vice-Président de l'innovation chez Samsung, dans son livre au titre volontairement provocateur : *L'Intelligence Artificielle n'existe pas* ⁵⁸.

À travers des films tels que *I, Robot*, *Her* et bien d'autres, Hollywood a très largement alimenté cette **fantasmagorie d'une machine qui deviendrait un jour plus intelligente que nous**, atteignant « l'ère de la singularité » et le stade d'une IA « forte », capable de penser comme nous, mieux que nous, et menaçant de nous remplacer.

Pourtant, l'IA ne saurait être autonome. Elle fait ce pour quoi elle a été entraînée. Cette IA « faible » est préférée par Luc Julia comme étant **une « Intelligence Augmentée »**, un **outil capable de faciliter notre travail** sur des tâches bien définies, pour lesquelles il se révèle au final plus efficace que nous. Cette Intelligence Augmentée nous permet de **traiter bien mieux que nous ne le ferions nous-mêmes de gros volumes de données**, ou que nous ne serions même capables de le faire, et aboutissant à prendre de meilleures décisions. Il faut donc voir l'IA comme un allié, car comme le souligne Fred Cavazza :



Vous n'êtes pas en concurrence avec les machines, sauf si vous refusez de travailler avec elles.

Un peu comme un bûcheron qui refuserait d'utiliser une tronçonneuse ou un comptable qui préférerait tout calculer à la main : d'un côté ils ne travailleront pas mieux, et de l'autre ils seront plus lents et moins précis.



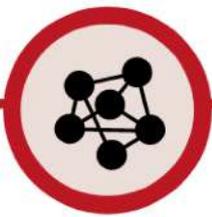
Fred Cavazza Conférencier et consultant en transformation digitale

2 méthodes d'apprentissage artificiel ont été créés pour simuler l'intelligence humaine :

- **L'apprentissage symbolique**, qui a été le premier pilier de l'IA à travers les moteurs de règles et les systèmes experts. Il crée des « arbres de décision » en s'appuyant sur une programmation logique et démontrable, pour automatiser des tâches de manière formelle et maîtrisée (cf cas d'application avec Golem.ai en p.245).

⁵⁸ <https://www.leslibraires.fr/livre/15270338-l-intelligence-artificielle-n-existe-pas-luc-julia-first>

- **L'apprentissage automatique** également appelé « connexionniste », basé sur des systèmes auto-apprenants qui s'entraînent à partir de grands volumes de données, et construisent des probabilités très efficaces mais bien moins faciles à expliquer (ce qu'on appelle la « boîte noire » du Deep Learning). Là où l'apprentissage symbolique exécute un programme informatique selon des paramètres établis par un humain, l'apprentissage automatique lui détermine ses propres paramètres pour parvenir à un résultat. Ses modèles continuent de se perfectionner après avoir été entraînés.



Apprentissage auto

apprend au fil du temps et des données



Perception

interprète la signification des données, notamment le texte, la voix et les images



Cognition

aboutit à des conclusions, décisions, actions voire interactions

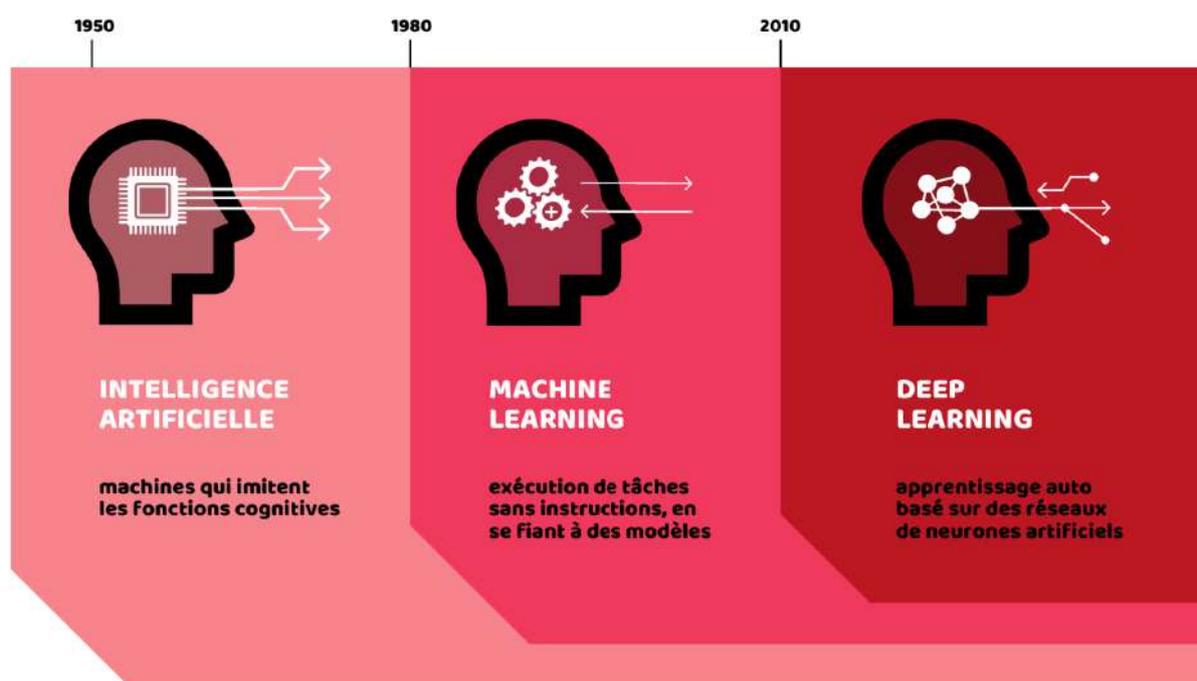
© Denis Lefebvre

L'apprentissage automatique domine les avancées récentes de l'IA, via deux méthodes :

- **Le Machine Learning** ou « Apprentissage Machine » :
 - Cette méthode consiste à créer et entraîner des algorithmes pour analyser, comprendre et prédire à partir des données.
 - Elle s'appuie sur une **méthode statistique pour regrouper, au sein d'un ensemble de données, des entités cohérentes** en vue de les segmenter/classifier, mais aussi **créer des prédictions** à partir de données historiques.
- **Le Deep Learning** ou « Apprentissage Profond », est un sous-domaine du Machine Learning qui accompagne l'essor de l'IA, plus à même de traiter des volumes de données toujours plus importants :
 - Cette méthode consiste à reproduire un cycle de décisions en se basant sur **des réseaux de neurones artificiels**. Ces réseaux sont dits « profonds » car ils

sont bâtis sur plusieurs couches de neurones, pour reproduire le traitement simultané d'informations complexes comme dans le cerveau humain.

- Ces neurones artificiels **se connectent les uns aux autres via des milliards de calculs informatiques, jusqu'à trouver la bonne configuration** (« paramètres ») pour répondre à une demande de traduction automatique, par exemple. À ceci près que là où un enfant qui apprend à parler aura entendu environ 45 millions de mots, un algorithme de Deep Learning tel que BERT, spécialisé dans le traitement du langage, aura besoin d'en assimiler 3 000 fois plus. Et là où BERT compte 1,5 milliard de paramètres, un autre algorithme NLP réputé tel que GPT-3 en compte 175 milliards... **Pour une précision toujours accrue, le Deep Learning pousse donc à des modèles toujours plus gros.** Mais alors que le cerveau d'un enfant finit par supprimer une partie de ses connexions neuronales non essentielles ⁵⁹, le Deep Learning ne fait pas encore preuve à ce jour de la même efficacité énergétique.
- Le Deep Learning est utilisé pour la reconnaissance d'images, la traduction automatique, mais aussi la robotique et bien d'autres champs d'application.



⁵⁹ <https://developingchild.harvard.edu/science/key-concepts/brain-architecture/>

1.2. Des modèles d'IA au besoin énergétique croissant

Pour parvenir à des modèles d'Intelligence Artificielle fiables et précis, il faut les nourrir de données, de beaucoup, beaucoup de données. Pour comprendre où se situe leurs besoins énergétiques, **essayons déjà de comprendre la façon dont ils fonctionnent.**

COMMENT SE CONSTRUIT UNE IA ?

Les modèles de Machine Learning suivent **un processus en 4 étapes**, qui visent à traiter des **données** pour les transformer en **informations**, à savoir en **données « intelligentes »**, et apporter une vraie valeur ajoutée à la stratégie Data recherchée par l'entreprise.

Phase n°1 : Préparation

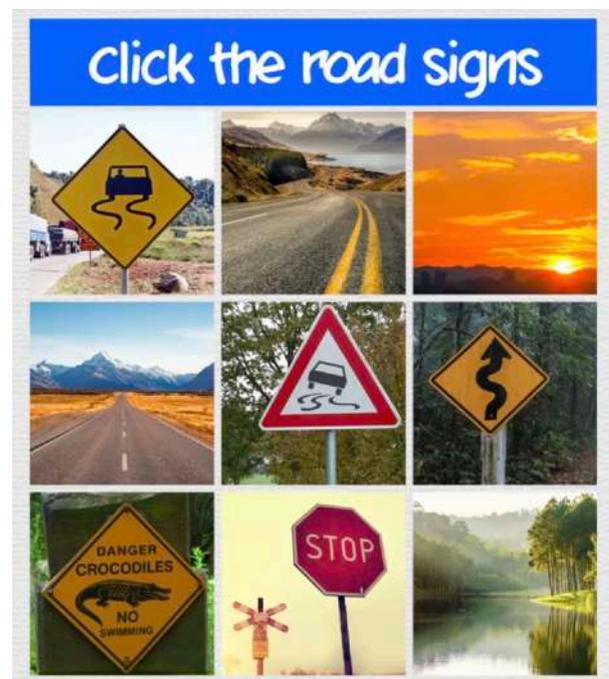
- Selon l'objectif fixé, **on sélectionne les données, le ou les algorithmes(s)** qu'on souhaite utiliser, et dans le cas du Deep Learning **l'architecture du réseau de neurones** sur lequel les entraîner.
- Comme nous allons le voir sur la phase n°2, l'apprentissage dit « supervisé » implique de **collecter et organiser les données avant de les entraîner.**
- **Le Data Labeling est une étape clé**, car des données mal étiquetées risquent de fausser les prédictions. Or, à ce jour, 60% des entreprises font face à des problèmes de qualité de données ⁶⁰. Mais les « biais » ne sont pas que dans la **précision du modèle** : ils sont aussi dans le risque de **reproduire des préjugés racistes, sexistes...** et toute forme de « discrimination technologique » ⁶¹ :



*Par exemple, les premiers algorithmes de reconnaissance faciale ne reconnaissent pas les peaux noires, car **un biais cognitif de la part des concepteurs** (majoritairement) blancs avait créé **un biais algorithmique.***



- **Ces biais sont l'un des défis majeurs de l'IA** dans son développement actuel, et au cœur des discussions **pour une IA responsable**. On mesure aisément le risque, par exemple, d'une IA d'aide au recrutement qui viendrait aggraver le problème de la discrimination à l'emploi. D'où l'importance de la **préparation éthique et transparente des données**.
- L'étiquetage vise à déterminer la « cible » que le modèle doit apprendre à identifier, en prédisant sa nature ou son comportement. Ainsi, **les données étiquetées contiennent les caractéristiques clés qui sont attendues**, en permettant de reconnaître un *pattern* (comme par exemple le visage d'une personne sur une photo), pour ensuite permettre de classer et regrouper ensemble tous les *patterns* identiques.
- L'étiquetage est opéré en premier lieu en manuel. C'est un travail laborieux : **un Data Scientist passe 80% de son temps sur la collecte, la préparation et le nettoyage des données** ⁶². Heureusement, des outils eux-mêmes bâtis sur du Machine Learning permettent à présent de l'assister, comme OpenRefine ou DataCleaner. L'outsourcing et le **crowdsourcing** sont aussi des méthodes très largement utilisées : sachiez-vous que quand on vous demande de prouver que vous n'êtes pas un robot, l'objectif est également d'entraîner les algorithmes à identifier des *patterns*, comme ici pour aider la reconnaissance visuelle des voitures autonomes ?!



⁶² Étude Cognilytica, 2020 : <https://www.forbes.com/sites/cognitiveworld/2020/02/02/the-human-powered-companies-that-make-ai-work/?sh=3b8d2d89670c>

Phase n°2 : Apprentissage

On entraîne les données selon 3 approches possibles :

- **Apprentissage supervisé :**
 - Comme évoqué à l'instant, la supervision humaine permet de **fournir au modèle les données type qu'il doit identifier**, incluant toutes ses variables. Les algorithmes sont donc entraînés avec des données qui contiennent la réponse attendue. Ils sauront donc **appliquer la même prévision aux données similaires qui leur seront ensuite soumises**.
 - Exemple d'utilisation : recommander des morceaux de musique sur la base de vos écoutes précédentes, vos likes et vos dislikes.
- **Apprentissage non supervisé**
 - Bien moins utilisé, il consiste à **ne pas faire d'étiquetage pour laisser le modèle chercher de lui-même** dans les données des éléments qui les rapprochent, et en faire des « clusters ».
 - Exemple d'utilisation : détecter des fraudes bancaires. Comme les transactions suspectes ne peuvent pas être connues à l'avance, et ne peuvent donc pas être préalablement étiquetées, l'algorithme va tenter d'identifier des anomalies pour détecter un risque de fraude.
 - Un **apprentissage semi-supervisé** permet quant à lui, une fois qu'un modèle supervisé est abouti, d'envoyer des données non étiquetées pour mettre le modèle à l'épreuve, et ainsi le perfectionner.
- **Apprentissage par renforcement :**
 - il s'agit d'un **apprentissage par la récompense**, qui fonctionne sur le principe du feedback en formation continue : si on donne au modèle une photo de chien et qu'il l'identifie comme étant un chat, on lui dit que ce n'est pas ça et il apprend de ses erreurs.

Phase n°3 : Paramétrage du meilleur modèle

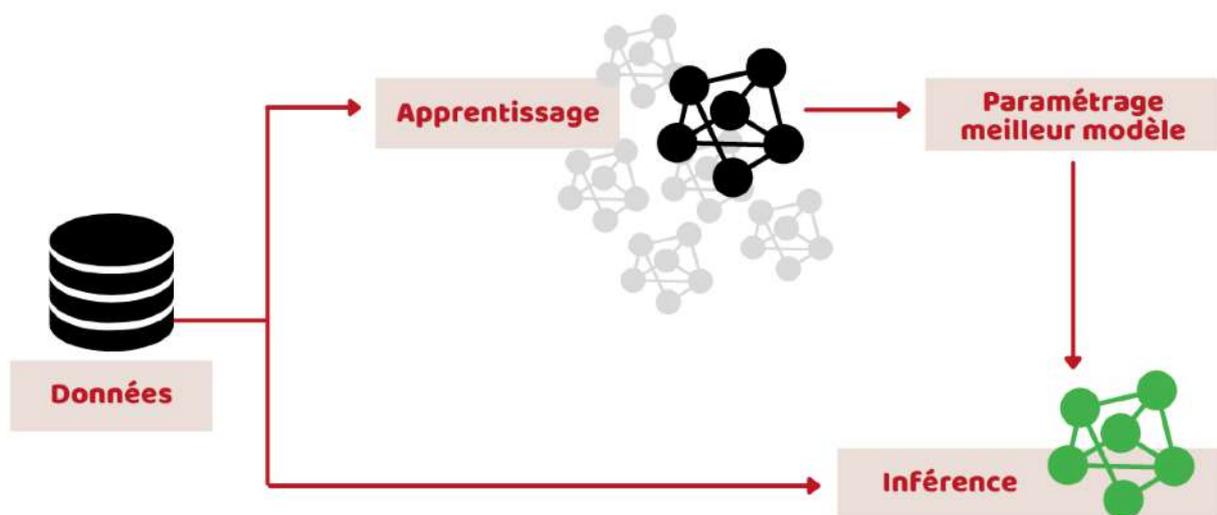
- Le principe du Machine Learning est de **laisser l'algorithme s'entraîner de façon itérative**, sur la base d'un jeu de données (« *dataset* ») initial, puis de nouveaux jeux de données à même d'éprouver ses résultats.
- Les **paramètres** constituent des **molettes d'ajustement** sur le modèle le plus abouti, parmi tous ceux qui auront été entraînés. On peut les régler jusqu'à obtenir les sorties souhaitées. **Plus il y a de paramètres, et mieux on maîtrise la précision du modèle.** Ils sont les rouages de la magie opérée par l'IA... car il y a bien un « truc » : les modèles de NLP, aussi perfectionnés qu'ils soient, ne comprennent pas la langue. Ils sont entraînés pour prédire le prochain mot à suivre. De nouveau, « l'Intelligence » Artificielle reste un abus de langage...
- Toujours est-il que les itérations successives du modèle permettent d'**aboutir à une prédiction parfaite sur les données observées.**

Phase n°4 : Inférence

- Une fois que le modèle est arrivé au bout de son cycle de formation initiale, il est ensuite **déployé en conditions réelles.**
- C'est ce qu'on appelle « **l'inférence** » : ce moment où l'IA passe de sa phase de « Recherche et Développement » à sa phase de « mise en production ». Elle va dès lors **se confronter à la vérité terrain grâce aux données récoltées en temps réel, qui viennent compléter les données d'apprentissage.** Elle peut ainsi re-calibrer ses paramètres en conséquence, et prendre en compte de nouvelles variables d'ajustement.
- **Prenons un cas concret : celui de la voiture autonome,** dont l'un des enjeux majeurs est bien sûr l'optimisation de la sécurité.

- Un algorithme est modélisé pour améliorer, disons, ses performances dans un virage, en prenant en compte diverses données : pression et usure des pneus, conditions de la route, etc.
- Le modèle est entraîné de manière itérative jusqu'à ce qu'il soit jugé assez performant, et prêt à être déployé.
- Il entre ensuite dans une phase d'inférence : des capteurs présents dans le véhicule traitent localement de nouvelles données collectées en temps réel, puis renvoyées vers le Cloud pour améliorer le modèle, qui sera ensuite re-déployé dans les autres véhicules.
- En somme, **l'intérêt de l'inférence est d'identifier des données utiles qui n'ont pas participé initialement à la formation du modèle.** L'inférence va jouer un rôle de plus en plus important avec la **démocratisation des objets connectés** (*Internet of Things* ou IoT). La valeur de l'IA se situera alors dans sa capacité à s'adapter en temps réel, bien au-delà de sa modélisation initiale.

Comment se construit une IA ?



© Denis Lefebvre

OÙ EST LA CONSOMMATION D'ÉNERGIE ?

Le processus d'apprentissage automatique requiert **deux postes de coûts énergétiques principaux : l'hébergement des données et la puissance de calcul**, sachant que c'est derrière l'empreinte carbone des **appareils et équipements qui les supportent** qui se révèle **la plus problématique**.

Hébergement des données

- **De larges banques de données sont nécessaires pour entraîner les algorithmes** : il en faut ainsi des millions pour de la reconnaissance d'images en Deep Learning.
- Dans le domaine du NLP, ce sont **majoritairement des données « non structurées »** qui sont collectées et exploitées : il s'agit de données aux formats très variés, tels que des emails, vidéos, fichiers texte ou son, posts sur les réseaux sociaux, PDF, enregistrements d'appels clients... collectées depuis n'importe quel support (tel que le chatbot, comme nous l'évoquions précédemment). À l'inverse, des données dites « structurées » équivalent à des données qu'on peut rentrer dans un tableau Excel, avec des lignes et des colonnes clairement identifiables. Or, selon MongoDB, **80 à 90% des données aujourd'hui collectées par les entreprises seraient des données non structurées** ⁶³, et Veritas Technologies estime que **52% de ces données seraient même des « Dark Datas »** ⁶⁴. Philip Carnelley, VP adjoint du cabinet d'études IDC Europe, définit ces *Dark Datas* comme des « données non structurées qui ne sont ni utilisées ni analysées, mais simplement stockées au fur et à mesure qu'elles sont générées par l'entreprise et son écosystème ». C'est la **baisse du prix du stockage numérique qui pousse ces entreprises à collecter toutes les données possibles sans discernement**, sur leurs serveurs ou dans le Cloud, au sein de ce que l'on appelle un « **Data Lake** ». Si ces données ne sont donc pas utilisées à instant T, leur collecte reste imputable à l'IA, dans le sens où elles sont conservées pour pouvoir potentiellement être exploitées dans le modèle a posteriori. Une **logique « longue traîne »** qui pourra

⁶³ <https://www.mongodb.com/unstructured-data>

⁶⁴ <https://www.veritas.com/dark-data>

permettre de « **réveiller la donnée** » pour, par exemple, fouiller efficacement grâce à l'IA dans vos archives sur le contenu d'anciens contrats signés, d'études oubliées etc.

- De plus, **la sauvegarde des données est toujours répliquée sur plusieurs serveurs**, pour pouvoir les solliciter même lorsque l'un d'eux tombe, ou s'il est l'objet d'un incendie, comme sur l'un des Data Centers de OVH en mars 2021 ⁶⁵. Ainsi, la règle suivie par les DSI est qu'**une même donnée soit sauvegardée 3 fois, sur 2 supports différents, et avec 1 copie en dehors du site d'activité principale**. À tel point que l'Académie des Technologies alerte en octobre 2020 que « le stockage et l'archivage des mégadonnées numériques par l'approche actuelle des centres de données ne sera **pas soutenable au-delà de 2040** », envisageant plutôt de stocker cette masse de données sur le support le plus pérenne et économe qui soit : l'ADN ⁶⁶.
- Les Data Centers qui stockent toutes ces données d'apprentissage et d'inférence, en assurant une haute disponibilité 7j/7 et 24h/24, représentent un **coût de fabrication et de production important et surtout grandissant**, sollicitant face à la **surchauffe des serveurs** des systèmes de refroidissement majoritairement alimentés en électricité.

Puissance de calcul

- Plus il y a de données, et plus les besoins de calcul et donc d'énergie augmentent.
- La phase de modélisation est très énergivore, puisqu'elle repose sur une **succession d'essais et d'échecs** jusqu'à atteindre un modèle prédictif suffisamment pertinent. Pendant cette formation, **des centaines voire des milliers de version d'un modèle** sont nécessaires. **Son architecture se complexifie d'autant plus en Deep Learning**. Les combinaisons se démultiplient, et le volume de calculs nécessaires qui va avec :
 - selon la quantité de neurones et des connexions entre eux ;
 - selon le nombre et de la vitesse de chargement des paramètres.

⁶⁵ <https://theconversation.com/retour-sur-lincendie-des-serveurs-dovh-la-sobriete-numerique-est-elle-possible-157193>

⁶⁶ <https://www.zdnet.fr/actualites/l-adn-une-technologie-de-rupture-passee-au-peigne-fin-par-l-academie-des-technologies-39917323.htm>

- Le traitement local de la donnée par **inférence** devrait, pour sa part, privilégier le « *Edge Computing* » en évitant le **transfert des données en continu vers le Cloud**. Mais à l'heure actuelle, les concepteurs d'IA succombent plutôt à la tentation de renvoyer toutes les données vers le Cloud, **sans les sélectionner, afin de s'assurer une sur-optimisation de leur modèle**. La démocratisation de l'IoT interroge sur la croissance exponentielle de flux de données sans limites, vers et depuis le Cloud.
- Le problème de la puissance de calcul se situe **aussi et surtout au niveau hardware** :
 - Les entraînements sont réalisés sur des **ordinateurs spécialisés** sollicitant des **processeurs graphiques appelés GPU, très gourmands en énergie**, bien que de nouveaux processeurs plus frugaux ont depuis émergé (p.232). Créés en 1999 par Nvidia, ils remplacent les traditionnels CPU (*Central Processing Unit*) pour réaliser des calculs complexes.
 - Lorsque les algorithmes sont ensuite déployés sur des ordinateurs plus petits, comme des ordinateurs personnels ou des smartphones, **l'apprentissage en continu requiert un besoin énergétique accru et sur-sollicite la batterie**.

2. Mesurer l'empreinte carbone de l'IA

C'est une période totalement inédite dans l'Histoire de l'humanité : nous n'avons jamais eu autant de capacité de stockage et d'analyse des données, qui se démultiplient au fil des évolutions technologiques. **L'IA, qui se nourrit à la donnée, s'inscrit largement dans le Big Bang du Big Data**, dont la croissance exponentielle suscite un **débat grandissant sur la question de son empreinte écologique**. L'heure du bilan carbone de l'IA a sonné : combien pèse-t-elle dans le coût CO2 global du numérique ?

2.1. Quel est le coût CO2 de l'Intelligence Artificielle ?

UNE ÉVALUATION DU BILAN CARBONE DE L'IA

Des chercheurs de l'Université du Massachusetts à Amherst ont mis en lumière en 2019 l'impact carbone de l'IA, en mesurant la **consommation énergétique du matériel informatique utilisé pendant l'entraînement des modèles de NLP**.

Comme nous l'avons vu précédemment, ce n'est pas un hasard si c'est le traitement du langage naturel qui est à l'étude. Le NLP enregistre en effet depuis 2017 les meilleures performances dans le domaine de l'IA. Il représente **un enjeu crucial sur des champs d'application aussi décisifs que la maîtrise du langage**, au cœur de toute interaction humaine. Ainsi de nouvelles perspectives s'ouvrent pour la traduction automatique, le chatbot conversationnel, la génération automatique de textes (médias, copywriting...).

Mais il nécessite pour ses entraînements des volumes de données gigantesques. Ainsi le modèle BERT (*Bidirectional Encoder Representations from Transformers*) de Google ⁶⁷, utilisé pour son moteur de recherche, s'est formé sur la base de 3,3 milliards de mots anglais, extraits de livres et de pages Wikipédia. **On a fait relire à BERT ces 3,3 milliards**

⁶⁷ « Pre-training of Deep Bidirectional Transformers for Language Understanding », 2019 : <https://arxiv.org/pdf/1810.04805.pdf>

de mots non pas 1 mais 40 fois, en masquant certains d'entre eux jusqu'à ce qu'il arrive à les deviner. C'est ainsi que se construit la prédiction du langage.

BERT fait partie des **4 modèles du langage mis sur le banc d'essai par les chercheurs de l'Université du Massachusetts**. Y figurent aussi Transformer, un autre modèle de Google étudié avec différentes tailles de paramètres, mais aussi ELMo et GPT-2, de OpenAI ⁶⁸.

Les chercheurs ont fait tourner les modèles une seule journée, et sur un seul processeur graphique, là où les réseaux de neurones doivent normalement être entraînés pendant des jours voire des semaines. **Sur la base de ce seul cycle d'entraînement**, et en ayant connaissance du nombre de jours nécessaires pour aboutir le processus de formation complet de chaque modèle, **ils ont pu reconstituer l'énergie totale consommée pour chacun d'eux**. Un équivalent CO2 est dès lors associé à ce calcul pour évaluer le coût écologique, ainsi qu'en dernière colonne le coût économique. **Voici le résultat :**

	Date of original paper	Energy consumption (kWh)	Carbon footprint (lbs of CO2e)	Cloud compute cost (USD)
Transformer (65M parameters)	Jun, 2017	27	26	\$41-\$140
Transformer (213M parameters)	Jun, 2017	201	192	\$289-\$981
ELMo	Feb, 2018	275	262	\$433-\$1,472
BERT (110M parameters)	Oct, 2018	1,507	1,438	\$3,751-\$12,571
Transformer (213M parameters) w/ neural architecture search	Jan, 2019	656,347	626,155	\$942,973-\$3,201,722
GPT-2	Feb, 2019	-	-	\$12,902-\$43,008

Note: Because of a lack of power draw data on GPT-2's training hardware, the researchers weren't able to calculate its carbon footprint.

Table: MIT Technology Review • Source: Strubell et al. • Created with [Datawrapper](#)

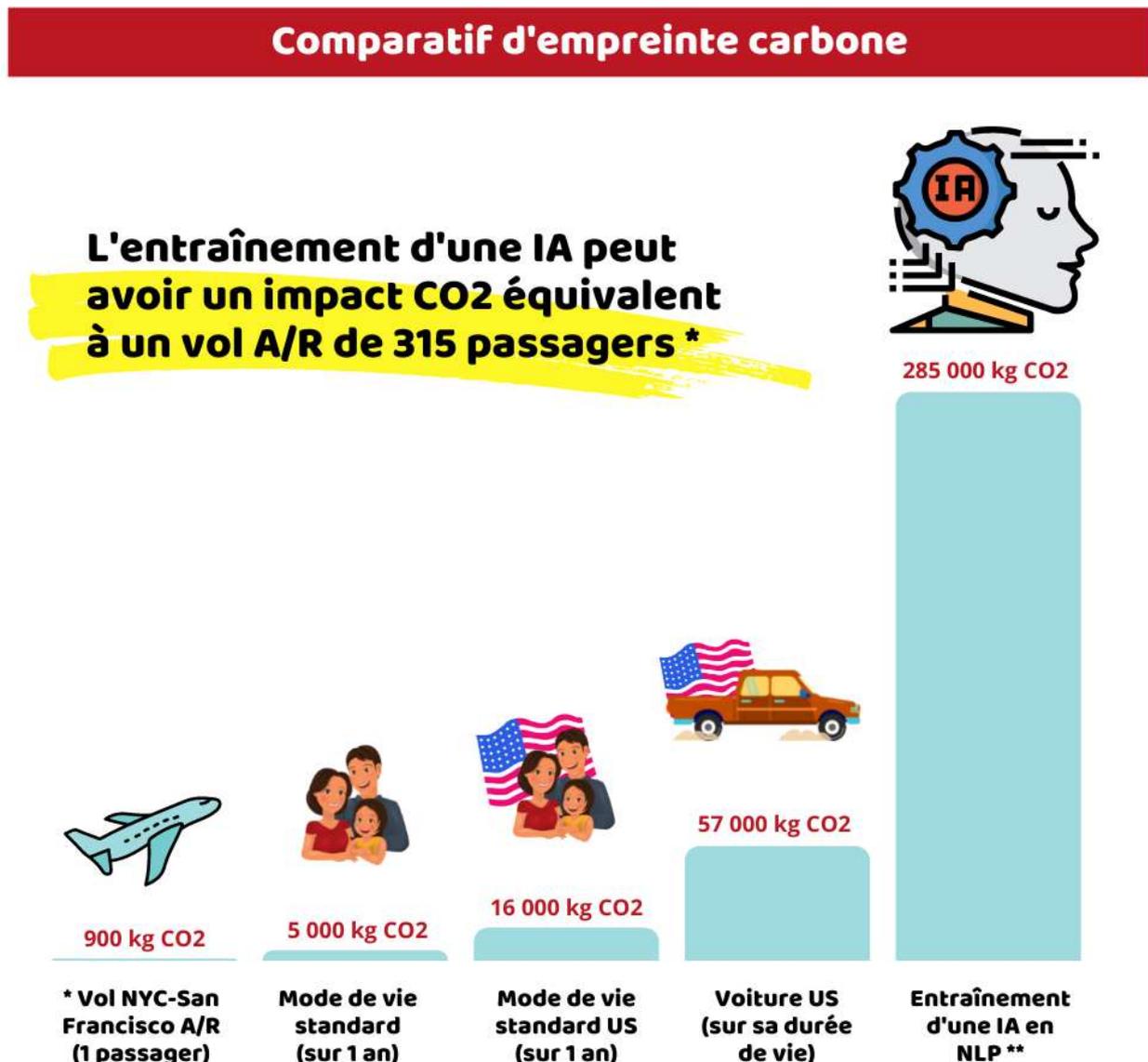
Le modèle le plus polluant se révèle aussi être le plus coûteux à financer, à savoir le **Transformer contenant 213 millions de paramètres**, associé à un processus de réglage dit « recherche d'architecture neuronale », qui vise une exhaustivité d'essais et d'erreurs

⁶⁸ « Energy and Policy Considerations for Deep Learning in NLP », 2019 : <https://arxiv.org/pdf/1906.02243.pdf>

pour tenter d'optimiser la conception de son réseau de neurones. Au final, **le gain de performance associé s'est révélé minime, pour un coût de calcul estimé entre 1 et 3,2 millions de dollars, et une empreinte carbone équivalent à :**

- 1 vol aller-retour NYC / San Francisco de 315 passagers (soit tout un Boeing 747) ;
- La consommation en CO2 d'une personne de 57 ans depuis sa naissance ;
- La durée de vie totale de 5 voitures américaines, fabrication comprise.

Les chiffres recueillis équivalent ainsi au comparatif suivant :



** Modèle d'IA Transformer (213M de paramètres) avec recherche d'architecture neurale
© Denis Lefebvre. Inspiré par le College of Information and Computer Science, Université du Massachusetts Amherst, Octobre 2019

Derrière ces chiffres quelque peu sensationnalistes se cachent plusieurs vérités, qu'il convient de mettre en lumière pour expliquer la démarche des chercheurs :

- Ce comparatif d'empreinte carbone se base sur un Transformer particulièrement gourmand en énergie, mais qui reste très peu utilisé. Or, **le second modèle le plus coûteux et bien plus courant, BERT, a un impact CO2 300 fois moins élevé.**
- L'étude se concentre sur les phases d'entraînement des modèles. Ces phases sont particulièrement coûteuses et énergivores, mais une fois le modèle abouti il sera en quelque sorte « **rentabilisé** » **par sa réutilisation en production**, non seulement pour l'entreprise qui l'a développé, mais aussi et surtout pour les développeurs d'autres organisations. L'écosystème IA est en effet **fortement ancré dans une logique open-source, où les avancées des uns servent toute l'industrie.** Le secret de l'entraînement reste bien gardé, mais les modèles eux sont bien souvent accessibles à tous, et ne nécessitent pas d'être refaits à partir de zéro. Il existe donc **un certain équilibre entre le coût des entraînements et l'utilisation des modèles qui en découlent.** Bien que le modèle open-source fasse ensuite l'objet d'ajustements pour coller au(x) besoin(s) de la nouvelle entreprise qui souhaite l'exploiter, il bénéficie en somme d'**une partie de coût énergétique mutualisé.**

Toujours est-il que le constat principal de cette étude, et ce qu'il faut en retenir, est que plus une IA est sophistiquée, plus son impact CO2 est fort. **Les coûts tant écologiques qu'économiques de la formation du modèle sont proportionnels à sa taille.**

Dans le cas du Transformer le plus polluant, ces coûts explosent du fait du raffinement de son paramétrage. C'est là le nœud du problème, sur lequel nous reviendrons dans la partie suivante. Pour l'heure, je vous propose de **mettre en perspective le coût carbone de l'IA au regard du numérique au global.**

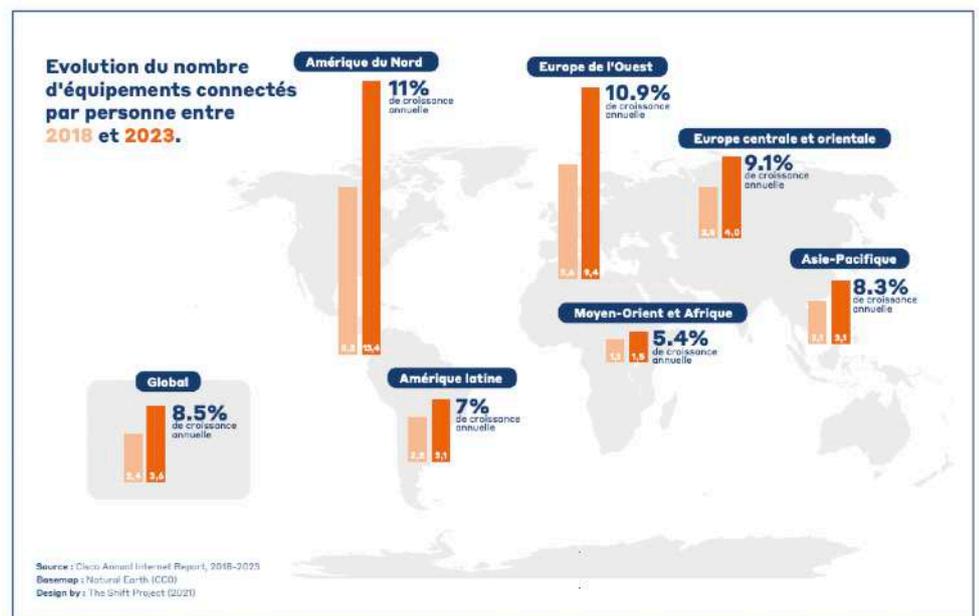
L'IA DANS LE BIG BANG DU BIG DATA

Nous avons vu où se situait la consommation d'énergie liée à l'Intelligence Artificielle, et son impact carbone spécifique aux phases d'entraînement. Mais **qu'en est-il de la part de l'IA dans l'empreinte environnementale du numérique au sens large ?**

D'après une étude de mars 2021 du Shift Project ⁶⁹, **le numérique au niveau mondial serait responsable de 3,5% des émissions de gaz à effet de serre (GES) en 2019**, avec un taux de croissance de l'ordre de **+6% chaque année** depuis 2013. GreenIT.fr l'estime même à 4%, soit 2 fois plus qu'il y a 10 ans, avec une projection à 5,5% d'ici 2025 ⁷⁰. En comparaison, le transport aérien civil était responsable de 2,5% d'émissions en 2018.

Les objectifs de « **l'Accord de Paris** », adopté en 2015 lors de la COP21 et ratifié par 175 États, visent à **maintenir le réchauffement climatique sous la barre des 2°C**, en réduisant de moitié l'exploitation des énergies fossiles d'ici 2030. Mais ce tout 1^{er} accord mondial en faveur de la préservation de l'environnement va être **mis à mal, notamment, par nos usages liés au numérique**. Plusieurs raisons à cela :

- **L'essor du marché de l'Internet des Objets (IoT)** dont le parc pourrait atteindre 55 milliards d'unités selon l'International Data Corporation, générant une croissance fulgurante du nombre de données (près de 73 zettaoctets, contre 18 en 2019) ⁷¹.

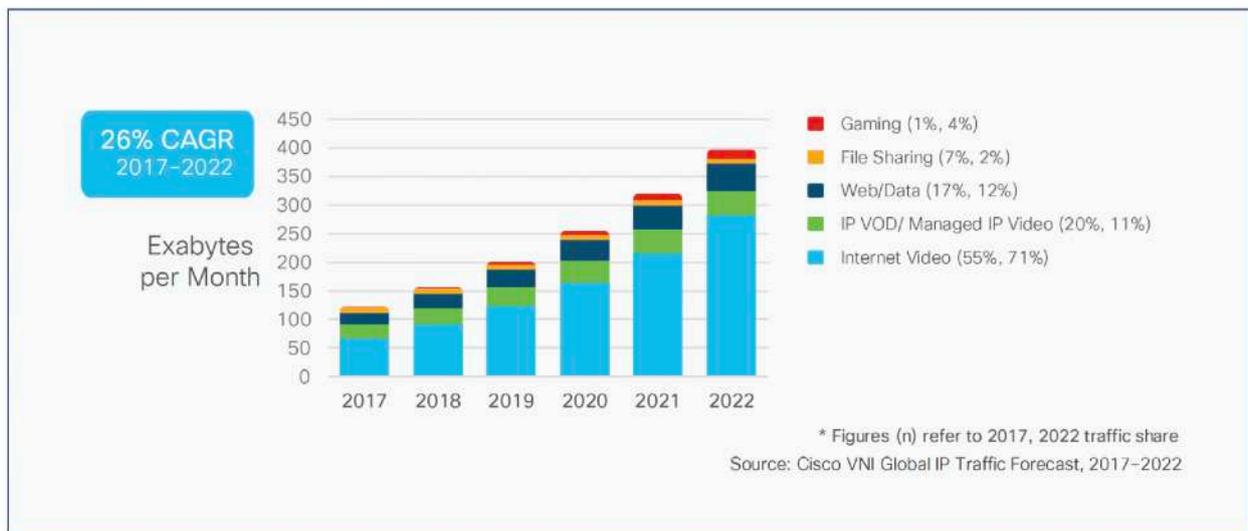


⁶⁹ https://theshiftproject.org/wp-content/uploads/2021/03/Note-danalyse_Numerique-et-5G_30-mars-2021.pdf

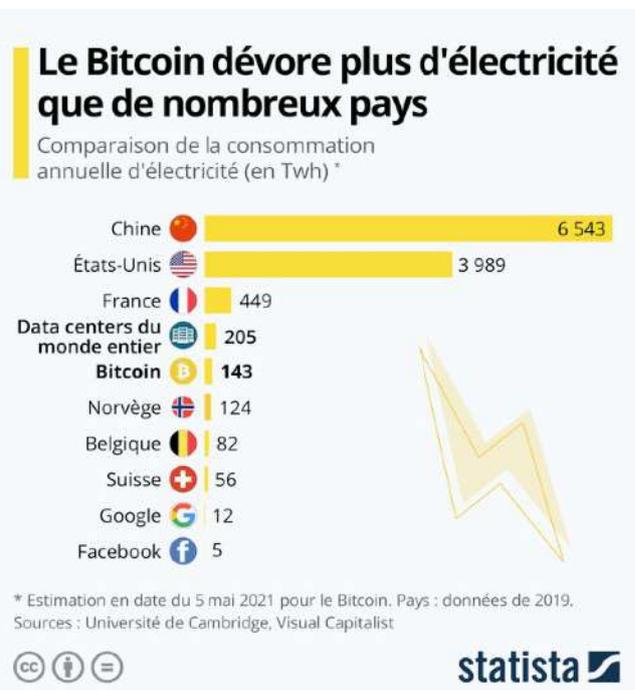
⁷⁰ <https://www.greenit.fr/2020/10/06/4-des-emissions-de-ges/>

⁷¹ <https://www.idc.com/getdoc.jsp?containerId=prAP46737220>

- **L'explosion des usages vidéos** notamment en HD (streaming, Cloud Gaming, VoD), **couplée aux réseaux mobiles de plus en plus puissants** (généralisation des forfaits 4G illimités, et à présent la 5G) font exploser les compteurs : chaque foyer américain consomme aujourd'hui en moyenne via sa box internet **4To de données par an, soit x38 depuis 10 ans**. Selon <http://archive.org>, au-delà des usages toujours croissants, le **poids des pages web** est lui aussi responsable, puisqu'il aurait été **multiplié par 4 en 10 ans** (450 Ko en 2010, 2 Go en 2020) ⁷²;



- **Le développement des usages de la Blockchain** : selon Cambridge Bitcoin Electricity Consumption Index ⁷³, le minage du Bitcoin, principale cryptomonnaie, engloutirait à elle seul **0,5% de l'électricité mondiale, soit près du tiers de la consommation électrique de la France chaque année**. Ce chiffre a quadruplé depuis 2017, avec une croissance de +40% par an. Le Bitcoin



⁷² <https://almanac.httparchive.org/en/2020/>

⁷³ <https://cbeci.org/>

est maintenant le Blockchain la plus critiquée pour la consommation d'énergie liée à son **minage, qui nécessite des calculs complexes pour assurer la sécurité de ses transactions décentralisées**. Les mineurs sont mis en compétition pour résoudre ces calculs. Des hangars remplis de milliers de racks de serveurs boostés aux GPU, aussi appelés « fermes de minage », déploient une puissance informatique considérable pour gagner les récompenses juteuses promises aux mineurs, et une **consommation d'énergie massive** comme nous pouvons le voir [dans cette vidéo](#). En juin 2021, la Chine a interdit le minage de crypto-monnaies sur son territoire ⁷⁴, certainement bien plus pour des raisons économiques que écologiques. Toujours est-il qu'il y a une volonté des acteurs de la Blockchain de passer au vert, avec la mise en place d'un Crypto Climate Accord ⁷⁵, et du passage du protocole *Proof of Work* évoqué à l'instant, vers celui de *Proof of Stake*, bien moins énergivore puisqu'il n'implique plus de mettre en jeu de la puissance de calculs. Il a été déployée en août 2021 sur la 2^e principale Blockchain, l'Ethereum, davantage axée sur les *Smart Contracts*, comme sur la prometteuse Elrond, 1^{ère} Blockchain *carbon negative* en Europe. ⁷⁶

Si le développement de nouveaux usages appelle sans cesse de nouvelles capacités de réseaux, poussant les opérateurs à se pencher d'ores et déjà sur la 6G par exemple ⁷⁷, **ces nouvelles capacités appellent à leur tour de nouveaux usages, et ainsi de suite.**

En outre, ce sont aussi les équipements supportant ces usages qui génèrent un impact environnemental conséquent. **Leur production représente près de 40% de l'empreinte carbone du numérique** en 2019. Le doublement de la taille des écrans TV, mais aussi la généralisation des smartphones fabriqués à partir de métaux rares, et dont la durée de vie courte, pour lesquels le renouvellement n'est que trop peu associé au recyclage, en sont la principale cause. Le nombre de **terminaux connectés devrait passer de 2,1 par individu en 2015, à 3,6 en moyenne en 2023.**

⁷⁴ https://www.lemonde.fr/economie/article/2021/06/19/les-mineurs-de-bitcoin-se-preparent-a-quitter-la-chine_6084808_3234.html

⁷⁵ <https://cryptoclimate.org/>

⁷⁶ <https://elrond.com/blog/elrond-carbon-negative-offsetra/>

⁷⁷ <https://www.phonandroid.com/6g-100-fois-plus-rapide-que-la-5g-la-chine-detient-la-plupart-des-brevets.html>

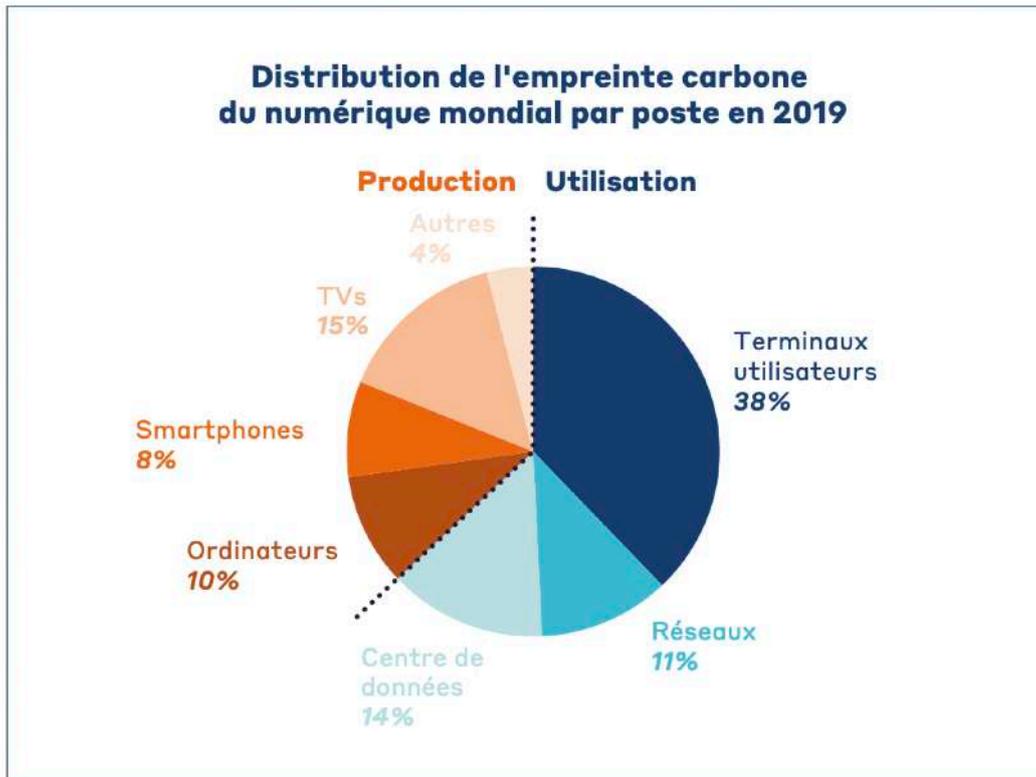


Figure 3 : Distribution de l'empreinte carbone du numérique mondial par poste en 2019 (The Shift Project – Forecast Model 2021)

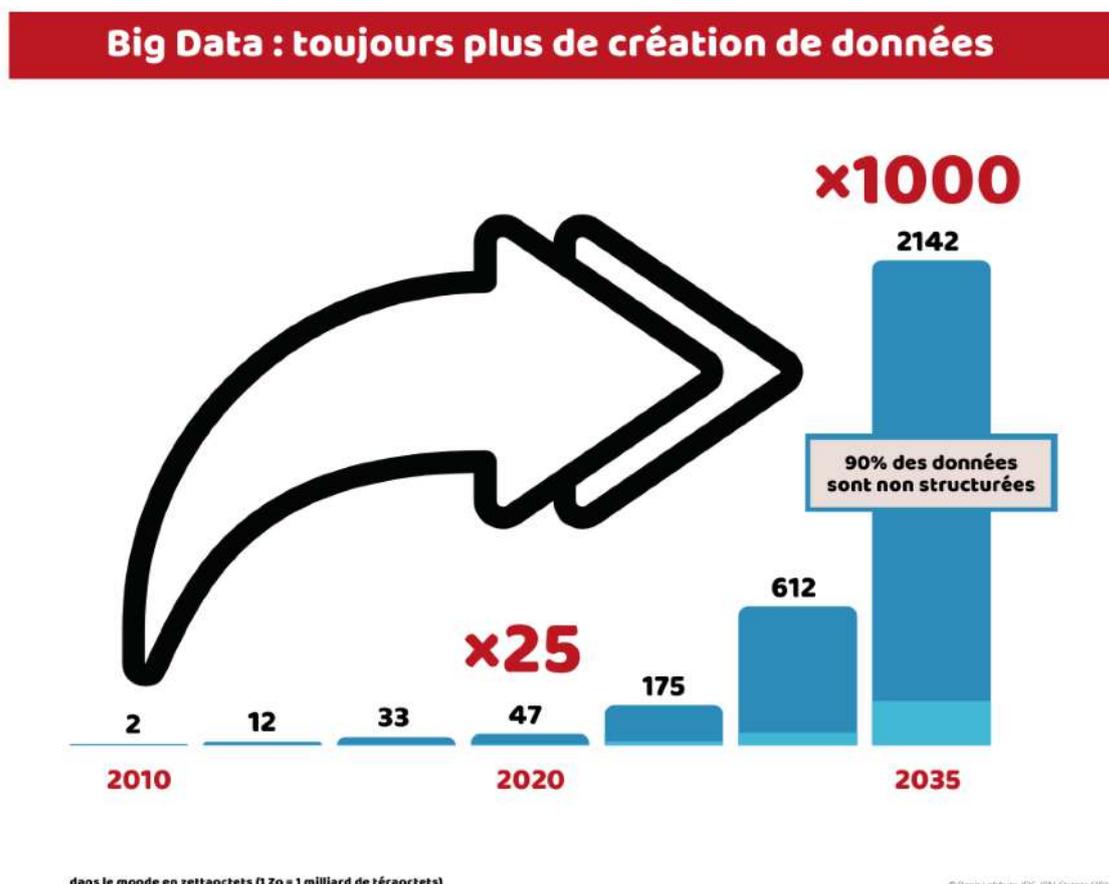
Si les **Data Centers** affichent encore une part relativement faible de l'empreinte carbone totale du numérique (14%), c'est **là où le poids environnemental risque de peser le plus dans les années à venir**, selon un rapport du Sénat, jusqu'à atteindre +86% d'ici 2040 ⁷⁸. La croissance du nombre de données stockées est d'ailleurs déjà de l'ordre de +40% par an, selon Cisco.

Face aux approches Cloud qui, bien que permettant de mutualiser les infrastructures, sont **très coûteuses sur la partie réseaux en terme de transmission des données**, une approche de *Edge Computing* basée sur le traitement en local des données, au plus près des capteurs IoT, a commencé depuis 2020 à se développer. L'explosion de la consommation des Data Centers a par ailleurs été limitée ces dernières années par la mise en place d'approvisionnements en énergies renouvelables, l'amélioration de leur rendement et de la performance de leurs systèmes de refroidissement, comme dans les pays scandinaves avec la géothermie (refroidissement naturel). Mais il faut être vigilant

⁷⁸ « Pour une transition numérique écologique », Sénat, 2020 : <https://www.senat.fr/rap/r19-555/r19-55511.html>

quant aux besoins toujours croissants de force de calculs et de stockage de données, qui ne cessent de grossir, et qui risqueraient de rendre ces gains inutiles.

L'explosion du stockage et du traitement des données, en particulier des données non structurées qui représentent 90% de leur collecte, risquent dans les années à venir de poser un sérieux problème. Selon IDC, **le volume des données mondiales devrait être multiplié par 45 entre 2020 et 2035, alors qu'il a déjà été multiplié par 25 en 10 ans.**⁷⁹



Et l'IA dans tout ça ? Dans l'interview qu'il m'a consacré (p.133), Laurent Cervoni, Directeur de la Recherche et de l'Innovation chez Talan, **tempère les chiffres alarmants** sur l'IA véhiculés dans les médias :

⁷⁹ <https://www.journaldunet.com/solutions/dsi/1424245-le-volume-de-donnees-mondial-sera-multiplie-par-45-entre-2020-et-2035-selon-statista/>



Des algorithmes aussi puissants ne sont pas mis en œuvre tous les matins. Il faut d'autant plus relativiser que **l'IA n'est qu'une partie du numérique**, qui lui-même représente environ 3,5% de la consommation énergétique mondiale.



Laurent Cervoni Directeur de la Recherche et de l'Innovation chez Talan

Il estime toutefois aussi que **l'IA peut et doit travailler à ne pas faire exploser sa facture énergétique**. Pour éviter une croissance exponentielle, les concepteurs doivent réfréner la tentation d'une précision toujours accrue de leurs modèles.

2.2. Le risque d'une croissance exponentielle ?

LA FOLLE TENTATION DU 1% SUPPLÉMENTAIRE

Plus les algorithmes s'entraînent plus ils ont de chance de devenir précis et performants. C'est la fameuse **quête du 1% supplémentaire, lorsque les niveaux de fiabilité des IA atteignent plus de 95%**.

Ce pourcentage est minime, mais il est clé pour mettre sur le marché le modèle le plus abouti. Pourtant, il est aussi **celui qui alourdit le plus la facture environnementale**.

Ainsi, 1 an après l'étude des chercheurs de l'Université du Massachusetts, OpenAI (soutenue notamment par Elon Musk, Microsoft et Amazon) dévoilait **un modèle NLP 116 fois plus gros que son prédécesseur GPT2**, contenant pourtant déjà 1,5 milliards de paramètres. **GPT-3, doté de 175 milliards de paramètres**, a nécessité pour s'entraîner l'équivalent de 355 années de calcul sur un seul processeur. Il a été programmé **sur l'un des 10 ordinateurs les plus puissants au monde**, appartenant à Microsoft, et pourvu de 10 000 cartes graphiques. Ce qui représente une **empreinte CO2 de l'ordre de 700 000**

kms en voiture, pour un coût exorbitant estimé à 10 millions de dollars... soit **3 fois plus que le modèle le plus coûteux de 2019** !⁸⁰

Des étudiants en informatique de l'Université de Copenhague, au Danemark, ont conçu un **logiciel nommé « Carbontracker »** qui permet d'**estimer la consommation d'énergie et les émissions de CO2 liées à la création des algorithmes de Deep Learning**.

Leur objectif est d'alerter sur la **course aux calculs dans laquelle les géants de la tech se sont engagés**, et le risque environnemental que cela engendre. Ainsi, ils pointent du doigt que les algorithmes doivent résoudre des équations de plus en plus complexes :



En 6 ans, de 2012 à 2018, le nombre de calculs nécessaires au Deep Learning a augmenté de 300.000% [...]

L'intelligence artificielle semble susceptible de se transformer en un problème de taille pour le climat si les tendances de l'industrie se poursuivent.



Raghavendra Selvan Co-créateur de Carbontracker, Université de Copenhague

Les chiffres continuent de leur donner raison aujourd'hui encore sur le **risque d'une croissance exponentielle dans le domaine du NLP notamment**, comme on peut le constater sur le graphique ci-contre (non exhaustif), avec par exemple :

- **BERT (Google)** entraîné sur **3 milliards de mots**.
- Puis **GPT-2 (OpenAI)** sur-enchérit en s'entraînant sur **40 milliards de mots** pour le surpasser.
- En 2019, **GPT-3 (OpenAI)** explose les compteurs à **500 milliards de mots**. Capable de générer des textes automatiquement avec un niveau de rendu correct, il suscite le débat et fait l'objet d'un véritable engouement médiatique. Au point que la MIT Technology Review en fait un **jalon de l'histoire de l'informatique**, en le comparant

⁸⁰ <https://www.ladepeche.fr/2020/11/05/avez-vous-une-idee-de-lempreinte-carbone-qui-se-cache-derriere-un-algorithme-de-deep-learning-9184196.php>

NLP : qui aura le plus gros modèle ?

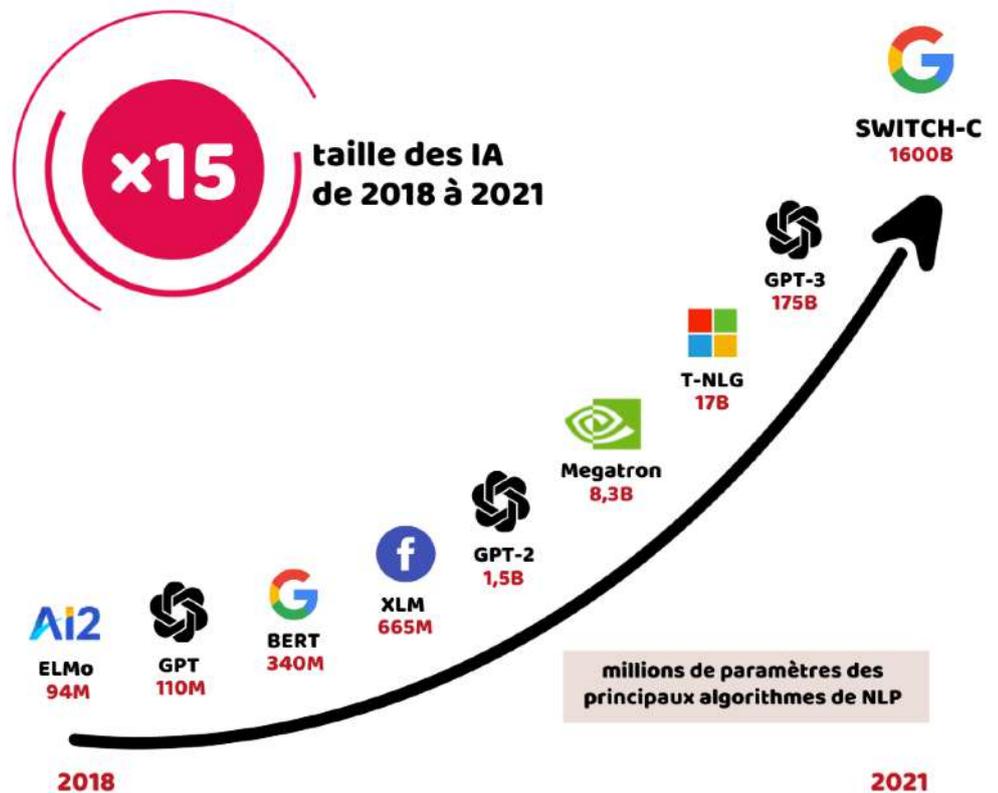
au super-ordinateur de IBM, *Deep Blue*, qui avait battu aux échecs Gary Kasparov en 1997, ou encore à *AlphaGo*, l'IA de Deepmind (Google) qui avait battu le champion de Go Lee Sedol. ⁸¹

- Pourtant, en janvier 2021, le lancement d'un nouveau modèle par Google rebat les cartes :

« **Switch-C** » affiche un nombre astronomique de 1,6 mille milliards de paramètres, soit **10 fois plus que GPT-3**, en traitant 4,5 milliards de mots par jour.

- Google annonce également début 2021 travailler sur un nouvel algorithme de NLP qui surpasserait BERT dans le traitement de documents longs : **SMITH**, qui n'est pas encore déployé, serait capable de comprendre le contexte d'un texte au-delà de la détection de ses mots-clés. Or, **comprendre un paragraphe ou un document entier en lieu et place d'une phrase unique implique de multiplier de façon exponentielle le nombre de connexions possibles entre les mots, pour en saisir le sens global**, bien qu'un traitement par lots soit envisagé.

Toujours est-il qu'il faut **mettre au regard de l'augmentation sidérante du nombre de paramètres l'accélération de la vitesse de calcul, qui permet de compenser le temps**



© Denis Lefebvre, inspiré par l'article de Microsoft "Turing NLG : a 17 billion parameter model by Microsoft" (février 2020) <https://www.microsoft.com/en-us/ai/turing-nlg-a-17-billion-parameter-language-model-by-microsoft>

⁸¹ <https://mobile.twitter.com/techreview/status/1367248219877515266>

d'entraînement et son coût associé. Ainsi Switch-C est **4 fois plus rapide** que le plus grand modèle de langage jusqu'alors développé par Google, le « T5-XXL ».

L'optimisation algorithmique de Switch-C notamment permet de simplifier le modèle pour **gérer davantage de paramètres à puissance de calcul égale.** La quête d'une utilisation optimale des ressources est classique dans l'informatique, à la fois pour en diminuer le coût, mais aussi et surtout pour aller plus vite. L'effort environnemental ne fait partie que d'un engagement plus récent lié à la dynamique du Green IT (p.205).

Toujours est-il que **cette sur-enchère n'apporte pas forcément un gain de performance à la hauteur de l'investissement.** Ainsi, une récente étude de cas sur un algorithme de reconnaissance d'images a démontré un **gain de seulement 3% pour un entraînement utilisant 2 fois plus de paramètres.** Pour les développeurs de modèles d'IA, le plus dur est de **savoir où s'arrêter, et pourquoi le faire** tant que les résultats offrent un avantage concurrentiel, à l'échelle des entreprises... comme des nations qu'elles représentent.

UNE BATAILLE ÉCONOMIQUE, ÉTHIQUE ET GÉO-POLITIQUE

La division IA de Google a été l'objet, depuis la fin de l'année 2020, d'une vague de licenciements et de démissions successifs.

Tout a commencé en Novembre, lorsque **Timnit Gebru**, débauché quelques temps plus tôt de l'Université de Stanford pour apporter plus d'éthique dans l'IA de Google, soumet un projet d'article pour l'*ACM Conference on Fairness, Accountability, and Transparency*.

Elle y **alertait sur les risques liés aux larges modèles de traitement du langage**, incluant le risque de la **sur-consommation énergétique**, mais aussi les **problèmes de biais.** En effet, en s'entraînant sur les données des langues des pays les plus riches, les modèles linguistiques s'imprègnent de leurs normes et cultures, en occultant les populations et les pays qui ont un accès restreint à Internet. Il en résulte un risque d'homogénéisation,

et de possibles **déviations linguistiques (biais sexistes, racistes...)**. Comme le souligne également la chercheuse française Aurélie Jean :



*La construction de modèles toujours plus complexes et abstraits rend ces **algorithmes moins explicables**. Ce qui peut poser un problème dans la compréhension de leur logique, de leurs réponses, mais aussi dans la détection de **possibles discriminations algorithmiques**. [...] **Relativiser cette course aux calculs toujours plus gros est alors un enjeu éthique.***

L'opacité grandit parfois aussi inutilement en augmentant, de manière somme toute relative, la précision de la réponse.



Aurélie Jean Docteur en sciences, entrepreneuse et ancienne chercheuse au MIT

Cette **opacité des algorithmes**, qui reste très certainement le secret le mieux gardé des géants de la tech, est sans doute ce qui a coûté sa place à Timnit Gebru. En Décembre, elle est poussée au départ après que son papier ait été rejeté sous prétexte de ne pas répondre aux « critères de publication ». 1 mois plus tard, Google annonçait en fanfare le lancement de Switch-C avec son trillion de paramètres...

En Février 2021, c'est au tour de **Margaret Mitchell**, la co-leader de la division éthique de l'IA, d'être licenciée. Elle voulait défendre Timnit Gebru en remontant le fil de ses échanges grâce à un script. Selon Google, en exfiltrant des documents confidentiels, elle aurait enfreint la politique de sécurité de l'entreprise. En Avril, le dirigeant et co-fondateur de Google Brain lui-même, **Samy Bengio**, démissionne...

En réaction aux avancées et positions respectives de Google et de OpenAI, s'est initié en Avril 2021 un **projet européen visant à aboutir un modèle de NLP qui serait à la fois collaboratif, multilingue, et moins biaisé que ses prédécesseurs**. « Big Science » est un projet inédit qui regroupe plus de 250 chercheurs français (CNRS, Inria, OVH, Airbus, Systran, Ubisoft...) et étrangers, et qui sera entraîné sur le super-calculateur Jean Zay à raison de 5 millions d'heures de calculs.⁸²

⁸² <https://itnews.com/articles/189893/big-science-le-projet-dia-multilingue-ethique-developpee-sur-le-supercalculateur-francais-jean-zay.html>

Il s'agit en effet de **ne pas laisser les géants américains comme chinois s'emparer seuls des IA qui seront les références de demain.**

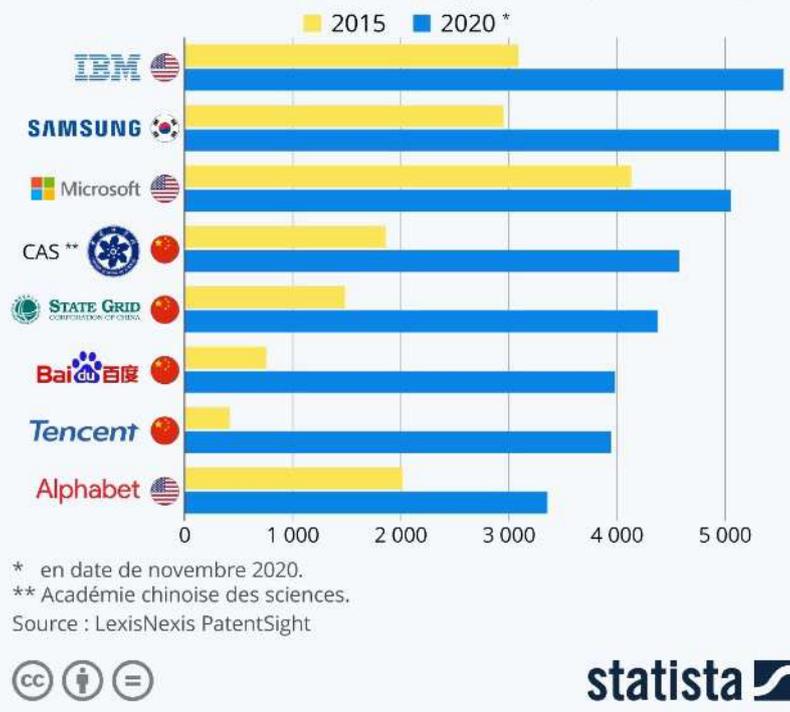
Les coûts croissants d'entraînement des modèles les rendent déjà **inaccessibles à des organisations plus modestes,** posant selon la journaliste Karen Hao un **vrai problème de droit à la concurrence et à l'innovation.**

Ce n'est pas un hasard si les universitaires, pour leurs travaux de recherche, s'associent aux GAFAMI (IBM inclus) comme aux

BATX chinois (Baidu, Alibaba, Tencent, Xiaomi) : ils n'ont tout simplement **pas accès aux ressources de calcul suffisants, réservées aux plus riches.**

La course à l'innovation dans l'intelligence artificielle

Entreprises et organisations détenant le plus de familles de brevets actifs liées à l'IA et à l'apprentissage automatique



Pour ne pas laisser les géants de la tech s'accaparer le choix des modèles développés, les types de données exploitées et l'utilisation qui en est faite, il convient de **ne pas se laisser dominer en investissant davantage dans le développement d'algorithmes plus frugaux d'un point de vue énergétique, qui serait aussi bien moins coûteux,** si l'on suit les conclusions des chercheurs de l'Université du Massachusetts et de Timnit Gebru :



*Il est plus que temps pour les chercheurs de **donner la priorité à l'efficacité énergétique et aux coûts afin de réduire l'impact négatif sur l'environnement** et l'accès inéquitable aux ressources.*



Au-delà de la **bataille éthique et économique**, c'est aussi une **bataille géo-politique** qui est en jeu. Le *Center for Data Innovation* a dévoilé un rapport sur les **capacités d'IA des États-Unis, de la Chine et de l'Union Européenne**. Depuis 2019, la Chine talonne ou devance les États-Unis dans plus de la moitié des 30 paramètres évalués. La qualité de sa Recherche sur l'IA, allié à sa suprématie en terme d'équipement en super-calculateurs (deux fois plus que les États-Unis), devrait bientôt en faire le pays le plus engagé sur l'IA.

En somme, le bilan carbone de l'IA restera maîtrisé tant que son déploiement à grande échelle ne s'accompagnera pas d'une débauche énergétique sur des modèles toujours plus gourmands. Je vous invite, après mon interview de Laurent Cervoni, à faire un **tour d'horizon de la façon dont l'IA peut, a contrario, nous aider à préserver notre planète.**



Laurent Cervoni Directeur
Recherche et Innovation chez Talan



Quand certains chercheurs alertent sur l'empreinte carbone de l'IA, est-ce qu'ils ne cherchent pas surtout à nous sensibiliser sur les risques liés à son développement exponentiel ?

« Un article dans la presse a mis en évidence une étude de chercheurs qui ont évalué la génération de CO2 pour un type d'algorithme, et qui démontre une consommation de l'IA qui est 5 fois celle d'une voiture pendant tout son cycle de vie. Évidemment, les médias se saisissent de ce genre de comparatif car c'est marquant. Mais des algorithmes aussi puissants ne sont pas mis en œuvre tous les matins. Il faut d'autant plus relativiser que **l'IA n'est qu'une partie du numérique, qui lui-même représente environ 3,5% de la consommation énergétique mondiale**. Certains chercheurs peuvent être des lanceurs d'alertes, mais ils sont surtout conscients de l'énergie consommée dans les Data Centers, les réseaux etc. Ils ont dans leur réflexion et leur démarche le souci d'optimiser les algorithmes. On voit bien que **les algorithmes les plus puissants qui sortent maintenant le sont avec une consommation énergétique de plus en plus faible**. C'est un peu le challenge que se fixent Google, Facebook, etc. Si vous prenez la dernière puce M1 de Apple, elle est dotée d'une puissance de calculs colossale, et pourtant l'ordinateur qui en est équipé a une autonomie encore plus impressionnante. On est donc bien **dans une recherche de concilier puissance de calculs et économie d'énergie**. Et c'est logique, car l'IA cherche à simuler le cerveau humain. Or, s'il s'agit de recourir à 50 Data Centers pour l'imiter, on restera toujours beaucoup trop loin des dizaines de watts seulement que consomme notre cerveau.

Est-ce que l'IA frugale n'est pas incontournable à l'heure du marketing éco-responsable ? Et comment la mettre en place concrètement ?

Vous êtes parfois contraints par les données qui sont mises à votre disposition. **Il y a des sujets où vous n'avez pas tant de données que ça.** Vous devez alors **trouver des modèles d'intelligence artificielle qui ne se basent pas sur des gros volumes de données.** Même si vous avez des données en quantité raisonnable, comme par exemple une banque qui vous donnerait un historique de 10 ans sur des milliers de clients pour faire de l'octroi de crédits, on reste très loin des volumes de données adaptés pour des modèles tels que GPT ou Transformer. Ces grands modèles utilisent des centaines de Gigas de données pour essayer d'être le plus générique possible, ce qui n'est pas votre objectif si vous traitez un sujet spécifique.

Par ailleurs, **une entreprise de taille normale ne va pas investir des millions pour construire un modèle.** S'il y a peu de données, ça lui coûtera moins cher d'embaucher deux ou trois analystes financiers pour les analyser. Et si, comme dans le domaine de l'imagerie, on peut effectivement se retrouver avec de grands volumes de données à traiter, alors la puissance de calcul nécessaire nous fera **privilégier les plateformes mises à disposition par Microsoft (Azure), Amazon (AWS), etc.** Car cette puissance de calcul, Talan ne l'a pas. Et nos clients non plus. Si on doit s'appuyer sur une infrastructure plus importante, notre objectif reste toutefois de **ne pas entraîner et faire tourner des machines pour le plaisir.** D'abord parce que ça a un coût, pour Talan comme pour les clients avec lesquels on travaille. Peu d'entreprises sont prêtes à dépenser des fortunes pour entraîner une IA. Donc pour limiter ce coût, ce que l'on fait c'est qu'**on pré-sélectionne en amont les données utiles.**

Au final, si on vise une IA « raisonnable », ce n'est pas que pour être le plus durable possible : c'est **d'abord pour des raisons économiques.** Il s'agit d'être pragmatiques. Ça ne sert à rien de partir sur des algorithmes de Deep Learning à un milliard ou à mille milliards de paramètres. Vous pouvez faire quelque chose qui tient avec 10 fois moins.

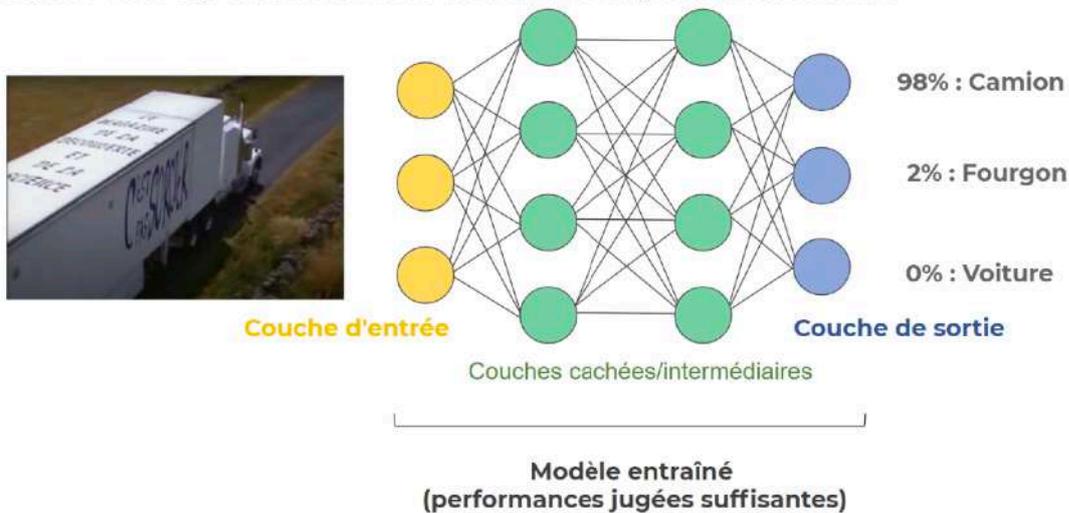
Comme en faisant du **Transfert Learning**, qui consiste à limiter l'entraînement en utilisant un modèle déjà entraîné.

Transfert Learning

Objectif : limiter l'entraînement en utilisant un modèle déjà entraîné

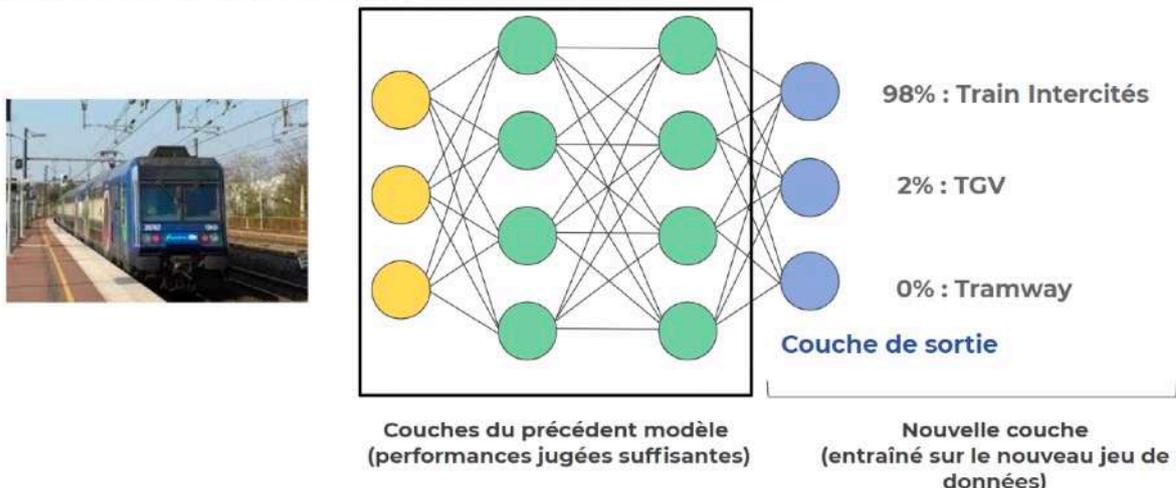
But (pour cet exemple) : Entraîner un modèle permettant la classification d'images de transport ferroviaire

À disposition : Modèle déjà entraîné servant à la classification d'images de véhicules routiers



Méthode

- Récupération des couches d'entrée & cachées du précédent modèle (pré-entraîné)
- Entraînement de la ou des couches de sortie(s) avec un "jeu" d'images ferroviaires



Approche de Transfert Learning exposée par Laurent Cervoni pour utiliser un apprentissage déjà fait, et éviter de ré-entraîner à partir de zéro. Ce modèle est intéressant à la fois d'un point de vue économique mais aussi écologique.

Ou encore en utilisant un **algorithme de régression**, qui n'a pas besoin de la puissance de calcul du Deep Learning, et qui permet de faire des choses totalement opérationnelles.

En somme, **en fonction du problème posé il faut trouver le bon algorithme d'IA**, et ne pas nécessairement chercher à assommer un moustique avec une enclume.

L'intelligence humaine est aussi de choisir le bon algorithme. Le Data Scientist ou Data Analyst doit être capable d'analyser les données et de se dire par exemple « là je n'ai pas besoin de faire un Deep Learning à dix-huit couches. ça va tenir avec un algorithme à deux couches ». C'est à l'humain de déterminer la bonne stratégie. »

Cas client



300

processus automatisables dans le cadre d'un **projet de Smart Automation**, à la croisée du Robotic Process Automation et de l'IA.

> 2 ETP

gain en temps homme sur les processus métiers ayant pu faire l'objet d'une robotisation.

L'IA au service de l'environnement

Nous avons vu qu'en se développant à grande échelle, les technologies d'IA étaient susceptibles d'avoir un impact carbone important, tant au niveau de la formation et le déploiement de ses modèles que des devices qu'elle utilise.

Pourtant, l'IA a également la capacité de se rendre utile à la cause environnementale. Elle serait même **notre meilleur allié pour répondre au défi majeur de notre temps : la sauvegarde de notre planète.**

Selon Capgemini, l'IA pourrait nous aider à réduire les émissions de gaz à effet de serre (GES) de 16% d'ici 2030, et nous permettre d'atteindre jusqu'à 45% des objectifs fixés par l'Accord de Paris.⁸³

Le bilan carbone de l'Intelligence Artificielle doit donc aussi se mesurer à l'aune de sa lutte **contre le dérèglement climatique, pour la protection de l'environnement et la préservation des espèces**, ainsi que **pour la transition énergétique.**

1. L'IA pour lutter contre le changement climatique

L'urgence climatique est au cœur de nos préoccupations. Selon le 6^e rapport du GIEC (groupe d'experts intergouvernemental sur l'évolution du climat)⁸⁴, **les phénomènes météorologiques graves vont s'intensifier**, avec des **sécheresses extrêmes jusqu'à 6 fois plus nombreuses**, et **+70% de précipitations extrêmes** si l'objectif de l'Accord de Paris de limiter le réchauffement climatique sous la barre des 2°C n'est pas tenu. Il est établi très clairement par le GIEC que les activités humaines sont responsables des dérèglements

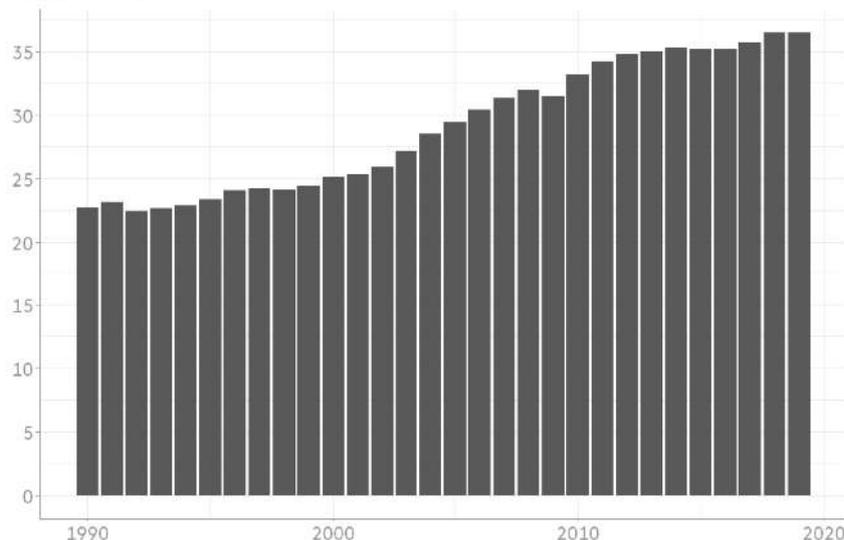
⁸³ https://www.capgemini.com/wp-content/uploads/2020/11/Climate-AI_Final.pdf

⁸⁴ <https://www.carbone4.com/rapport-giec-points-cles>

observés, et que les effets seront irréversibles si des réductions importantes d'émissions de CO2 et d'autres GES ne sont pas tenues dans les prochaines décennies.

Emissions de CO2 par année dans le monde

En gigatonnes



Source : La Banque Mondiale - Crédits : franceinfo

Pour nous aider à gagner cette bataille contre les degrés supplémentaires, l'IA va pouvoir s'appuyer sur les cas de catastrophes naturelles passés, pour **constituer un historique de données sur lequel modéliser des analyses**. Grâce à cela, nous pourrons mieux anticiper et prévenir les événements futurs.

Mais **l'IA pourra aussi s'appuyer sur les données captées en temps réel**. Ces données de « flux » sont obtenues par des outils de collecte allant de la **vision par ordinateur** (*Computer Vision*) aux **capteurs thermiques et sensorielles**, grâce au développement de l'IoT. Elles permettront d'aider à la prise de décision en s'appuyant sur la vérité terrain.

Ces outils sont à présent **à l'œuvre partout : dans les mers, les sols, les airs, et même depuis l'Espace**, pour aider la reconnaissance d'images à prendre de la hauteur. Grâce à eux, **l'IA peut s'attaquer un à un à tous les enjeux environnementaux** : la dépollution, la conservation des espèces, la biodiversité, et bien plus encore.

ImpactIA a ainsi identifié que **l'IA peut fortement aider à atteindre 9 des 17 Objectifs de Développement Durable (ODD) fixés par l'ONU** ⁸⁵ :



1.1. L'IA pour dépolluer

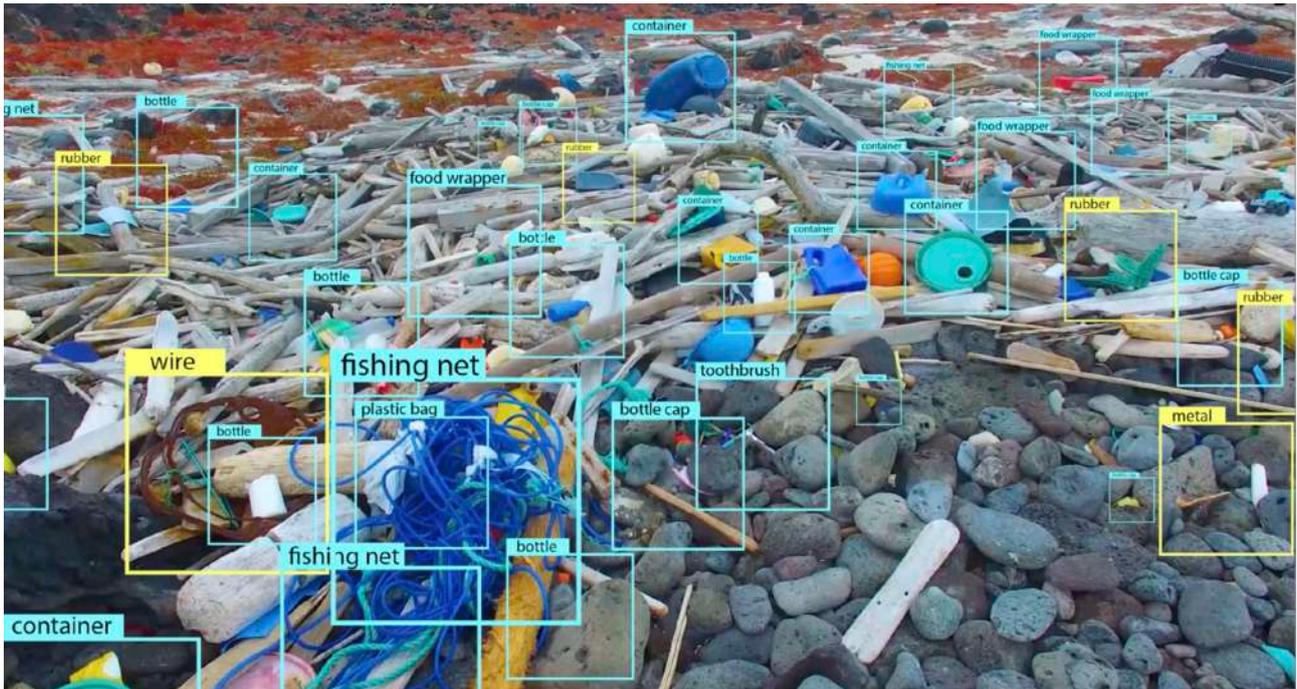
Chaque année dans le monde, ce n'est pas moins de **275 million de tonnes de déchets plastiques** qui sont créés, et qui mettront chacun près de 500 ans à se décomposer. La pollution touche aussi bien les mers que l'air et les sols. Il faut **prendre la problème à la racine et produire moins de plastique**. Mais il faut bien **aussi s'occuper des plastiques déjà présents dans nos écosystèmes**.

DÉPOLLUER LES MERS

Tous les ans, c'est près de **10 millions de tonnes de plastiques qui se déversent des rivières vers les mers et les océans**, se désagrègent en milliards de petits morceaux et menacent tout l'écosystème marin. Face à ce problème majeur, **citons quelques projets qui ont recours à l'IA pour identifier les plastiques et dépolluer les eaux**.

⁸⁵ <https://www.impact-ai.fr/fr/2021/05/28/ai-for-green-une-serie-speciale-par-impact-ai/>

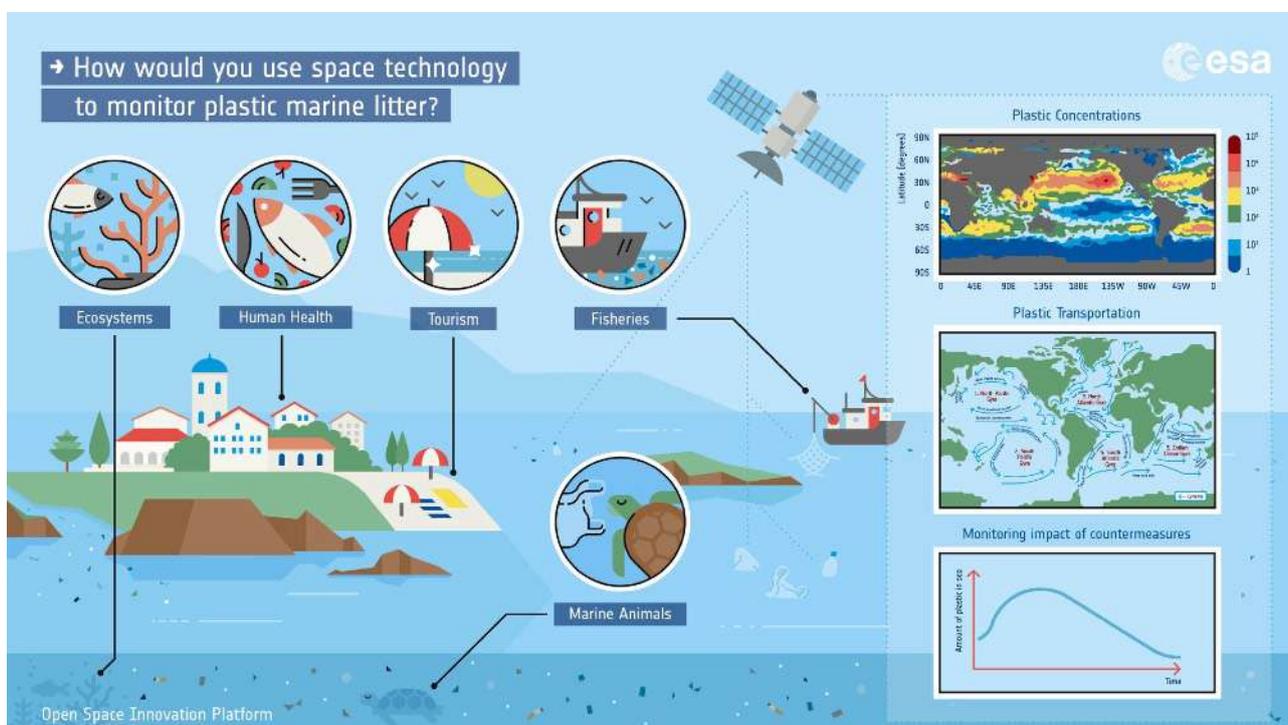
Identifier les plastiques par drones et images satellites



- **Ellipsis** (anciennement « The Plastic Tide Project ») s'appuie sur les vidéos et images récoltées à partir de drones, satellites, sous-marins ou images de vidéosurveillance sur les plages, dans les rivières et dans les océans partout dans le monde. Un **algorithme de Machine Learning** est alors utilisé par **pour identifier les déchets plastiques et les classer** selon qu'ils sont des bouteilles en plastique, des bouchons de bouteille, des filets de pêche, des brosses à dents, etc. même sur des plastiques dégradés. Le taux de précision est de 93% et permet même d'identifier la marque. Ce taux est suffisamment élevé pour **établir de façon fiable une carte thermique globale des déchets plastiques dans un environnement spécifique**, en comparant son évolution dans le temps. Ellipsis travaille alors en concertation avec les gouvernements, les entreprises et les organisations à but non lucratif afin de sensibiliser et agir.



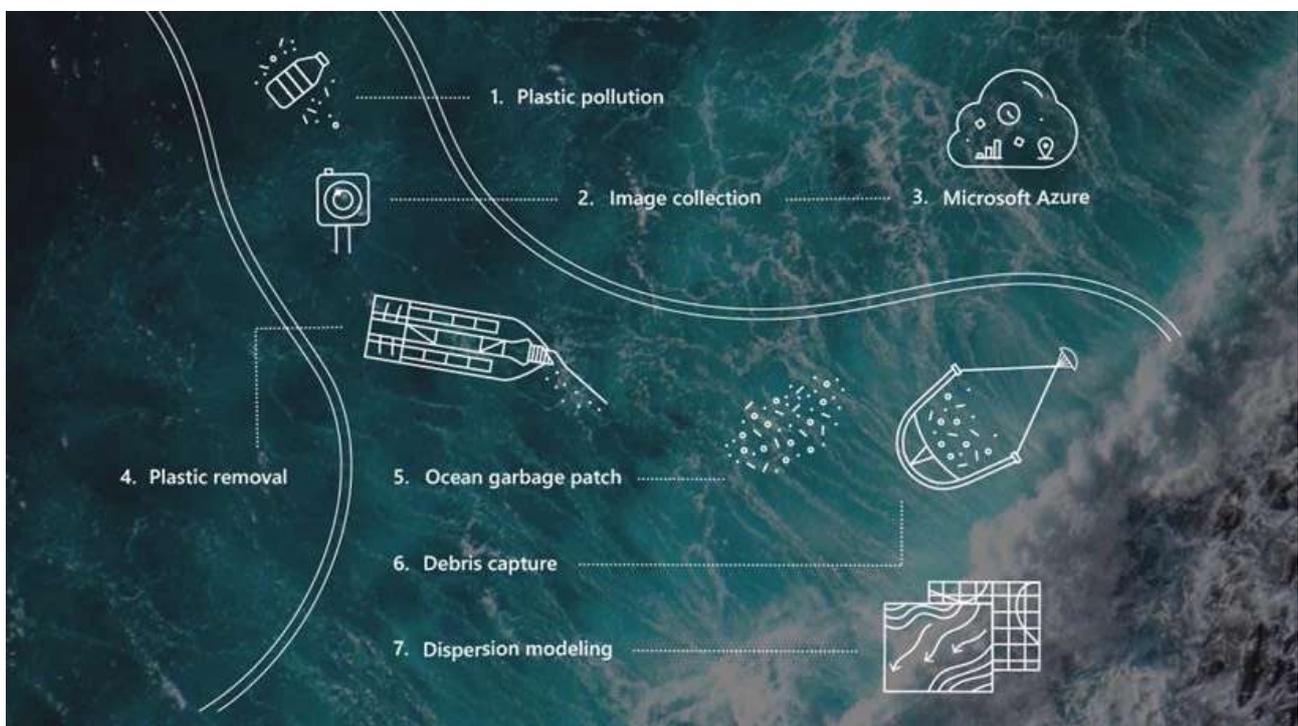
- Pour faciliter la surveillance du plastique plus particulièrement au milieu de l’océan, **L’Agence Spatiale Européenne** met à disposition des satellites en orbite autour de la Terre. Ces **satellites**, parmi lesquels Sentinel-2, sont **dotés de caméras très haute résolution**. Des équipes de chercheurs dont ceux du Plymouth Marine Laboratory (Royaume-Uni) analysent les données récoltées grâce à l’algorithme qu’ils ont développé. Cet algorithme permet non seulement d’**identifier les concentrations de plastique** – en réussissant à le distinguer des autres matières dans 86% des cas – mais aussi les sources des courants marins déversant ces plastiques. Ainsi, l’objectif est le plus possible de remonter à l’origine du problème.



Capter les plastiques avant qu’ils n’atteignent les océans

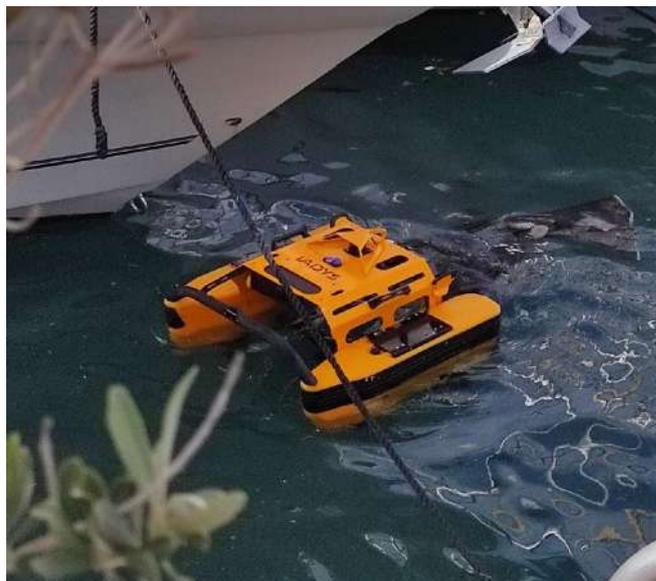
80% des pollutions plastiques présentes dans les océans proviennent d’environ 1000 cours d’eau. Pour résoudre le problème de la pollution des océans, il faut donc pouvoir **intercepter les plastiques en amont, depuis les cours d’eau et les rivières.**

- C'est l'objectif que s'est fixé **The Ocean Cleanup**, une ONG néerlandaise créée en 2013, qui avait été très médiatisée à partir de 2019 et la création de son bateau à énergie solaire « L'Interceptor ». Ce bateau autonome, équipé d'un gigantesque filet orienté selon le vent et les courants par des robots flotteurs, et conçu pour que les animaux marins ne soient pas capturés, s'attelait initialement à nettoyer les plastiques déjà présents dans les océans. En particulier ceux du « **septième continent** », **cette poubelle flottante** à mi-chemin entre Hawaï et la Californie, et qui fait 3 fois la taille de la France. L'Interceptor est ainsi capable d'**extraire 50 000 kg de plastique par jour, soit plus de 15 000 tonnes par an**. Pour « intercepter » les plastiques avant qu'ils n'atteignent les océans, The Ocean Cleanup s'est **associé en 2020 à Microsoft afin de développer un modèle de Machine Learning basé sur sa solution Azure** ⁸⁶.
L'algorithme permet d'identifier les déchets plastiques directement dans les rivières, et de les distinguer de déchets organiques comme le bois ou les feuilles, afin de ne capturer que les éléments qui doivent l'être.



⁸⁶ <https://www.microsoft.com/en-us/ai/ai-for-earth-the-ocean-cleanup>

- Un bateau robotisé d'une toute autre taille a été conçu par la société française IADYS : le **Jellyfishbot**. Ce robot électrique de 20 kg, doté d'Intelligence Artificielle, identifie et collecte les déchets (plastiques et hydrocarbures) à la surface de l'eau, **sur des zones plus petites et plus difficiles d'accès** : canaux, ports, lacs, bases de loisirs aquatiques... Le Jellyfishbot est programmé sur une zone de nettoyage déterminée par points GPS, et peut recueillir en toute autonomie pendant 6 heures jusqu'à 80 litres de macro-déchets et 30 litres d'hydrocarbures.



DÉPOLLUER LES SOLS

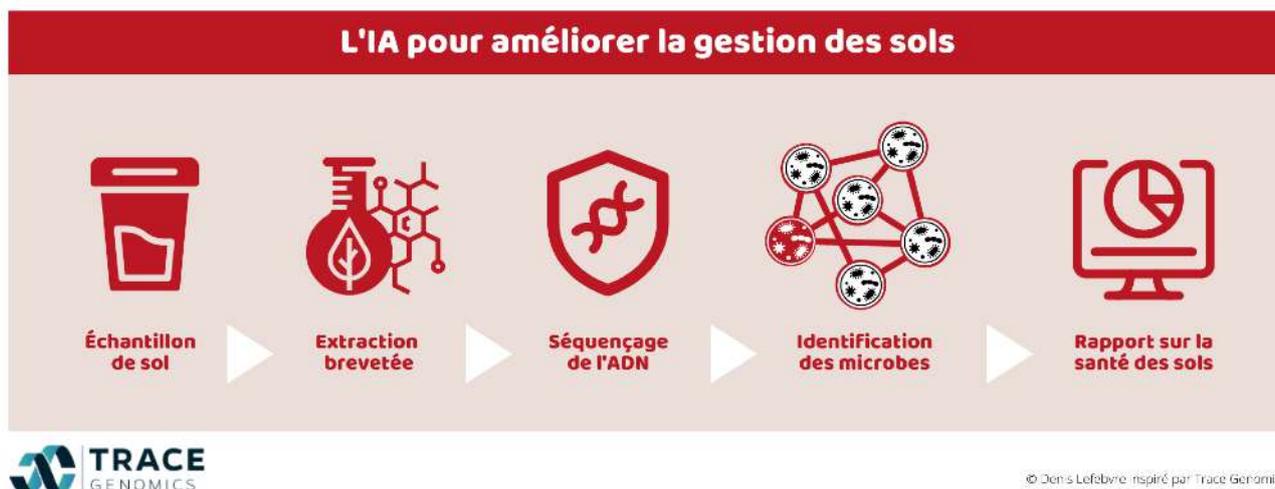
L'industrialisation, l'urbanisation, le forage, l'intensification de l'agriculture font partie des principales causes de la pollution des sols, du fait de la contamination liée par exemple aux engrais et produits chimiques, aux plastiques et aux déchets. En Chine, c'est 16% de l'ensemble des sols, et 19% des sols agricoles qui se retrouvent pollués.

Cette pollution menace la biodiversité des sols, leur capacité de filtrage et leur équilibre nutritif. **Les eaux souterraines se contaminent, et contaminent à leur tour les denrées produites, ainsi que les animaux en pâturage.** Toute la chaîne est dès lors impactée, et nuit à la sécurité alimentaire.

La folle progression de la population mondiale, qui devrait atteindre les 9 milliards d'habitants d'ici 2050 ⁸⁷, impose de **travailler à des rendements plus durables**. L'impact économique de sols plus sains, qui améliorent la productivité et la qualité des récoltes, est un argument de poids pour pousser les agriculteurs à franchir le pas.

⁸⁷ <https://www.nationalgeographic.fr/environnement/2019/06/notre-planete-abriterait-plus-de-9-milliards-de-personnes-dici-2050>

C'est ainsi que **Trace Genomics** utilise l'IA pour aider les agriculteurs à améliorer la **gestion de leurs sols**. Les algorithmes de Trace Genomics analysent et comparent les données recueillies à partir d'un échantillon pour **évaluer l'état de santé du sol**, en prenant compte de son pH, les millions de microbes et de bactéries qui s'y trouvent, etc. Les données biologiques et chimiques sont croisées avec les données d'exploitation des terres, pour **en tirer des recommandations à même d'améliorer leur rendement**.



DÉPOLLUER L'AIR

Selon une étude réalisée par des chercheurs allemands ⁸⁸, **la pollution atmosphérique serait responsable de 800 000 morts tous les ans en Europe, et 8,8 millions à l'échelle mondiale**. Elle raccourcirait l'espérance de vie de 3 ans principalement en causant des maladies cardiovasculaires. **La pollution de l'air tue davantage que le tabac**. C'est une menace invisible à laquelle nous sommes tous soumis. Si la pollution naturelle (feux de forêt, éruptions volcaniques...) n'est pas étrangère au problème, deux tiers de cette mortalité atmosphérique reste liée à la pollution humaine.

Or, comme nous allons le voir, **l'IA peut infléchir autant sur la pollution naturelle que sur la pollution liée à nos émissions de combustibles**.

⁸⁸ « Loss of life expectancy from air pollution compared to other risk factors », 2020 : dx.doi.org/10.1093/cvr/cvaa025

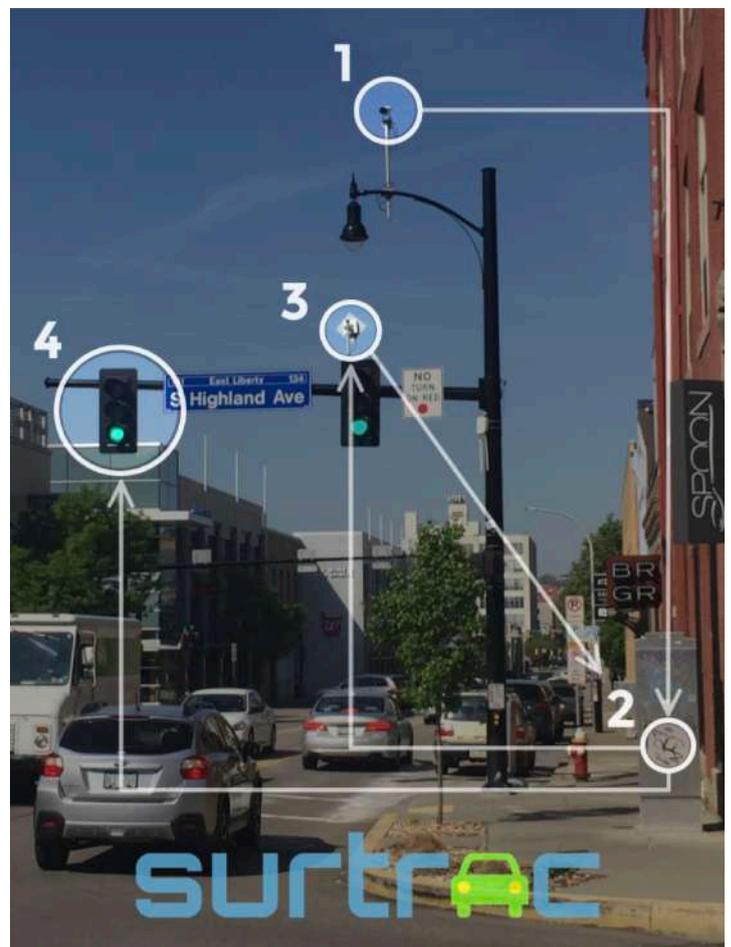
Fluidifier le trafic pour réduire les émissions de GES

14% des émissions de CO2 dans le monde sont imputables au secteur des transports. Sur la route, l'IA est déjà à l'œuvre pour alimenter les véhicules autonomes, et à travers la technologie V2X présente par exemple sur la Golf 8 de Volkswagen, qui lui permet de communiquer avec son environnement (infrastructures, véhicules, piétons, réseaux). Mais l'IA permet aussi de **limiter nos émissions en réduisant les embouteillages et la consommation de carburant qui en découle.**

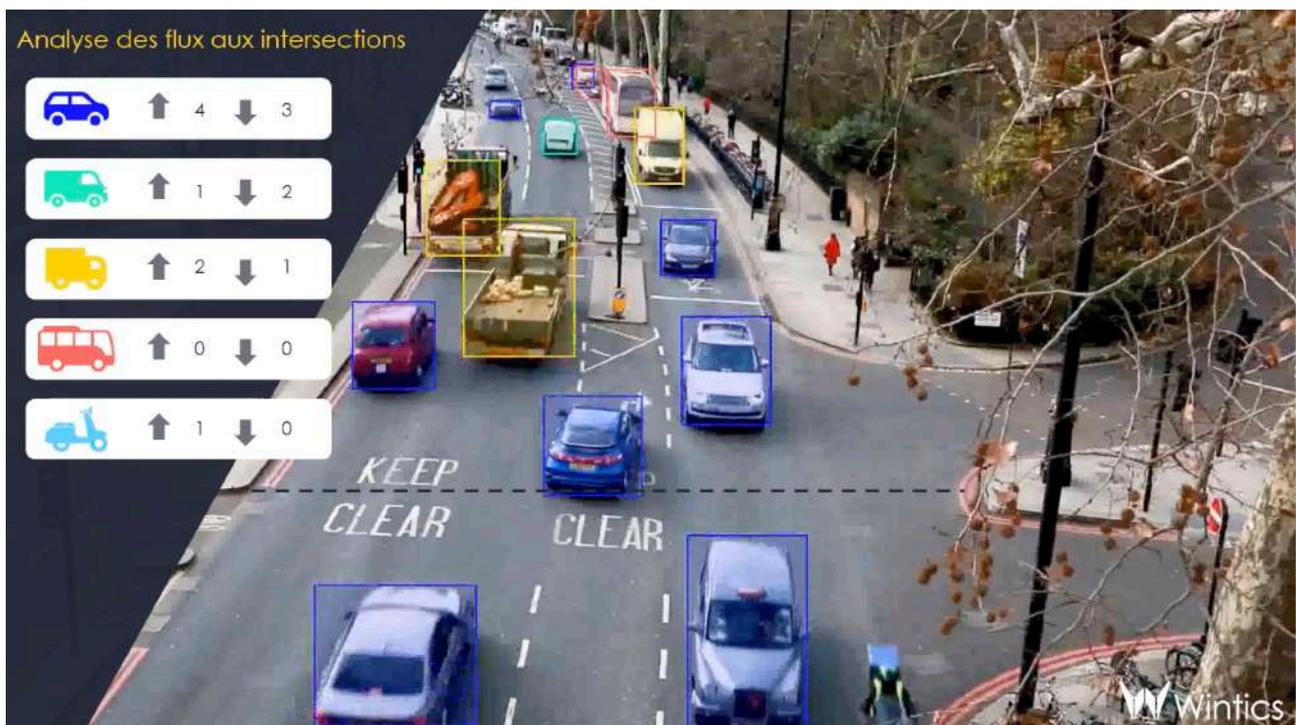
- Les feux intelligents **Surtrac de Rapid Flow** permettent de réduire les émissions de CO2 de 20% en fluidifiant le trafic routier urbain, en réduisant le nombre d'arrêts, de freinages et d'accélération, et le temps d'attente aux intersections (jusqu'à -40%). **Alliant robotique et intelligence artificielle**, cette solution œuvre activement aux enjeux de la Smart City.

Alors, comment ça marche ?

1. **Détection de l'environnement** grâce à des radars et caméras ;
2. **Traitement de l'information en temps réel grâce à des algorithmes** à même de générer un plan d'optimisation adapté du trafic ;
3. Envoi des commandes pour **coordonner l'ensemble des autres feux de circulation** ;
4. Envoi des commandes sur le feu de circulation **présent à l'intersection.**



- La startup française **Wintics** s'appuie sur le **Deep Learning pour analyser, grâce à la reconnaissance d'images, les vidéos de trafic**. Elle peut ainsi classifier les véhicules, analyser les flux aux intersections etc. pour **aider les villes à optimiser leur mobilité urbaine**, et mieux gérer et faciliter l'explosion de la mobilité douce (vélo, trottinette électrique...). Wintics veille d'ailleurs à l'impact environnemental du traitement de ses données, **en ne renvoyant pas les flux vidéos captés vers un serveur centralisé**, mais au contraire en plaçant son logiciel directement depuis le boîtier sur place. Ce traitement **en Edge Computing** évoqué précédemment, traitant les données près de la source de captation, limite considérablement la consommation requise (environ 10 watts pour le boîtier, et seulement quelques kilooctets pour ne renvoyer vers le serveur Wintics que les statistiques de comptage).

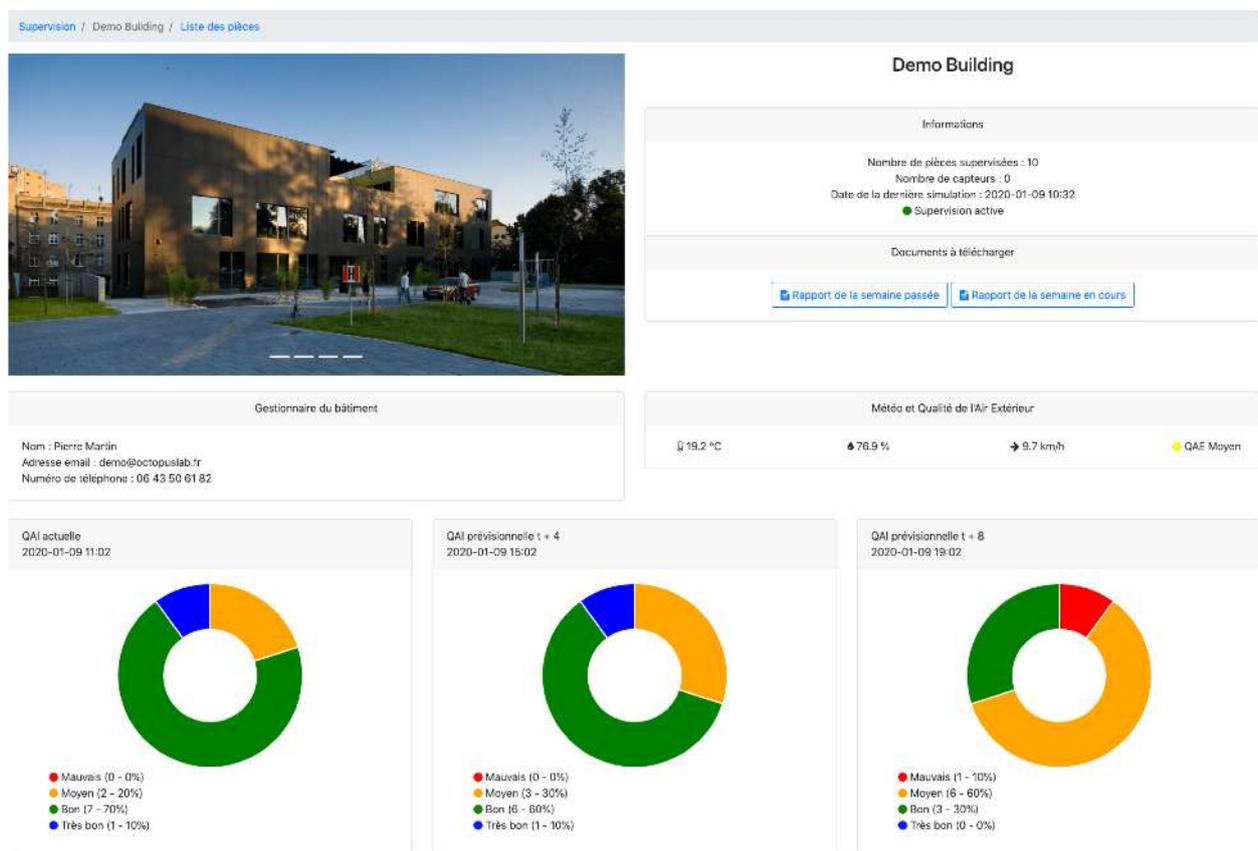


- Le projet intrapreneurial **Cyclope.ai** de Vinci élargit ce champ d'action aux péages et aux autoroutes. Ainsi sa solution « Tunnelwatch » permet, grâce à des algorithmes de Deep Learning, d'améliorer la qualité des **informations remontées en temps réel depuis les tunnels** : trafic, détection d'incidents avec un taux de réussite de 98%, etc.

Améliorer la qualité de l'air intérieur

L'air intérieur est **5 à 8 fois plus pollué que l'air extérieur**. La crise de la Covid-19 a mis sur le devant de la scène la nécessité de travailler à la ventilation de nos intérieurs.

Détecter le risque de propagation du coronavirus dans les bâtiments d'entreprise est un argument que la société **Octopus Lab** a bien sûr mis en avant pour promouvoir INDALO Supervision, sa solution de conception et de **maintenance prédictive visant à améliorer la qualité de l'air**. Les algorithmes de INDALO Supervision ont appris à **anticiper les pics de CO2** en croisant toutes les données liées aux matériaux présents dans le bâtiment, le système de ventilation, la densité d'occupation des pièces et la pollution extérieure. L'air peut ainsi être modulé en conséquence par un **pilotage intelligent de la ventilation**.



1.2. L'IA pour préserver la biodiversité

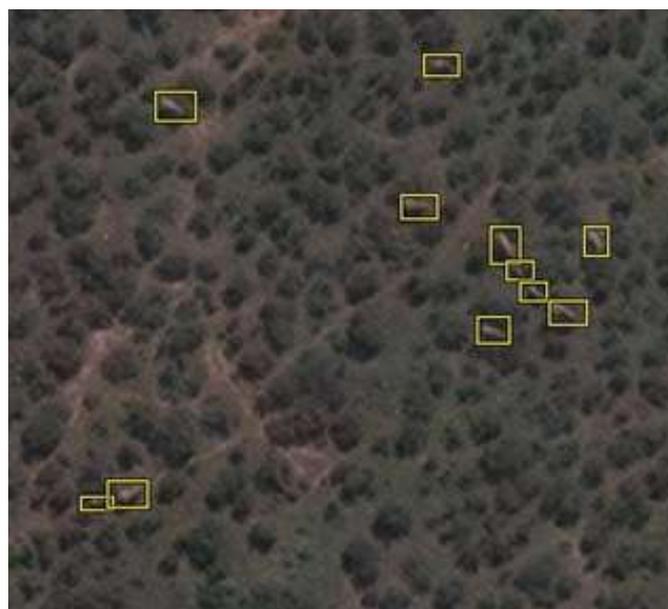
Selon le Rapport Planète Vivante 2020 de la WWF ⁸⁹, **entre 1970 et 2016 c'est 68% de la faune sauvage qui a purement et simplement disparu**. Ce pourcentage a été établi sur la base de données collectées sur 21 000 populations d'oiseaux, de poissons, de mammifères, d'amphibiens et de reptiles partout dans le monde. Ce sont dans les zones tropicales que le **déclin de la biodiversité** est le plus important, même s'il touche tous les continents. Les causes sont nombreuses : changement climatique, pollution, espèces invasives, malaises, chasse, braconnage, dégradation des habitats naturels...

L'IA peut aider là encore à **cartographier les espèces et leur environnement**, pour à la fois **mieux les comptabiliser**, mais aussi **comprendre ce qui les menace pour agir**.

PRÉSERVER LES ESPÈCES MENACÉES

Comprendre les écosystèmes pour mieux les protéger

- La collecte de données sur le terrain se fait plus facile, plus sûre et plus précise grâce au déploiement d'un large éventail de solutions : capteurs, drones, radars... Les résultats sont largement meilleurs que par la détection humaine. Par exemple, les satellites permettent d'obtenir des images de haute définition, et peuvent couvrir des territoires de plus de 5000 km² en quelques minutes, sans perturber à aucun moment les espèces. À titre d'exemple, **les satellites WorldView-3 et 4 détectent et comptabilisent automatiquement les éléphants en Afrique**, grâce à un



⁸⁹ https://www.wwf.fr/sites/default/files/doc-2020-11/20200910_Rapport_Planete_Vivante_WWF.pdf

algorithme développé par une chercheuse de l'Université de Bath (Royaume-Uni). La position et les **schémas de migration** sont suivis plus efficacement, et permettent d'**aider les pouvoirs publics à protéger cette espèce menacée**.

- Au fond des océans, citons notamment le poisson robot **Soft Robotic Fish** conçu par des chercheurs du MIT. Le Soft Robotic Fish prend l'apparence d'un petit requin, il est contrôlé à distance et permet d'**approcher les créatures marines pour les observer sans les déranger**. Sa batterie en lithium lui permet en outre d'être bien moins polluant que les habituels véhicules sous-marins. D'autres robots tels que ceux de Notilo Plus (Seasam et iBubble) peuvent être cités, de même que ceux de Subsea Tech. Ces solutions permettent de **cartographier les récifs coralliens menacés d'extinction** (projet Allen Coral Atlas), mais aussi la totalité des fonds marins en HD, comme ambitionne de le faire le Shell Ocean Discovery XPrize d'ici 2030. L'Ocean Data Alliance cherche également à mutualiser les efforts en utilisant le **Machine Learning pour développer des outils de surveillance et de prédiction** à partir de données open source.



- À la surface, **le cachalot est sous bonne protection avec l'IA de Cap Gemini et Amazon Web Services (AWS)** au service de la biologiste marine Lisa Steiner ⁹⁰. Chassé pendant des siècles et aujourd'hui perturbé dans son environnement naturel, le cachalot est officiellement une espèce menacée depuis 2018. Il n'en resterait plus qu'une centaine de milliers dans les mers. Pour les préserver, la solution est de **comprendre leur parcours de migration afin de détourner les routes maritimes**. Pour cela, Lisa Steiner photographie la forme de la queue unique à chaque cachalot, qui permet de les identifier individuellement. Grâce à un réseau de Deep Learning pré-formé basé sur les services de AWS et Cap Gemini, leur identification est largement facilitée avec un taux de précision de 97,5%, et des correspondances accrues.

fluketracker Welcome Upload Match Whales Browse Pictures About Login/Register

Do these whales match?

Select Uploaded Picture

New Image Number: 6 of 7

Whale ID: -
 Picture owner: LISASTEINER
 Coordinates / Place: -
 Date: -


 PM-WWA-MArietti-20190510-711711.jpg

Best Matching Picture Number: 1

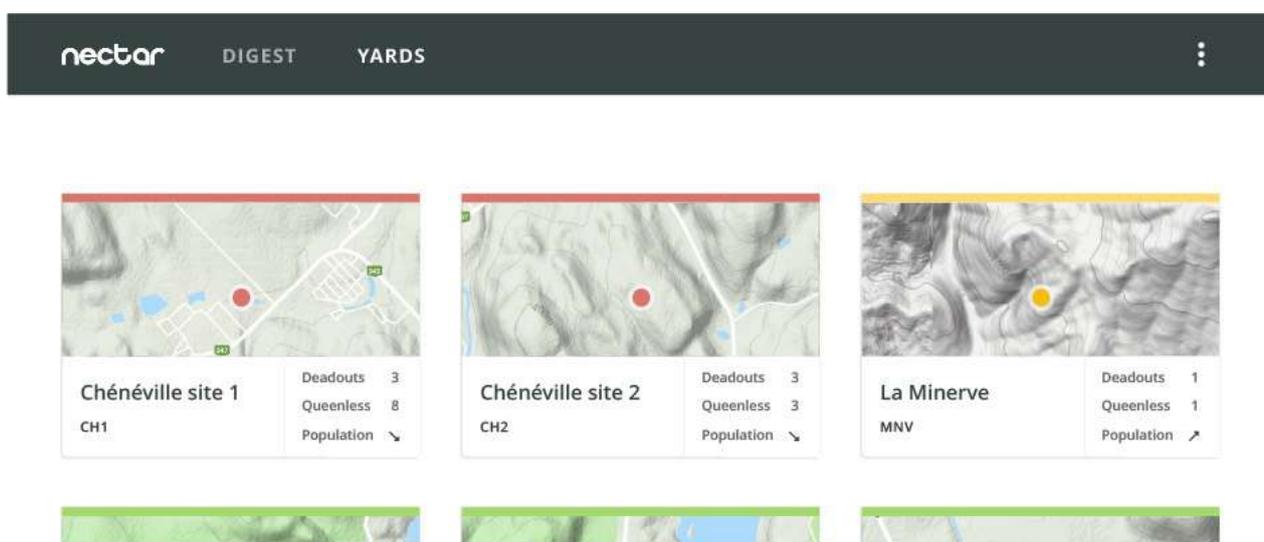
Whale ID: 6351 SHOW MORE PICTURES
 Picture owner: LISASTEINER
 Coordinates / Place: 38.3522783, -28.4250850 OPEN IN MAP
 Date: TUE, 15 SEP 2020 11:13:29 GMT
 Similarity-Score: 0.76


 PM-WWA-20200915-051crp.jpg

- **Les abeilles sont elles aussi menacées d'extinction avec jusqu'à 80% de mortalité dans certaines ruches d'Europe.** Or, 80% de nos plantes et de nos fleurs dépendent de la pollinisation des abeilles pour leur survie. C'est donc un enjeu majeur auquel

⁹⁰ <https://www.capgemini.com/fr-fr/cas-client/intelligence-artificielle-protection-cachalots/>

l'IA peut là encore répondre. Ainsi la startup **Nectar** collecte grâce à une solution IoT (capteur In-Hive) **des données directement dans les ruches**, comme la température, l'humidité, le son ou encore les déplacements. **Des algorithmes d'apprentissage automatique traitent alors ces données pour aider les apiculteurs à suivre l'état de santé de leurs ruches, et agir pour les re-dynamiser** : détection de mortalité, de la présence ou non de la reine, évolution de la population... Ainsi les colonies peuvent rester saines et productives, enrayant leur déclin.



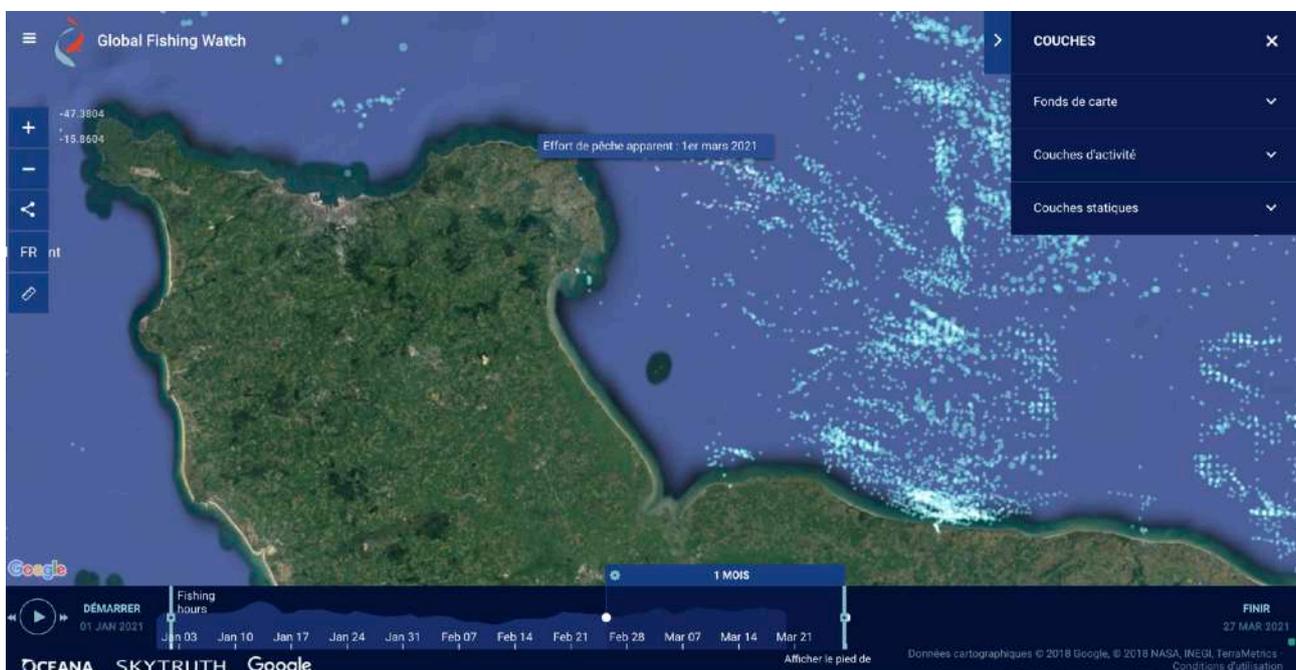
Prévenir les attaques menaçant les espèces

La chasse et le braconnage peuvent être surveillés grâce à des capteurs en temps réel, mais aussi anticipés grâce à l'optimisation des points de contrôle, et au développement d'**outils prédictifs**. Ainsi le parc national de Port-Cros a équipé ses équipes de patrouille avec la solution d'Intelligence Artificielle « EYSEEA AMP » de BSB Marine ⁹¹. EYSEEA AMP peut **alerter sur des présences de bateaux non autorisées et de comportements suspects en milieu protégé**.

⁹¹ <https://thegoodai.co/2021/02/23/le-parc-national-de-port-cros-et-bsb-marine-developpent-une-ia-pour-la-preservation-des-aies-marines-protgees/>

La **pêche illégale**, non déclarée ou non réglementée, représente à ce jour jusqu'à 26 millions de tonnes, soit un tiers du total de la pêche légale ⁹². Cette surpêche menace 55% des récifs du monde entier, ainsi que des espèces capturées par erreur telles que les requins ou encore les tortues ⁹³.

Pour suivre les mouvements et les informations d'environ 60 000 bateaux de pêche commerciale, et surveiller la surpêche illicite, **Google a développé grâce à des algorithmes d'apprentissage automatique un carte interactive de l'activité de pêche mondiale**, en quasi temps réel et accessible à tous : le Global Fishing Watch.

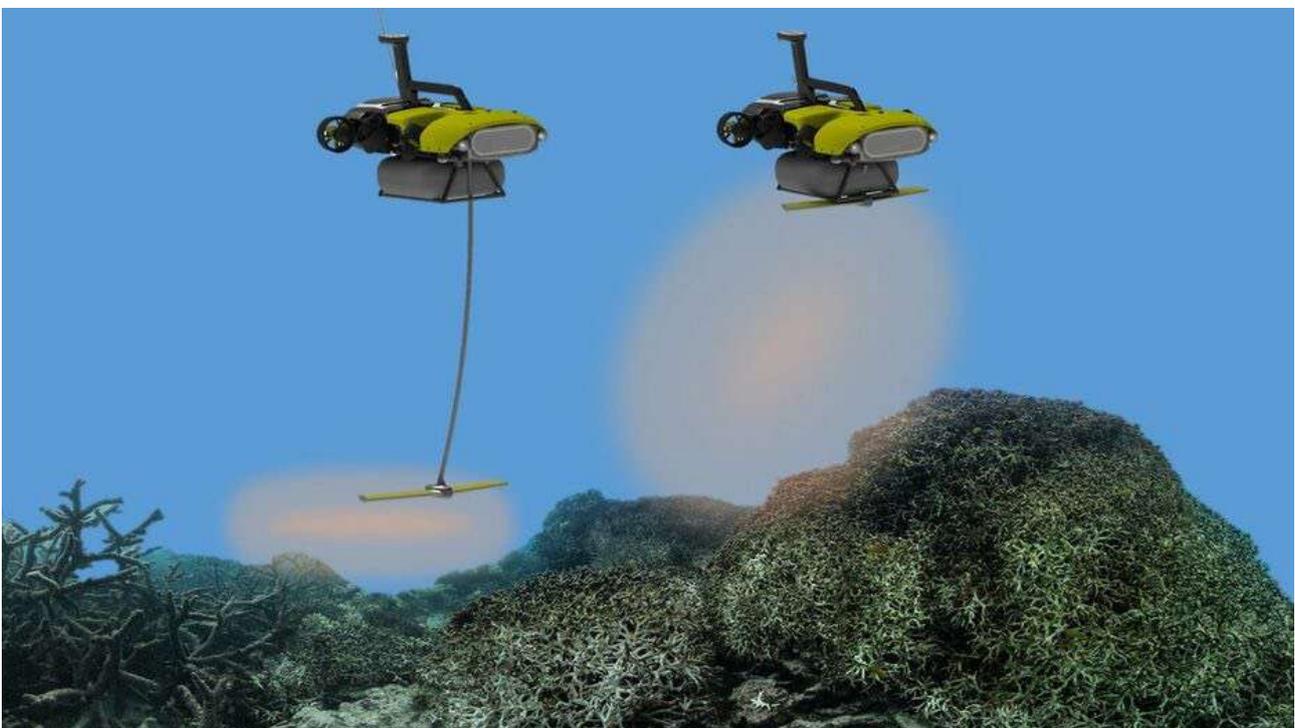


⁹² <https://worldoceanreview.com/en/wor-2/fisheries/illegal-fishing/>

⁹³ www.reefresilience.org/coral-reefs/stressors/local-stressors/overfishing-and-destructive-fishing-threats/

PRÉSERVER LES ENVIRONNEMENTS MENACÉS

- **La propagation d'espèces invasives menace les récifs coralliens**, et notamment l'acanthaster, une étoile de mer dévastatrice, sans prédateur naturel. Un tiers des récifs de la zone du Pacifique sont touchés par **l'acanthaster, qui a détruit 21% de la Grande Barrière de corail au cours de ces trente dernières années** ⁹⁴. En cause, le réchauffement des eaux, l'enrichissement en nutriments notamment d'engrais venant des eaux côtières et qui nourrissent les larves d'acanthasters. Jusqu'à présent, seul un décompte manuel permettait de surveiller leur propagation, avant de les enlever là où ils sur-consomment les récifs au-delà de leur capacité de régénération, ou de leur injecter du jus de citron ou du vinaigre. Le **RangerBot**, un drone sous-marin conçu par l'Université du Queensland (Australie), est **capable d'identifier avec un taux de précision de 99,4% cette « couronne d'épines »** pouvant aller jusqu'à 70 cm, et de **la neutraliser à l'aide d'une toxine mortelle**. Le **LarvalBot** est une autre version du RangerBot qui vise à identifier, toujours à l'aide d'une IA, les meilleurs endroits où déposer des larves de corail et ainsi reconstruire les récifs décimés.



⁹⁴ <https://www.brut.media/fr/international/acanthasters-ces-etoiles-de-mer-qui-devorent-les-coraux-43a078ca-b293-4a3d-9e2a-0bcba8cfdeb9>

- D'après l'ONU, **les forêts tropicales font l'objet de 90% d'exploitation illégale**, et un cinquième de la forêt amazonienne a été décimée sur les 50 dernières années ⁹⁵, principalement en raison d'une surexploitation agricole. Or, alors qu'elles ne couvrent que 2% de la surface du globe, les forêts tropicales sont **à l'origine de la production de 20% de notre oxygène**. Les détruire, c'est couper l'air que l'on respire, et menacer au passage tout leur écosystème : faune, flore, jusqu'aux indigènes... en accélérant le réchauffement climatique. Car **la déforestation est responsable de 17% des émissions mondiales de carbone**, alors qu'en comparaison tous les transports réunis génèrent 14% des émissions de CO2.

Pour s'attaquer à ce problème majeur, **The Rainforest Connection** utilise un **système de surveillance bioacoustique** à partir de panneaux solaires et de réseaux cellulaires, permettant d'alerter les patrouilles en temps réel lorsque sont perçus des bruits de tronçonneuse ou de véhicule. La lutte contre le braconnage est aussi facilitée par cette solution. Vu la surface et la densité d'une forêt tropicale, il ne serait **pas possible sans l'IA de détecter de telles intrusions**.



© Denis Lefebvre inspiré par Rainforest Connection

⁹⁵ https://www.maxisciences.com/deforestation/jusqu-a-90-de-la-deforestation-serait-d-origine-illegale_art26864.html

1.3. L'IA pour prévenir les catastrophes naturelles

Nous l'avons compris, l'IA se met au service de la sauvegarde de notre planète en permettant non seulement d'en avoir une **photographie générale en temps réel, mais aussi en créant des prédictions qui s'appuient sur différentes simulations**. Grâce à elle, nous pouvons détecter plus facilement et plus rapidement les changements climatiques pour alerter les autorités, leur permettre d'anticiper l'impact de catastrophes naturelles à venir, et de travailler à des mesures en amont permettant de les limiter.

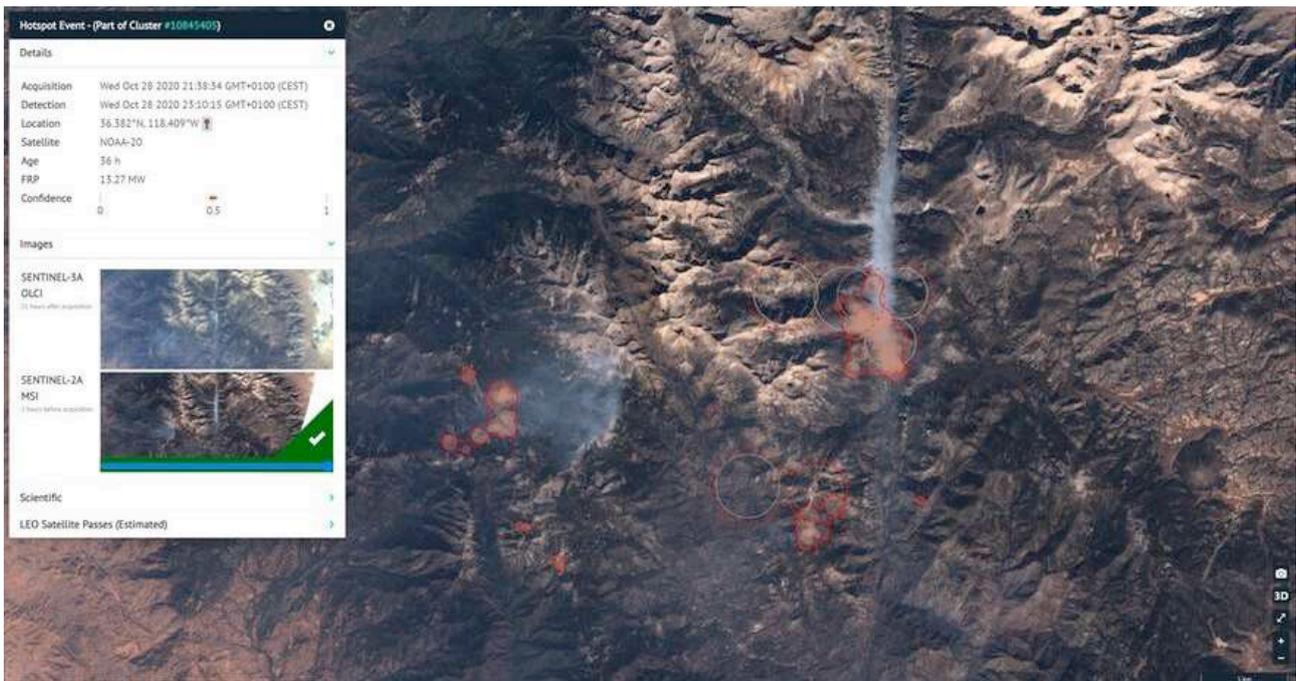
ANTICIPER LA DÉTECTION DES FEUX DE FORÊT

Depuis le début des années 1970, **les incendies de forêt en Californie ont été multipliés par huit**, et la superficie brûlée annuellement a augmenté de près de 500%. Selon une étude publiée dans *Earth's Future*, cette forte progression est liée à l'aménagement du territoire, mais aussi aux effets du changement climatique, notamment la sécheresse. ⁹⁶

La start-up **Descartes Labs**, basée dans le Nouveau-Mexique, s'est lancée dans **analyse géo-spatiale afin de localiser en seulement 9 minutes les départs d'incendie**. Pour cela, elle s'appuie sur les images du satellite GOES-16 qui fournit un flux vidéo toutes les 5 minutes, que Descartes Labs peut traiter et analyser dans les 4 minutes suivantes. Grâce à **plusieurs algorithmes qui prêtent chacun attention aux différentes propriétés des incendies** – présence de fumée, changement dans les données infrarouges thermiques montrant des points chauds, etc. – la **probabilité d'un démarrage d'incendie** qui est remonté aux gardes forestiers sur place pour vérification. Ce système de surveillance par l'Espace, dénommé « Wildfire Watch », leur permet en outre de **suivre efficacement l'évolution du feu** pour mieux l'arrêter, et anticiper l'évacuation de bâtiments menacés.

⁹⁶ <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019EF001210>

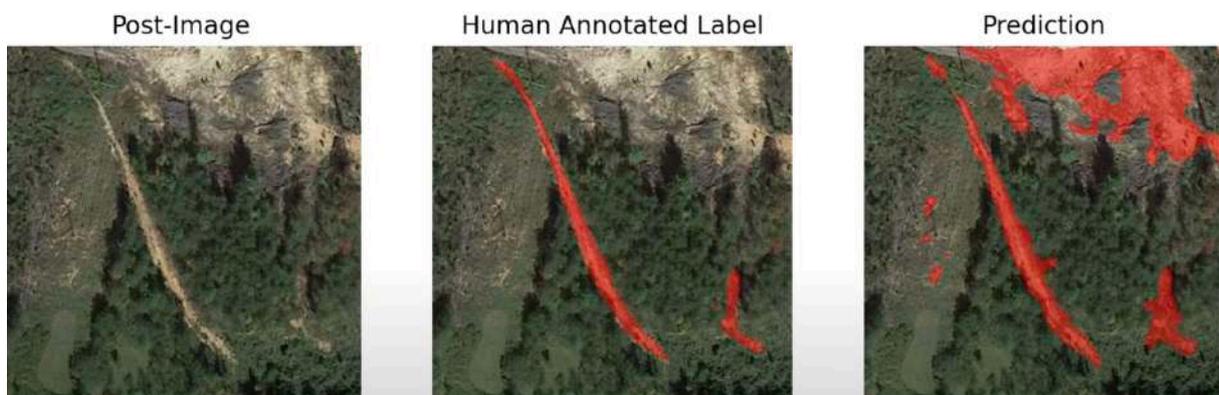
La startup allemande Orora Technologies en fait de même partout dans le monde.



ANTICIPER LES TSUNAMIS, LES SÉISMES...

Détecter de futurs glissements de terrain

Le projet universitaire deepLDB est l'un des 20 lauréats du Google AI Impacts Challenge 2019. L'équipe de chercheurs s'est fixée pour mission de prévenir le risque sociétal et environnemental lié aux glissements de terrain. Pour cela, deepLDB s'appuie sur des images satellites pour créer une base de données à même de **modéliser, à partir de l'apprentissage automatique, un outil de prédiction des risques.**



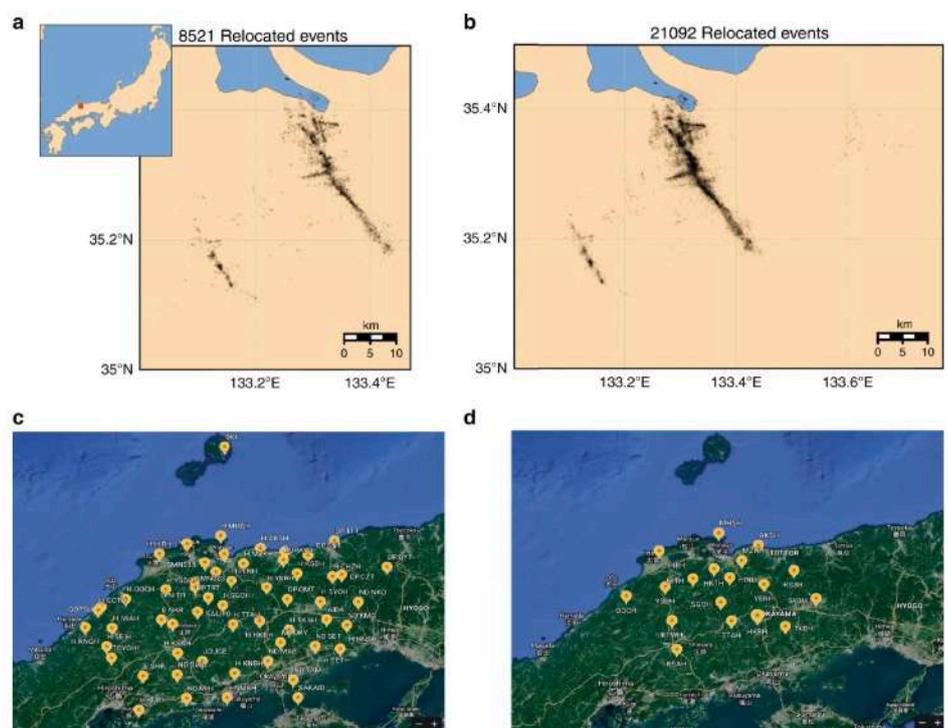
Anticiper les séismes et les tsunamis

- Les plus grands tremblements de terre se produisent environ une fois par an. Mais chaque jour, c'est plusieurs centaines d'épisodes sismiques très légers qui se font ressentir dans le monde, avec des impacts importants sur les villes.

Earthquake Transformer (Université de Stanford) **détecte les signaux faibles des tremblements de terre grâce à l'Intelligence Artificielle.** Les algorithmes développés sont **inspirés de ceux à l'œuvre lors du déclenchement vocal des assistants virtuels.** L'objectif est de réussir à distinguer les bruits sismiques des autres bruits perturbant leur identification, comme par exemple celui des vagues, des camions, etc. Sur la base de plus d'un million de données de sismogrammes enregistrées depuis deux décennies dans le monde, Earthquake parvient à mieux comprendre les minuscules signaux annonciateurs d'une brèche dans la croûte terrestre. En appliquant leur algorithme sur le séisme Tottori de magnitude 6,6 survenu il y a 20 ans au Japon, Earthquake a réussi à détecter et localiser plus de 20 000 événements sismiques, soit **2,5 fois plus que ceux détectés manuellement**, tout en n'utilisant que le tiers des données remontées par les stations sismiques (à droite sur le visuel ci-contre). L'IA pourrait être en

mesure d'aider à **anticiper les tremblements de terre en temps réel, et permettre aux autorités locales de mieux s'y préparer.**

From: Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking



Seismicity of Tottori region between 6 October and 17 November 2000. **a** relocated events in Fukuyama et al.²³ using manual phase picks by JMA. **b** relocated events using the automatic phase picker (EQTransformer) of this study. **c** Distribution of 57 seismic stations used by JMA and Fukuyama et al.²³ **d** distribution of 18 stations used in our study to detect and locate earthquakes in Tottori region.

- Conséquence directe des séismes : les risques de tsunami, comme celui qui a frappé le Japon en 2011 en causant 20 000 morts, en provoquant en cascade un accident nucléaire sans précédent à Fukushima. Pour **anticiper les risques de tsunamis** et leur impact sur les bâtiments, routes et infrastructures, les Universités de Tohoku et Tokyo au Japon utilisent le **super-calculateur le plus puissant du monde, 2,8x plus rapide que celui d'IBM**. Grâce à son énorme puissance de calculs, « **Fugaku** », conçu par Fujitsu en partenariat avec l'Institut de Recherche Riken, est capable de **simuler les conditions d'inondation avant même qu'un tsunami ne touche terre**, après avoir entraîné son algorithme de Deep Learning sur des milliers de scénarios ⁹⁷.

Ainsi, ONG, startups, chercheurs et géants de la tech peuvent travailler conjointement à l'élaboration de **modèles d'IA pour mieux circonscrire et prédire les risques** de pollution, les catastrophes naturelles, protéger notre environnement et sa biodiversité.

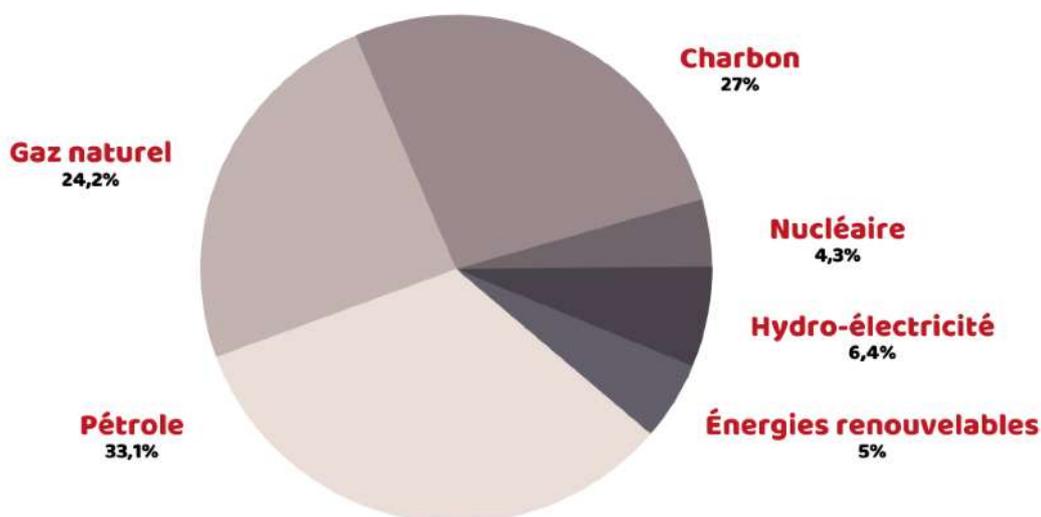
Mais la démarche ne peut pas être que curative : l'inertie du système climatique est telle que les 30 prochaines années de réchauffement sont irréversibles, et seront le résultat des combustibles fossiles déjà brûlés. Il faut donc se préparer et s'adapter aux risques climatiques prévus, mais il faut aussi et surtout s'attaquer aux problèmes à la racine, pour infléchir leur intensification. **L'Intelligence Artificielle va nous y aider, en nous offrant les outils à même de réussir notre transition énergétique.**

⁹⁷ <https://www.actuia.com/actualite/le-supercalculateur-fugaku-et-un-modele-dia-au-service-de-la-prevention-dinondations-dues-aux-tsunamis/amp/#click=https://t.co/q98gJ6sl62>

2. L'IA au service de la transition énergétique

La production électrique mondiale repose en grande majorité sur les combustibles fossiles tels que charbon, pétrole et gaz naturel, qui sont la plus importante source de polluants et d'émissions de GES.

Mix énergétique mondial



© Denis Lefebvre inspiré par BP, *Statistical Review of World Energy 2020*

À l'heure de l'intensification du réchauffement climatique, la transition énergétique est devenu non plus un objectif, mais une nécessité. L'Intelligence Artificielle est un allié de poids pour y parvenir : **grâce à l'IA, il devient plus facile de détecter et de corriger le gaspillage énergétique, de prédire pour la favoriser l'offre et la demande des énergies renouvelables, d'améliorer les propriétés thermiques** des appareils de chauffage, de climatisation, de ventilation, ainsi que des bâtiments.

Mais l'IA peut aussi **capter l'énergie produite par la donnée**, comme au sein des Data Centers, **pour la redistribuer intelligemment**, notamment au service de la Smart City. Dès lors, et plus que jamais, « l'IA est la nouvelle électricité »...

2.1. L'IA pour améliorer l'efficacité énergétique

D'après Capgemini, **l'Intelligence Artificielle devrait permettre d'améliorer l'efficacité énergétique de l'ordre de 15% dans les 3 à 5 prochaines années** ⁹⁸.

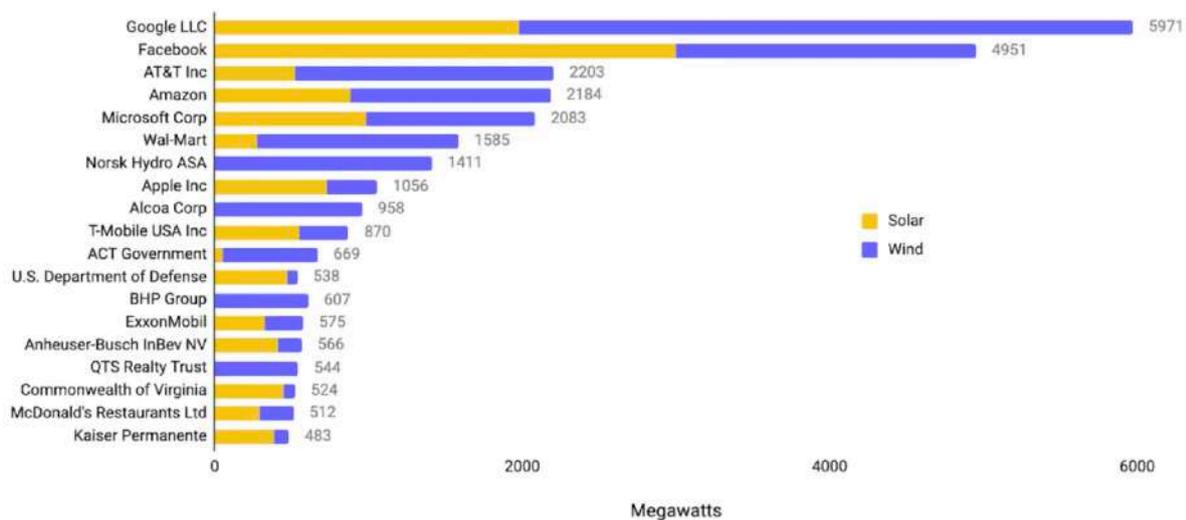
L'analyse de grandes quantités de données collectées à travers les systèmes électriques des infrastructures (Data Centers, bâtiments...), mais aussi des appareils IoT (compteurs et thermomètres intelligents, capteurs sur les turbines d'éoliennes ou sur les panneaux photovoltaïques...) permettent de **mieux comprendre et prévoir la consommation, mais aussi la production, la distribution, la commercialisation des énergies renouvelables, et la maintenance** des systèmes qui y sont associés.

AMÉLIORER LA RENTABILITÉ DES ÉNERGIES À FAIBLE IMPACT CARBONE

- **Les énergies « vertes »** (solaire, éolienne, hydroélectrique) utilisent les éléments naturels pour générer de l'électricité. Elles **ont l'avantage d'être présentes partout** dans le monde, et d'alimenter jusqu'aux endroits les plus isolés. Et surtout, elles ne génèrent **pas de carbone**. Elles semblent donc taillées pour répondre aux enjeux de la nécessaire transition énergétique.
- **Mais leur adoption est freinée par un problème majeur : ces énergies sont soumises aux aléas météorologiques**, qui rendent leur production instable, peu fiable et donc moins rentable. Le secteur de l'énergie fonctionne en effet sur l'anticipation, afin de garantir une continuité d'approvisionnement. **Pour viabiliser le marché des énergies renouvelables, il faudrait pouvoir prévoir la quantité de ressources disponibles.**
- C'est le défi auquel s'est attaqué Google pour atteindre son objectif « zéro carbone », qui est à présent légion chez l'ensemble des GAFAM pour réduire leur empreinte (p.205), en alimentant notamment leurs Data Centers en énergies plus « vertes ».

⁹⁸ https://www.capgemini.com/wp-content/uploads/2020/11/Climate-AI_Final.pdf

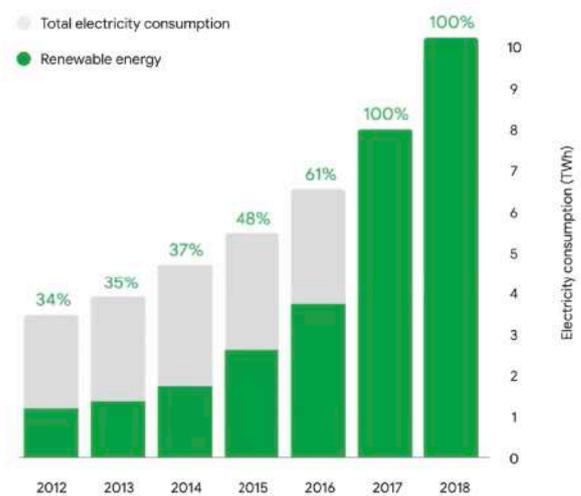
- Google est devenu le 1^{er} acheteur industriel d'énergie renouvelable, en s'alimentant auprès de plus de 50 parcs solaires et éoliens dans le monde, principalement aux États-Unis et dans le Nord de l'Europe. En 2019, le géant américain déclarait vouloir investir un montant record de 2 milliards d'euros dans les énergies renouvelables ⁹⁹.



Source: Bloomberg New Energy Finance (BNEF), as of January 2020
 Note: BNEF rates solar capacity in DC, while Google uses AC

Google

- Grâce à cette démarche, Google a franchi une étape significative en réussissant à faire **correspondre sa consommation d'énergie annuelle**, en constante augmentation, **avec son approvisionnement annuel en énergie renouvelable**. Amazon et Microsoft, quant à eux, n'en étaient « que » à 50% d'énergies renouvelables en 2018.

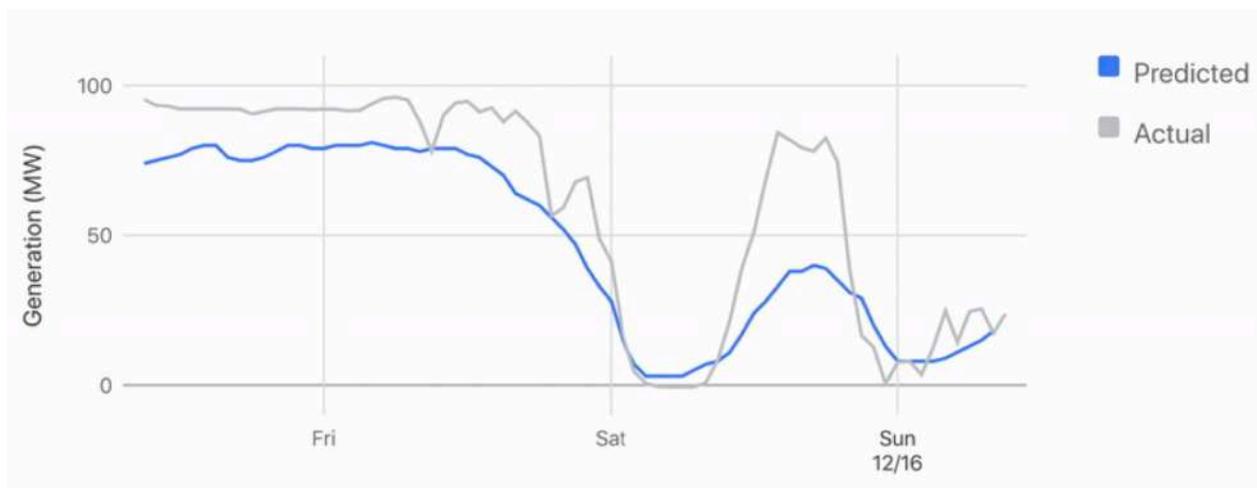


Google

- Pour réussir ce pari, il fallait que **Google aide ses fournisseurs d'énergie verte à viabiliser leur approvisionnement dans ses réseaux électriques**. Pour cela, il s'est appuyé sur les algorithmes de sa filiale britannique DeepMind. En combinant

⁹⁹ <https://siecledigital.fr/2019/09/23/google-investira-2-milliards-deuros-dans-les-energies-renouvelables/>

l'apprentissage par réseau de neurones aux données météorologiques, ainsi qu'à celles collectées sur les turbines d'éoliennes, **Deepmind a pu créer un modèle prédictif pour anticiper la production éolienne 36h à l'avance.** Les fournisseurs peuvent ainsi **garantir une livraison électrique optimale, mieux prévoir l'offre et la demande et ainsi réduire leurs coûts d'exploitation.** En outre, la vente par anticipation leur a permis de faire croître le revenu de leur parc éolien de +20%.



- Le modèle de Deepmind participe activement aux **Smart Grids** en combinant les technologies du numérique et de l'électricité. En se mettant au service de l'énergie éolienne, Google **augmente la viabilité du marché des énergies renouvelables** pour ses propres besoins, mais aussi et surtout pour **aider leur adoption plus largement** :



*Nous espérons que ce type d'approche d'apprentissage automatique pourra renforcer l'analyse de rentabilité de l'énergie éolienne et **favoriser l'adoption d'une énergie sans carbone sur les réseaux électriques du monde entier.***



Sims Witherspoon Responsable du programme Deepmind

- D'autres initiatives similaires se sont lancées par ailleurs pour **prédire grâce à l'IA la production d'électricité éolienne et solaire, comme en France avec meteo*swift**, ou à travers le partenariat entre Météo-France et Atos ¹⁰⁰. De quoi aider la France à atteindre son objectif d'une alimentation électrique à 32% d'origine renouvelable d'ici 2030, alors qu'elle en est aujourd'hui à un peu plus de 20%.

RÉDUIRE LA DÉPENSE ÉNERGÉTIQUE DES DATA CENTERS

- Selon IDC ¹⁰¹, il y aurait à présent **8 millions de Data Centers dans le monde**, et leur **consommation double tous les 4 ans**. Alors que le stockage des données absorbe aujourd'hui **3% de la production mondiale d'électricité**, et représente à lui seul 2% des émissions de GES – soit autant que le secteur aérien –, il pourrait nécessiter **3 fois plus d'électricité d'ici 2030**, et avoir une **empreinte carbone jusqu'à 7 fois plus importante d'ici 2040**. Soit autant que celle des États-Unis aujourd'hui... Il est vrai qu'avec le déploiement successif des nouveaux réseaux (5G, 6G...), la croissance des flux vidéos (Youtube, Netflix...) connus pour être les plus énergivores, l'essor des cryptomonnaies, la multiplication des données collectées et la puissance de calcul requise par l'IA... le besoin énergétique ne va que continuer de croître.
- **La dépense énergétique des Data Centers équivaut déjà à plus de 60% de leurs coûts d'exploitation** ¹⁰². C'est donc un enjeu économique majeur. L'alimentation en énergies renouvelables permet de réduire la consommation électrique de l'ordre de 9%, mais il s'agit aussi d'optimiser les **systèmes de refroidissement**, qui **consomment environ 40% de l'énergie des Data Centers**, nécessitent aussi de gros équipements industriels (climatiseurs, pompes...). Les serveurs de Google, par exemple, risquent en effet la **surchauffe** avec en moyenne **63 000 recherches par seconde à traiter...**

¹⁰⁰ <https://www.actuia.com/actualite/energie-atos-et-meteo-france-sassocient-pour-creer-une-plateforme-de-prevision-des-productions-delectricite-renouvelable/>

¹⁰¹ <https://www.lebigdata.fr/data-centers-environnement>

¹⁰² <https://influenceursduweb.org/impact-energetique-des-datacenters/>

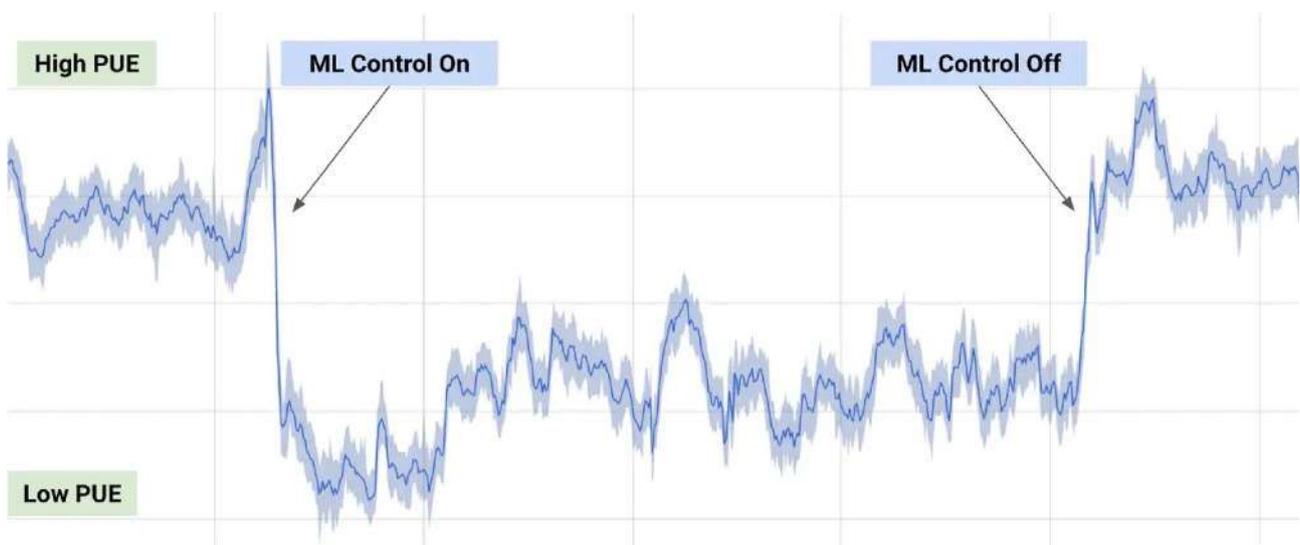
- Ce n'est donc pas un hasard si les Data Centers des géants de la Tech fleurissent dans les Pays Nordiques, comme en Norvège où se trouve le plus grand Data Center au monde (600 000 m²), ou **en Finlande où les serveurs de Google sont refroidis à l'aide de l'eau glaciale de la Mer Baltique** ¹⁰³. Microsoft quant à lui est allé jusqu'à immerger pendant 2 ans un conteneur rempli de serveurs dans les fonds marins, à 35 mètres de profondeur à proximité des côtes écossaises. Le « **Project Natick** » ¹⁰⁴ a aussi pu être alimenté en énergie solaire et éolienne, et il serait 8 fois plus fiable que sur Terre, grâce à une atmosphère intérieure très azotée qui limite la corrosion (comme c'est le cas avec l'oxygène), et les risques de panne ou d'incendie. Vu le succès de ce premier test, **les Data Centers sous-marins de Microsoft devraient se multiplier**, même si leur impact sur l'écosystème marin n'est pas clairement établi.



¹⁰³ <https://www.lebigdata.fr/google-investissement-data-center-finlande>

¹⁰⁴ <https://www.numerama.com/tech/648236-sous-leau-depuis-deux-ans-le-data-center-sous-marin-de-microsoft-a-tenu.html>

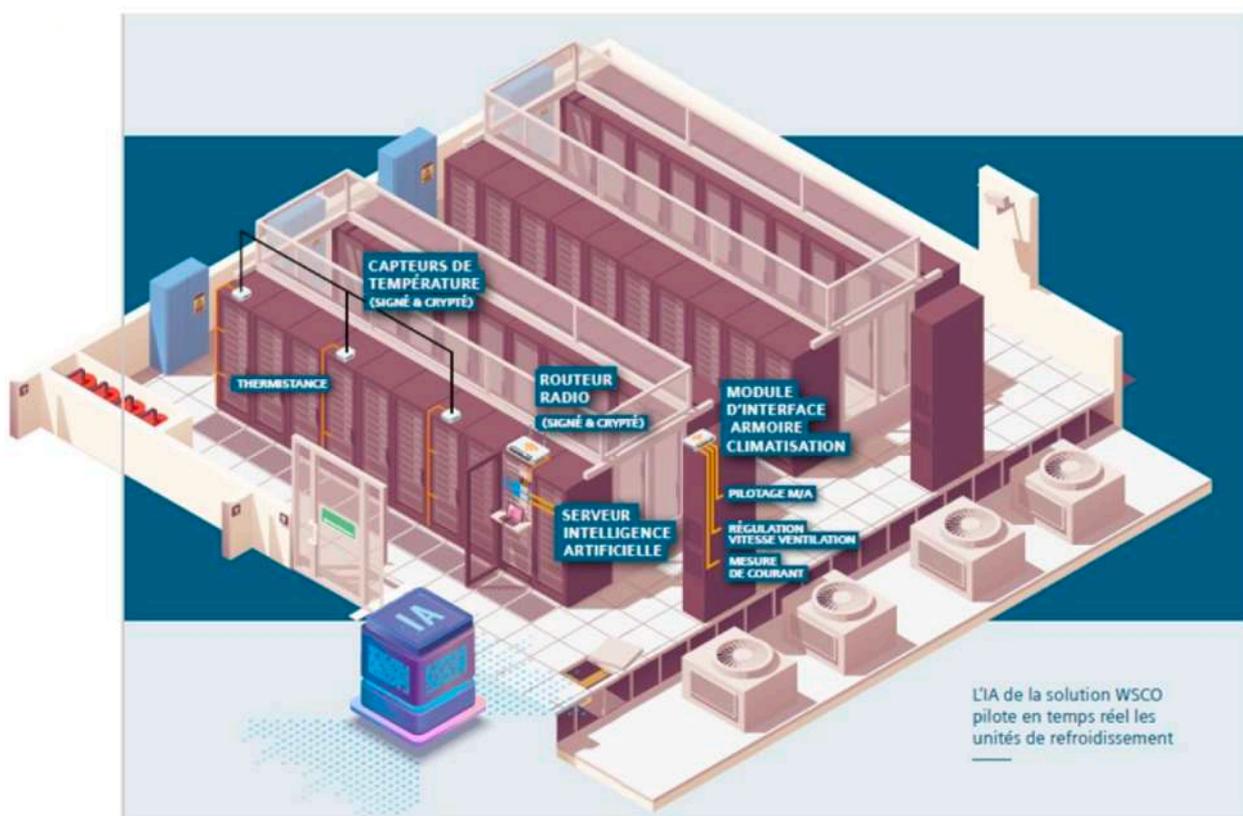
- Mais pour assurer des performances optimales en étant au plus près des utilisateurs, tout en respectant les réglementations sur l'hébergement des données, les Data Centers ne peuvent pas tous être situés dans les zones froides. Utiliser l'air extérieur pour refroidir les serveurs reste possible jusqu'en région parisienne, comme a pu le démontrer Scaleway et son procédé astucieux utilisant l'évaporation de l'eau sur des filtres à air, pour un **système de refroidissement sans climatisation**.¹⁰⁵
- Mais **le Machine Learning peut aussi contribuer largement à optimiser le système de refroidissement des serveurs**. Ainsi Google a de nouveau fait appel à Deepmind et son algorithme d'apprentissage par renforcement, qui a appris pendant 2 ans par essais et erreurs, et a permis une **réduction de 40% du montant de l'énergie utilisée pour le refroidissement**. Le graphique ci-dessous montre bien l'impact du Machine Learning, activé puis désactivé lors d'une journée de test, sur le PUE (*Power Usage Effectiveness*) du Data Center, indicateur d'efficacité énergétique.



- La diminution de la facture énergétique des Data Centers permise par l'IA se mesure aussi en terme de **maintenance prédictive**. En effet, les conséquences financières d'une panne peuvent être importantes. Selon une étude Gartner de février 2020, **le**

¹⁰⁵ <https://www.01net.com/actualites/on-a-visite-dc5-le-data-center-de-scaleway-au-systeme-de-refroidissement-unique-et-ecologique-2043880.html>

coût de l'arrêt d'un Data Center se situe en moyenne à \$5600 /minute ¹⁰⁶, en plus des pénalités de *Service Level Agreement* conclu avec ses clients, et d'autres coûts indirects (réputation, perte de clients et de données...). L'**anticipation des risques de panne**, et l'**aide au pilotage** pour limiter les causes d'indisponibilité liées pour 25% à des erreurs humaines, permettent d'assurer une continuité de service optimale. Ainsi, **Siemens utilise la solution WSCO et son Intelligence Artificielle**, pour non seulement piloter en temps réel la température des salles serveurs, grâce à des capteurs, mais aussi réagir en moins de 1 minute à la détection de dysfonctionnements. ¹⁰⁷



Je vous invite à présent à découvrir **mon interview de Jérôme Totel, Vice Président chez Data4**, qui nous explique les actions déployées pour limiter la dépense énergétique de ses 42 Data Centers en France, en Italie, au Luxembourg et en Espagne, notamment en ayant recours à l'Intelligence Artificielle.

¹⁰⁶ <https://www.entrepreneur.com/article/345826>

¹⁰⁷ <https://datacenter-magazine.fr/siemens-optimise-lefficacite-energetique-dans-les-data-centers-grace-a-la-solution-intelligente-white-space-cooling-optimization-wsco/>



Jérôme Total

Vice Président de la Stratégie de Data4



Quelles sont les actions mises en œuvre par Data4 pour réduire l'empreinte carbone de ses Data Centers ? Est-ce que vos clients font de cet engagement durable un vrai critère de choix ?

« Le Data Center est un gros consommateur d'énergie. L'une des choses les plus importantes qu'on fait, depuis 3 ans déjà c'est **réinjecter 100% de notre consommation énergétique dans le réseau de notre opérateur d'énergie, en énergie renouvelable.**

C'est ce qu'on fait en France, en Italie, et en Espagne où l'on est en train d'ouvrir un Data Center. On a la chance en France d'avoir une énergie très décarbonée puisque c'est du nucléaire, avec un impact CO2 assez faible. Mais ça dépend aussi de chaque pays.

Ensuite on a une autre initiative qu'on a lancée il y a 3 ans : un projet intitulé **Smart Data Center**, et qui permet à nos clients d'acheter un logiciel de Data Center Infrastructure Management (DCIM), qui est connecté à leur portail Data4, et qui leur permet de **suivre en temps réel tout ce qu'il se passe dans le Data Center, notamment dans leurs propres espaces d'hébergement.** Car vous le savez, un Data Center c'est un bâtiment où il y a plein de capteurs partout pour vérifier que tout fonctionne en permanence. Quand il y a un problème, un opérateur le reçoit sur son écran et le traite le plus rapidement possible. On utilisait ces données en interne, mais on s'est dit qu'il serait intéressant de les partager avec nos clients. On a donc fait un long travail de modélisation numérique de tous nos bâtiments, avec nos groupes électrogènes, climatiseurs, etc. Et on a rajouté dans le logiciel **une mesure de l'impact environnemental de chaque équipement, sur**

l'ensemble de son cycle de vie : fabrication, opération et destruction, recyclage. Sur un bâtiment qui fait 4000 mètres carrés au total, on est donc capable de dire à notre client, grâce à un ensemble de calculs : si vous nous avez acheté 1/4 d'hébergement dans le bâtiment, **l'impact de votre informatique chez nous est de tant de CO2, tant d'eau, tant de ressources rares et tant d'énergie**. Ça fonctionne et on l'a lancé depuis fin 2020.

En parallèle de ça, on est en train de travailler avec une Junior Entreprise pour définir **une quinzaine de KPIs pour mesurer encore plus finement l'impact environnemental de nos bâtiments et des personnes qui y travaillent**. Par exemple, l'utilisation du béton a un impact important sur le CO2. Ça fait 15 ans que la société existe, on a aujourd'hui deux générations de Data Centers, et ça nous permet de savoir d'où on part et où on veut aller. Quel sera demain le bâtiment qu'on construira ? Est-ce que ce sera toujours le même ? Peut-être que oui, peut-être que non. On travaille aussi beaucoup avec nos fournisseurs, comme Schneider Electric, pour qu'ils nous transmettent des PEP, qui sont **des fiches techniques très détaillées et qui affichent l'impact environnemental de leurs**

Environmental impact assessment of the information system hosted at Data4

1000 cabinets - 2,000 sqm
Report generated with data collected in 2019

Multiples actions ▾

Environmental impact



991 t CO2 eq
849.65 Paris / New York flights (per passenger)



3,180 m3
0.85 olympic pools

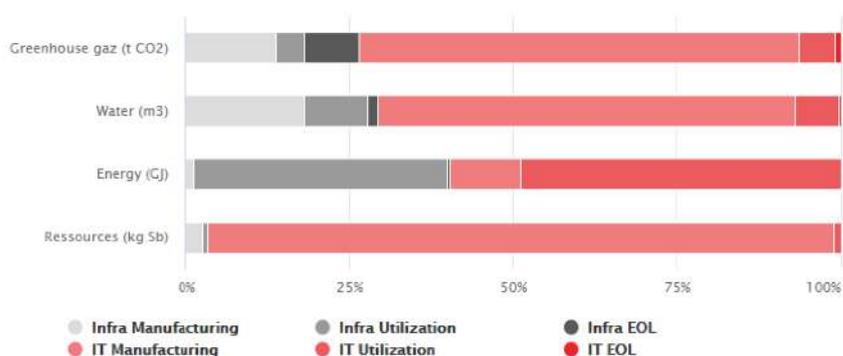


79,420 GJ EP
2,760.42 people for one year



229.21 kg Sb eq
6.37 kg of pure gold

Impacts breakdown



Le projet "Smart Data Center" s'inscrit dans le programme RSE ambitieux de Data4, en faveur notamment du développement durable.

équipements. On injecte toutes ces données dans notre modèle de *Smart Data Center*, pour être le plus précis possible.

Au final oui, **nos clients souhaitent s'assurer que tout est fait par le Data Center pour limiter son impact.** Parmi ceux qui achètent des espaces de colocation chez nous pour installer leurs équipements informatiques, on a des PME ou des grands comptes du CAC40, mais on a aussi des grosses plateformes numériques, incluant de la TV à la demande ou encore des fournisseurs de Cloud. Certains ont parfois construit leurs propres Data Centers. Quand ils achètent de la colocation ailleurs, il faut donc **montrer patte blanche sur tous ces aspects là, qui sont importants pour eux, et qui sont donc aussi importants pour nous.**

Avez-vous recours à l'IA pour limiter l'énergie consommée par les serveurs ?

C'est plus que d'actualité puisqu'on est en train de faire **4 POC avec 4 fournisseurs différents sur 2 sujets d'IA.**

Le premier, c'est autour de la maintenance prédictive. Aujourd'hui, des hébergeurs comme Data4 s'engagent sur un taux de disponibilité de plus de 99% pour l'électricité, et sur des plages de température pour la climatisation, pour assurer le bon fonctionnement des serveurs. Donc ça nous intéresse de voir si les technologies à base d'IA pour de la maintenance prédictive permettraient d'améliorer ce taux de disponibilité.

Le second sujet, c'est sur l'utilisation de l'IA pour améliorer l'efficacité énergétique. On est en train de tester sur un bâtiment en production le gain énergétique en appliquant des modèles, notamment à base de jumeau numérique. Est-ce qu'il y a un gain ? Si oui, de combien ? On est en train de regarder tout ça de près, en nous appuyant sur les 5 000 à 10 000 capteurs installés dans chacun de nos Data Centers, et qui permettent d'obtenir assez de données à analyser par des algorithmes.

Face à la surchauffe des serveurs : les refroidir ou utiliser la chaleur produite ?

C'est une démarche dans laquelle on s'inscrit puisque nous avons **investi dans la société Qarnot** à Bordeaux. Ils ont installé dans des logements sociaux **des radiateurs qui sont des mini Data Centers**, et qui permettent à la fois de faire du calcul HPC pour des banques, par exemple, tout en chauffant des citoyens. Maintenant, il y a **des contraintes pour nous en tant que fournisseur de Data Centers**. Car la chaleur, elle ne se transporte pas sur des kilomètres. Il y a beaucoup de déperditions de chaleur, donc il faut être dans un environnement assez proche des citoyens. Est-ce que le Data Center de demain sera en plein centre-ville ? J'en doute, mais ça dépend de la taille dont on parle. »



D4

42

Data Centers en France, en Italie, au Luxembourg et en Espagne.

220

clients opérateurs Cloud et Telecom disposant de 123ha de surface de stockage, faisant de Data4 un leader européen.

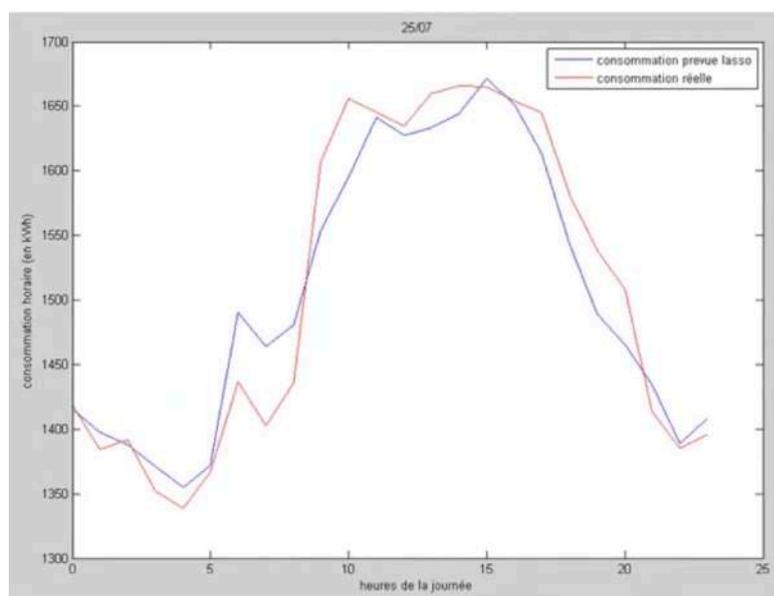
100%

d'alimentation par des sources d'énergie décarbonées d'ici 2030 pour atteindre la neutralité carbone de ses Data Centers.

AUTRES AXES D'OPTIMISATION ÉNERGÉTIQUE

Gérer l'optimisation énergétique des bâtiments

- Le secteur du bâtiment génère à lui seul **44% de la consommation d'énergie en France, et un quart des émissions de CO2** ¹⁰⁸. L'efficacité énergétique est donc clé. L'Etat a pris des mesures incitatives pour **lutter contre les passoires thermiques** en octroyant des primes énergie pour la rénovation. Mais l'Intelligence Artificielle peut aussi permettre de mieux gérer la dépense énergétique des bâtiments.
- Lors d'un webinar de Planet Tech'Care en février 2021 ¹⁰⁹, Juliette Mattioli, Senior Expert IA chez Thales, a exposé **comment l'IA a permis d'analyser puis d'optimiser la consommation énergétique du bâtiment Thales Research and Technology** sur le plateau de Saclay. 72% de ce site de 3600m² est constitué de salles blanches, qui nécessitent 4 Centrales de Traitement d'Air (CTA) pour assurer un empoussièrément minimal, une température et une hygrométrie idéale. Or, ces CTA représentent 80% de la consommation énergétique du bâtiment, et l'un de ses plus gros postes de dépenses : 2M€ de facture d'électricité par an, et des pénalités de dépassement au-delà des seuils prévus de consommation pouvant aller jusqu'à +80k€. Pour diminuer ses coûts, Thales a alors mis en place un **modèle de prévision de la consommation énergétique par apprentissage**, tenant compte en particulier des conditions météorologiques. 3 ans de données historiques sur le traitement de l'air extérieur par les CTA ont pu être exploitées pour créer ce modèle de prédiction journalière de consommation. **Il s'est révélé fiable**, et permet à présent d'**anticiper toute sur-**



¹⁰⁸ <https://www.ecologie.gouv.fr/energie-dans-batiments>

¹⁰⁹ <https://systematic-paris-region.org/evenement/planet-tech-care-atelier-5-ia-et-consommation-energetique-par-systematic/>

consommation en régulant les activités sur le site, tout en remontant des anomalies révélées par des incohérences entre prévision et consommation réelle. Il devient ainsi un indicateur pour savoir s'il faut changer les filtres d'empoussièrement.

- L'optimisation énergétique entre **jusque dans les maisons individuelles**, et devient l'affaire de tous. Ainsi les **thermostats intelligents de Netatmo ou Ween**, récemment rachetés par les fabricants de matériel électrique Legrand et Atlantic, permettent de faire des économies d'énergie allant jusqu'à 25%. Ces thermostats permettent de **dresser le profil thermique d'une maison** selon ses caractéristiques, ainsi que la météo locale, la détection de fenêtres ouvertes, etc. Grâce aux algorithmes d'apprentissage développés par Ween, son thermostat intelligent, connecté aux smartphones de ses habitants, est même capable de **détecter leur éloignement de la maison et de réduire la température, puis la remonter lorsqu'ils sont sur le point de revenir**, en calculant leur temps de retour en fonction de la circulation.



Réduire la consommation énergétique des devices

- L'IA peut également aider à diminuer la consommation énergétique des tablettes, smartphones et ordinateurs. Les mises à jour d'applications et d'OS, mais aussi le téléchargement automatique de nouveaux podcasts par exemple, font appel à des tâches de calcul intensif en sur-sollicitant la batterie. Des chercheurs de l'Université de Melbourne ont ainsi eu recours au **Machine Learning pour anticiper le moment le**

plus opportun pour faire ces mises à jour, en prédisant le moment où un utilisateur risque de déverrouiller son mobile. ¹¹⁰

- Une fois vérifiées les mises à jour disponibles, le chargement des données pourrait se faire quand l'utilisateur ne sollicite pas son appareil, et notamment la nuit. De quoi économiser la batterie et limiter le renouvellement des appareils.



Au lieu de vérifier si des mises à jour sont disponibles toutes les 5 minutes, le mobile pourrait le faire **une seule fois, juste avant un déverrouillage**.

L'utilisateur bénéficierait de **rechargements moins fréquents**, et d'une **meilleure expérience**. De plus, le système tournerait à sa vitesse optimale.



Vassilis Kostakos Professeur Interaction Homme-Machine, Université de Melbourne

2.2. L'énergie des données : « l'IA est la nouvelle électricité »

Pour réduire encore le coût énergétique des Data Centers, une autre idée commence à faire son chemin : **et si, au-delà de chercher à refroidir la chaleur produite par les serveurs, on s'en servait aussi comme d'une source nouvelle d'énergie ?**

De quoi assurer la transition énergétique tout en limitant drastiquement la facture liée au besoin grandissant en calculs et en hébergement des données, au-delà du seul cas d'usage de l'Intelligence Artificielle.

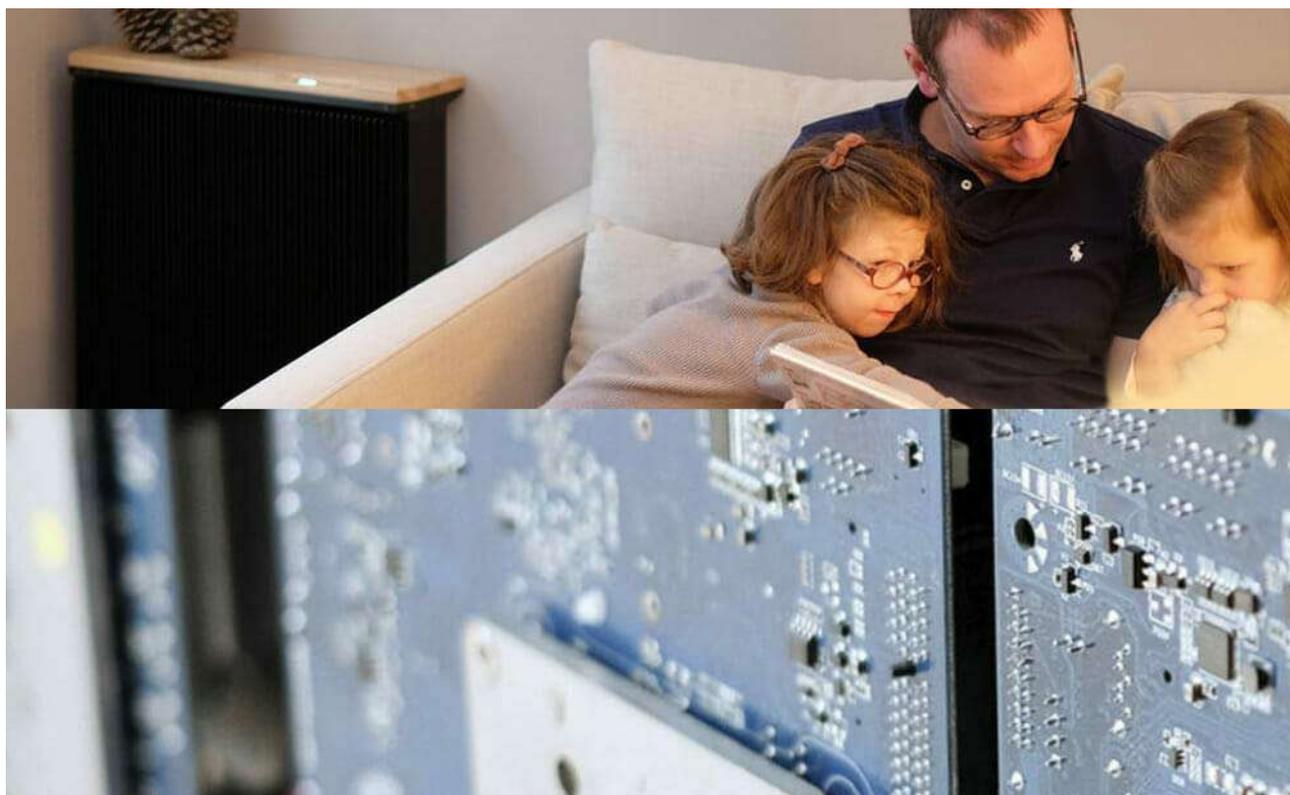
C'est la mission que se sont fixés différents acteurs, parmi lesquels **Qarnot et son radiateur-ordinateur qui exploite la chaleur fatale informatique**.

¹¹⁰ <https://www.lemondeinformatique.fr/actualites/lire-ameliorer-l-efficacite-energetique-des-mobiles-avec-l-ia-74529.html>

DES SERVEURS EN GUISE DE RADIATEURS DANS LES HLM

Depuis 2010, la société française Qarnot s'est lancée sur une idée simple : **chauffer les bâtiments grâce à la puissance des calculs** ! Là où la chaleur dégagée par les serveurs était une contrainte, Qarnot l'a vu comme une ressource, capable de servir tout aussi bien les hébergeurs et producteurs de calculs que les infrastructures en quête d'un chauffage écologique.

Très concrètement, la chaleur informatique est gérée depuis la plateforme de Qarnot pour être distribuée vers **des radiateurs-ordinateurs et des chaudières numériques embarquant des microprocesseurs**. Ce qui en fait des **mini Data Centers à domicile**, capables pour le premier de chauffer une pièce d'environ 30 m², mais aussi de gérer toute une gamme domotique : surveillance de la qualité de l'air, borne de recharge mobile, borne wifi gratuite, etc.



Les clients de calculs paient à Qarnot des heures de puissance informatique, qui sont ensuite redistribuées gratuitement sous forme d'énergie en grande majorité à des

logements sociaux. La répartition des calculs permet de gérer la distribution du chauffage à travers tout le « parc » de Qarnot, selon les saisons et les usages.

Le chauffage aux données réduit drastiquement l’empreinte carbone des besoins en calcul informatique, dans un cercle écologique et économique vertueux. Qarnot a d’ores et déjà déployé ses installations majoritairement dans des HLM en Ile-de-France et Gironde (300 logements sociaux fin 2020), mais aussi dans des incubateurs, collèges et des bureaux. Les chaudières numériques sont utilisées à Nantes par exemple dans des bains-douches pour les sans-abris, et elles chauffent 10 000 foyers de la ville de Kankaanpää, en Finlande, en étant intégrées directement au réseau local. Un partenariat entre Qarnot et Engie devrait permettre de voir se déployer peu à peu cette initiative.

D’autres projets du même type ont ainsi vu le jour :

- La **Mairie de Paris** a ouvert depuis 2019 **son propre Data Center**, en plein cœur de Paris. Sur 760 m², dans les sous-sols du 18^{ème} arrondissement, il contient pas moins de 300 serveurs physiques et 3500 serveurs virtuels. La Mairie de Paris a souhaité investir à hauteur de 16M€ dans ce projet, pour garder la main sur les données des Parisiens et assurer à la fois leur sécurité et leur conformité avec le RGPD. La Mairie s’est par ailleurs inscrite dans une démarche d’économie circulaire **en récupérant la chaleur produite par ses propres serveurs pour alimenter le réseau de chauffage** d’habitations du quartier, mais aussi de serres situées sur le toit du bâtiment.¹¹¹
- La Mairie avait déjà conclu depuis 2017 un partenariat avec la société Stimergy pour **chauffer en partie la piscine de la Butte-aux-Cailles**, dans le 13^e arrondissement. La « salle serveurs » éco-responsable de Stimergy est installée directement dans le sous-sol du bâtiment, et permet d’économiser 45 tonnes de CO₂ par an. De façon concrète, **les serveurs sont immergés dans un bain d’huile**, ce qui permet de ne pas avoir besoin de système de refroidissement énergivore, tout en récupérant la chaleur produite pour la rediriger vers le ballon d’eau chaude, en faisant monter la

¹¹¹ <https://www.lebigdata.fr/paris-data-center>

température jusqu'à 60°C. Les **mini Data Centers écologiques** de Stimergy sont aussi déployés dans des écoles et des logements sociaux.

- La valorisation de la chaleur excédentaire est ainsi vouée à devenir un enjeu clé des Data Centers du monde entier, comme c'est encore le cas avec celui de **DigiPlex** à Ulven, en Norvège. Ce Data Center est non seulement alimenté en énergie durable, mais il fait office de projet pilote en étant **relié directement au système de chauffage urbain de la ville d'Oslo**. Il devrait couvrir d'ici 2021 la demande annuelle de chaleur de plus de 5 000 appartements.

CHAUFFER LES HÔPITAUX ET LES ÉGLISES AUX BITCOINS

Comme nous l'avons vu précédemment, **la Blockchain requiert de fortes puissances de calcul, et son minage est très énergivore**. Selon le *Bitcoin Electricity Consumption Index* de l'Université de Cambridge, le Bitcoin aurait consommé 2/3 de la consommation de tous les Data Centers du monde entier en 2019, et plus que 159 pays dans le monde.¹¹² Visual Capitalist estime même que **le Bitcoin nécessiterait 10 fois plus d'énergie que Google !**¹¹³

Après ces annonces fracassantes au Printemps 2021 sur l'impact écologique du Bitcoin, le marché des cryptomonnaies dans son ensemble a chuté pour entrer en *Bear market*, sous l'impulsion d'un tweet de Elon Musk suspendant les paiements Tesla en Bitcoin.¹¹⁴ L'Etat de New York avait aussi présenté un projet de loi visant à suspendre le minage de cryptomonnaies le temps d'évaluer leur empreinte environnementale, bien que 65% du minage se faisait alors en Chine... avant qu'il n'y soit à son tour interdit. Nous l'avons vu cependant, d'autres acteurs comme Ethereum ou Elrond œuvrent à **une Blockchain plus écologique** (p.123).

¹¹² <https://cbeci.org/>

¹¹³ <https://www.visualcapitalist.com/visualizing-the-power-consumption-of-bitcoin-mining/>

¹¹⁴ <https://www.macg.co/aillieurs/2021/05/elon-musk-decouvert-les-problemes-environnementaux-du-bitcoin-121423>

Toujours est-il que pour limiter son impact, **la piste inédite d'une Blockchain qui serait aussi productrice d'électricité a vu le jour**. Car après tout, les radiateurs-ordinateurs de Qarnot chauffe bien les HLM de Bordeaux grâce aux données de marché boursier de BNP Paribas... Plusieurs initiatives sont à noter :

- La start-up française **Tresorio**, qui développe des calculateurs pour la Blockchain, a installé en 2018 **une partie de ses serveurs dans les sous-sols de l'hôpital de Metz-Thionville**. Une chaudière récupère la chaleur thermique produite par les serveurs. Une armoire de 63 cartes graphiques permet de chauffer à elle seule une surface d'environ 250 m². Ce projet, mené en partenariat avec Dalkia, filiale d'EDF, a depuis été déployé sur d'autres structures, et la startup a depuis lancé une **plateforme de « Cloud écologique » à destination des entreprises, incluant un service de minage éco-responsable**, fonctionnant sur le même principe que pour le CHR de Metz.



- Une **expérimentation plus inattendue** a lieu depuis 2020 **en l'église de Saint-Adrien, au Québec**. Les bâtiments religieux sont connus pour être particulièrement difficiles à chauffer. La réhabilitation de l'église Saint-Adrien en pôle créatif par Pierre-Philippe Côté, un artiste québécois, aurait nécessité \$60 000 /an de frais de chauffage. Avec

Patrick Lussier, ils ont alors l'idée d'installer **35 serveurs de données dans les sous-sols de l'église. Ils y minent de la cryptomonnaie, et la chaleur qui est dégagée sert à chauffer tout le bâtiment.** Fort du succès de ce premier test, ils ont lancé le projet BTU, qui ambitionne de **préserver le patrimoine bâti grâce aux serveurs de traitement des données, tout en œuvrant pour la transition énergétique** en repensant la façon dont nous pouvons nous chauffer.



Nous arrivons au terme d'un **bilan carbone de l'IA qui, au final, pourrait avoir un impact plus positif que négatif sur l'environnement**, si l'on en juge l'étude de PwC qui estime que l'IA pour l'action climatique a le **potentiel de réduire les émissions mondiales de GES d'ici 2030 jusqu'à 4%** ¹¹⁵, là où sa consommation actuelle ne représente qu'une partie des 3,5% du numérique.

Ceci, comme nous l'avons vu, **sous réserve que les modèles d'IA soient développés et déployés de façon plus frugale.** Les chercheurs y travaillent dans leurs laboratoires, de même que les concepteurs de hardware.

Mais c'est aussi de la responsabilité des entreprises qui ont recours à l'IA, de **faire les choix de modèles et d'infrastructures à même d'assurer leur propre éco-responsabilité**, à l'heure où elles se voient soumises à la transparence de leur empreinte carbone.

Comment donc, dans le cas qui nous intéresse, tendre vers une IA marketing durable ? **Comment bénéficier des super-pouvoirs de l'IA marketing tout en préservant notre planète ? Suivez dans ce troisième chapitre mon mode d'emploi pratique !**

¹¹⁵ « How AI can enable a sustainable future », PwC, 2019 : <https://www.pwc.co.uk/sustainability-climate-change/assets/pdf/how-ai-can-enable-a-sustainable-future.pdf>

CHAPITRE 3

Une IA durable pour un marketing éco-responsable



L'IA marketing à l'ère de l'éco-responsabilité

Dans ma dernière partie, je vais vous proposer un **mode d'emploi pratique pour que vous puissiez répondre aux enjeux de durabilité dans le cadre du déploiement de vos projets d'IA marketing**, tels que ceux décrits dans le chapitre 1.

Le framework proposé est **adapté pour correspondre aux profils d'entreprises de toute taille** : celles qui développent leurs modèles d'IA marketing elles-mêmes, comme celles qui passent par des solutions SaaS tierces et clés en main.

Le modèle d'une IA marketing durable s'impose d'autant plus que les entreprises sont à présent soumises, comme nous allons le voir à présent, **à la nécessité d'adopter une démarche éco-responsable**.

1. La déferlante du marketing éco-responsable

D'après vous, qui arrive **en tête des 11 tendances marketing à suivre en 2021** d'après Hubspot ? ¹¹⁶ Le « **marketing éco-responsable** », aussi appelé « marketing écologique » ou « marketing vert ».

Il est **l'une des branches les plus dynamiques du marketing responsable**, qui englobe également une dimension éthique et sociale pour promouvoir les valeurs de la marque, et en faire ***in fine* un argument de notoriété et de vente**.

Le marketing éco-responsable vise donc à **valoriser le positionnement écologique d'un produit ou de la marque elle-même, et/ou à communiquer sur les engagements de la marque pour œuvrer en faveur de la protection de l'environnement**, dans le choix de ses fournisseurs, ses modes de fabrication... et jusqu'à ses modes de communication.

¹¹⁶ <https://blog.hubspot.fr/marketing/tendances-marketing>

1.1. De nouvelles habitudes de consommation

C'est un phénomène récent et grandissant : les entreprises orientent de plus en plus leur stratégie marketing autour de la question environnementale. Très peu le faisaient jusqu'alors, ou de façon très fragmentée et de manière non pro-active.

Mais **face à la prise de conscience collective des impacts du changement climatique, les consommateurs font évoluer leurs pratiques d'achat vers le « mieux consommer »**. Les marques doivent dès lors se plier à ce nouveau paradigme.

L'ENVIRONNEMENT : UN ENJEU MAJEUR POUR LES FRANÇAIS

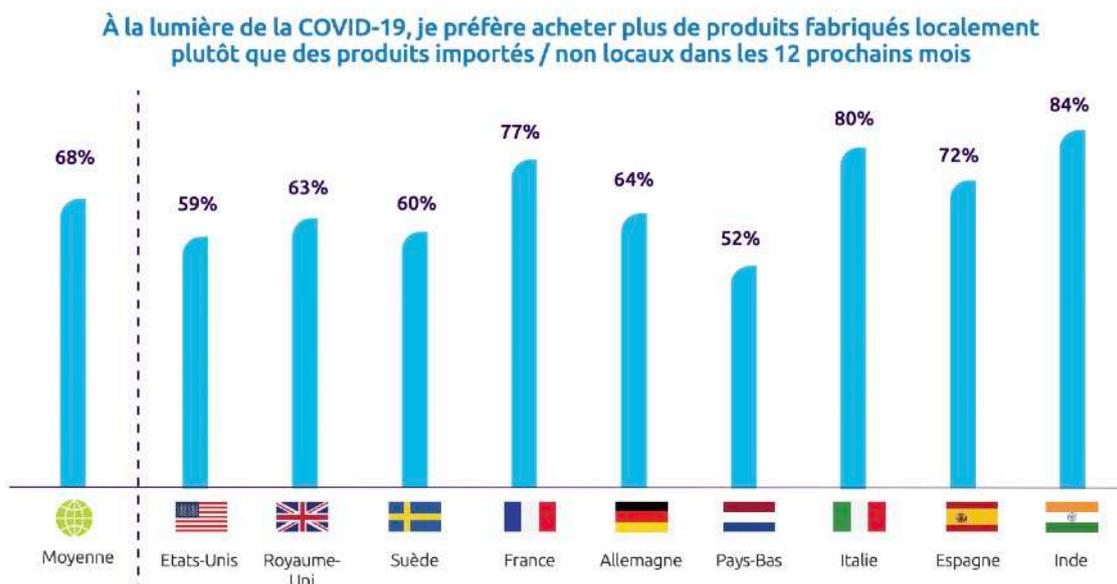
Le baromètre 2020 de l'ADEME, l'Agence gouvernementale de la Transition Écologique, montre que **l'environnement représente la deuxième préoccupation la plus importante pour les Français** (12%), derrière l'emploi (22%) et devant la sécurité des biens et des personnes (7%).¹¹⁷

La crise sanitaire et économique liée à la Covid-19 aurait pu faire passer ce sujet au second plan. Mais c'est tout le contraire : **l'opinion comprend que le dérèglement climatique peut exacerber la prolifération de nouveaux virus**. Et en effet... la **fonte progressive du permafrost**, un sol gelé sur près d'un quart de l'hémisphère nord-est en Alaska, en Sibérie et dans l'Arctique, risque à la fois de libérer d'énormes quantités de carbone (jusqu'à 1400 gigatonnes), mais aussi des bactéries et autres virus jusqu'alors enfermés dans la glace. Son dégel va largement amplifier les émissions de gaz à effet de serre, et il représente donc une menace sanitaire à l'échelle mondiale.

Les modes de consommation évoluent en conséquence. Déjà, face à la fermetures des frontières et aux pénuries liées à la dépendance vis-à-vis de producteurs étrangers (biens de consommation, masques...), les populations ont eu tendance à se replier vers un approvisionnement plus local, valorisant les circuits courts plutôt que l'importation :

¹¹⁷ <https://www.ademe.fr/representations-sociales-changement-climatique-21-eme-vague>

Figure 2. La préférence des consommateurs envers le local est particulièrement marquée en France



Source : Capgemini Research Institute, Sustainability in Consumer Products and Retail Survey, March 2020, N=7,520 consumers.

En matérialisant une menace concrète, capable de désorganiser les États et de faire trembler l'économie mondiale, **la Covid-19 a donc exacerbé l'enjeu environnemental qui, déjà avant elle, animait de façon croissante les consommateurs.**

Cette prise de conscience écologique est poussée depuis de nombreuses années par des ONG telles que Greenpeace, dont l'activisme parfois radical a eu du mal à faire rallier l'opinion. C'est plutôt le monde des startups, porté par une génération Millenials en quête de sens, qui a su trouver la façon de **démocratiser le « mieux consommer »**.

Nombre de projets comme Yuka ou TooGoodToGo dans l'alimentaire, Clear Fashion ou Fairly Made dans la mode, Spareka pour la réparabilité de l'électroménager, etc. **aident les consommateurs et les industriels à adopter des démarches plus responsables.** Les mesures d'impact se font à diverses échelles : santé, environnement, réemploi, etc.



CLEAR FASHION

Application d'aide au consommateur dans le choix des marques et vêtements les plus éco-responsables.



Solutions à destination des industriels de la mode au service du sourcing éco-responsable, de la transparence et de la mesure d'impact.



Application qui permet de scanner les produits alimentaires et d'obtenir une information sur leur impact santé.

De grandes marques leur ont bien sûr emboîté le pas, à l'image par exemple de La Redoute de La Redoute pour la seconde main vestimentaire, ou les magasins dédiés à l'occasion ouverts par Ikea ¹¹⁸ ou encore E.Leclerc ¹¹⁹.

L'initiative de Time for The Planet est également intéressante : ce fonds à but non lucratif a pour ambition de **mettre l'entrepreneuriat au service de l'urgence climatique**, en mobilisant 1 milliard d'euros pour **créer 100 entreprises à même de lutter efficacement contre le réchauffement climatique**. Bien sûr, si ces entreprises ont vocation à être « more than profit », elles doivent rester des startups à même de perdurer, et donc de générer du chiffre d'affaires. Leur enjeu est donc de **parvenir à réconcilier écologie et économie à travers des modèles économiques rentables**. Créée en avril 2019, cette organisation propose aux particuliers et aux entreprises de devenir actionnaires dès 1€. Elle a déjà réussi à mobiliser plus de 20 000 actionnaires pour 4M€ levés.



Enfin, le cadre réglementaire cherche aussi à imprimer une démarche de durabilité :

- **à l'échelle des Nations à travers l'Accord de Paris sur le Climat**, adopté par 195 pays lors de la COP21 (2015) puis ratifié par 188 pays. Il vise à réduire les émissions de gaz à effet de serre pour **limiter le réchauffement climatique sous la barre des 2°C**

¹¹⁸ <https://newsroom.inter.ikea.com/news/the-world-s-first-second-hand-ikea-store-opens-in-sweden/s/b1aa5e3d-a9e8-4816-828d-72af9b914106>

¹¹⁹ <https://www.capital.fr/entreprises-marches/leclerc-occasion-prevoit-100-ouvertures-en-2020-1358297>

par rapport aux niveaux préindustriels. Pourtant, **ces promesses ne sont pas tenues pour l'instant**, les émissions ayant atteint un niveau historique en 2019, avec 59 Mds de tonnes de CO₂, soit **+5% par rapport à 2015 où a été adopté ce tout 1^{er} traité international pour le Climat**. Le *Green New Deal* européen et américain vise quant à lui à fixer des objectifs continentaux d'ici 2030, sur le chemin de l'Accord de Paris.

- **à l'échelle nationale à travers la Convention Citoyenne pour le Climat** : 150 citoyens français tirés au sort ont constitué fin 2019 une assemblée, fixant un objectif de 40% de réduction des émissions de GES d'ici 2030, à travers 149 mesures dont 75 ont été mises en œuvre début 2021 avec la « Loi Climat ». Mais là encore, certains dénoncent une loi qui dénature la majorité des propositions de la Convention Citoyenne. La **loi AGECE** (Anti-Gaspillage pour une Economie Circulaire), quant à elle, avait déjà lancé 1 an plus tôt la limitation puis l'interdiction de la vente des produits en plastique à usage unique (gobelets, assiettes, pailles...) d'ici 2040.

La vague d'éco-responsabilité devient donc **un nouvel impératif pour les marques**, qui doivent s'en emparer pour rester en phase avec les attentes de leurs clients.

LES MARQUES S'ADAPTENT À CE NOUVEAU PARADIGME

Traditionnellement, le marketing est à l'opposé d'une démarche éco-responsable, puisqu'il pousse à consommer toujours plus et à créer des besoins là où il n'y en a pas forcément. Il est le **moteur de l'achat impulsif, et parfois superflu**.

Mais les marques doivent faire évoluer leurs pratiques à l'ère des « consomm'acteurs », qui ont bien compris que leurs choix pouvaient avoir un impact sur le tissu économique local, et l'état de notre planète.

Selon une étude d'avril 2021 du cabinet de sondage Appinio et de Circul'R, spécialiste de l'économie circulaire, **47% des Français interrogés vérifient les engagements d'une**

marque avant de l'acheter. Près de 49% sont sensibles aux matériaux de l'emballage, et 48% au pays d'origine du produit. ¹²⁰ « Au lieu d'être incités à consommer davantage, les Français attendent des marques, entreprises et distributeurs qu'ils les guident dans leurs choix pour consommer moins et mieux », indique pour sa part l'ADEME suite à la sortie de son Baromètre 2021 de la consommation responsable. ¹²¹

Le prix reste cependant, aujourd'hui encore, le principal facteur de décision. 75% de ceux qui ont acheté des produits reconditionnés ou recyclés l'ont fait en raison de leur budget, plus que pour la cause environnementale. ¹²² En outre, les prix souvent plus élevés des produits soucieux de leur impact restent **un frein pour 63% des Français.** À titre d'exemple, Poule House, qui œuvre au bien-être animal en évitant que les poules pondeuses soient abattues à l'âge de 18 mois du fait de leur baisse de productivité, commercialise sa boîte de 6 œufs à 3,99€, et jusqu'à 5,99€ pour ses œufs bio. C'est deux fois plus que le prix moyen. **Si les clients sont prêts à payer plus cher un produit éco-responsable, l'adoption passera aussi par l'accessibilité financière.** Le prix pouvant être réduit en diminuant par exemple les coûts de production (par le recyclage) et/ou de distribution (en privilégiant les circuits courts).



¹²⁰ <https://www.e-marketing.fr/Thematique/insights-1092/Breves/Seulement-tiers-Fran-ais-pensent-que-engagement-marques-est-sincere-359711.htm>

¹²¹ <https://presse.ademe.fr/2021/05/14eme-barometre-de-la-consommation-responsable-2021.html>

¹²² Etude KPMG x FEVAD « E-commerce et RSE : The Green Deal », 2020 : <https://www.fevad.com/etude-kpmg-x-fevad-e-commerce-et-rse-the-green-deal/>

Toujours est-il que l'engagement des marques devient un véritable argument de vente pour un nombre grandissant des consommateurs. Ce qui les pousse donc à passer d'un marketing traditionnel à un marketing responsable, opérant les changements résumés ci-dessous par l'ADEME ¹²³ :

Marketing conventionnel		Marketing responsable
<ul style="list-style-type: none"> • Pousser à la surconsommation • Inventer des problèmes • Répondre à la demande 		<ul style="list-style-type: none"> • Challenger le consommateur • Répondre aux enjeux de la société • Développer la demande pour une offre plus responsable
<ul style="list-style-type: none"> • Mauvaise image, méfiance • Marketing vert en réaction • Créer de la valeur pour la marque 		<ul style="list-style-type: none"> • Confiance restaurée, image renouvelée • Marketing proactif • Créer de la valeur pour la marque et l'ensemble des parties prenantes
<ul style="list-style-type: none"> • Les produits/services verts restent une niche (pour les écolos-bobos) • Le développement durable : une mode 		<ul style="list-style-type: none"> • Ils concernent de plus en plus de consommateurs (mainstream) • C'est une réalité qui va se renforcer
<ul style="list-style-type: none"> • Les produits/services verts coûtent beaucoup plus chers • Ils sont forcément moins efficaces 		<ul style="list-style-type: none"> • Il existe des moyens de rembourser les innovations • Ils sont de même qualité ou supérieure
<ul style="list-style-type: none"> • Le développement durable est un sujet anxieux • C'est une contrainte 		<ul style="list-style-type: none"> • C'est vivre mieux : santé, bien-être, environnement, partage • C'est une opportunité

© Denis Lefebvre, inspiré par l'Ademe (Étude exploratoire sur le marketing responsable, Septembre 2013)

¹²³ https://presse.ademe.fr/wp-content/uploads/2013/11/ademe_etude-marketing-responsable_rapport.pdf

Plus spécifiquement, **le marketing éco-responsable recouvre deux axes**, qui pèseront d'autant plus qu'ils seront combinés :

- **Une conception éco-responsable des produits ou services**
 - Fabrication limitant l'impact industriel sur l'environnement, tout en assurant un véritable bénéfice pour le client : matériaux plus résistants et solides, meilleure qualité, meilleure durabilité... ;
 - Packaging privilégiant des matériaux biodégradables.
- **Une communication « verte »**
 - Campagnes sur les engagements de la marque (traitement des déchets...);
 - Supports de communication en phase avec les enjeux liés au développement durable : papier recyclé, canaux de diffusion et de vente adaptés, numérique responsable (p.205), etc.

In fine, entre marketing traditionnel et marketing (éco-)responsable, **l'enjeu reste le même : améliorer son image de marque, et ainsi augmenter ses ventes**. Mais les entreprises sont-elles sincères dans leur démarche éco-responsable ?

1.2. Coup de com ou véritable engagement ?

Selon l'étude « Le développement durable, facteur déterminant de préférence d'achat » de Cap Gemini ¹²⁴, **la moitié des consommateurs français se montrent suspicieux quant aux prétentions éco-responsables des marques**. 58% déclarent ne pas pouvoir s'assurer du caractère vraiment durable de leurs produits, mettant en avant un problème de traçabilité et de transparence. Le sondage de Circul'R et Appinio ajoute que **seulement 35% des Français pensent que les actions des marques pour l'environnement sont sincères**. Ils sont pourtant 80% à estimer que leur engagement est nécessaire.

¹²⁴ https://www.capgemini.com/fr-fr/wp-content/uploads/sites/2/2021/03/Synthèse_Sustainability-CPR_FR_Def1.pdf

LES DANGERS DU GREENWASHING

Le « greenwashing », ou « écoblanchiment », est un **procédé marketing** qui consiste à créer et à diffuser « tout message publicitaire pouvant induire le public en erreur sur la qualité écologique réelle d'un produit ou d'un service, ou sur la réalité de la démarche de développement durable d'une organisation. » ¹²⁵

En d'autres termes, **le greenwashing renvoie une image trompeuse de responsabilité écologique**. En voici 4 illustrations :



- **Air France** a annoncé en 2019 vouloir compenser 100% des émissions CO₂ sur ses vols intérieurs, en plantant 70 millions d'arbres pour capturer l'équivalent des 4,8 millions de tonnes de CO₂ émises chaque année par le trafic aérien domestique. Air France veut se donner une image « verte », comme l'illustre la publicité ci-dessus, en optant pour la **compensation sans chercher à diminuer ses propres émissions**. Car

¹²⁵ <https://librairie.ademe.fr/consommer-autrement/1757-guide-de-la-communication-responsable-9791029709661.html>

planter des arbres pour compenser, c'est bien... mais non seulement les quantités ne sont pas suffisantes, le rythme de croissance des arbres ne permet de capter le CO2 que bien plus tard, en décalage avec le réchauffement climatique, et surtout ce sont des actions à la source qui sont attendues de Air France pour limiter son empreinte carbone. ¹²⁶ Greenpeace a d'ailleurs partiellement repeint en vert l'un des avions de la compagnie, en mars 2021, pour dénoncer ses pratiques de greenwashing. ¹²⁷

- **Volkswagen** vante la qualité écologique de ses véhicules, comme dans la publicité ci-dessus de 2007, en détournant le « Roulez propre » et en l'associant à l'image de « Monsieur Propre »... alors que la marque sera **épinglée en 2015 sur le scandale lié à la dissimulation de ses émissions polluantes** (jusqu'à 40 fois la norme autorisée) de 11 millions de ses voitures diesel.
- **Le Chat (Henkel)** en 2009 lave sa publicité ci-dessus « plus vert que vert » en utilisant abusivement le terme « écologie », et en **se dotant de ce qui ressemble à un label, « Eco Efficacité », qui n'en est en fait pas un...** 5 ans plus tard, la marque rectifie le tir après l'obtention d'un vrai label écologique européen, l'Ecolabel.
- **Mc Donald's** passe en 2009 le fonds de son **logo du rouge au vert** dans les pays européens où il est installé, et où la conscience écologique grandit. Si la chaîne de fast-food a effectivement réduit ses émissions de GES de 35% depuis 2005, cet engagement cache le fait qu'elle est le **1^{er} consommateur mondial de bœuf, l'une des viandes ayant le plus fort impact carbone** (35,8 kg, contre 6 kg pour le poulet).

Pourtant, **le risque du greenwashing est double : non seulement il alimente la méfiance des consommateurs, mais il freine également les actions réelles et concrètes d'autres entreprises** qui pourraient être mises dans le même bain.

¹²⁶ <https://reporterre.net/Compensation-carbone-le-gouvernement-rend-obligatoire-une-fausse-solution-au-lieu-de-reguler-le-trafic-aerien>

¹²⁷ <https://www.greenpeace.fr/action-lavion-vert-ne-sauvera-pas-le-climat/>

La loi Climat et Résilience de mai 2021 vise entre autres à fixer un **cadre légal pour limiter le greenwashing**, et à le reconnaître comme pratique publicitaire trompeuse. Parler de « durabilité » en bariolant son message de vert ne veut pas dire grand chose. Cette approche neuromarketing vise surtout à déclencher un stimuli émotionnel.

Pour **prouver la sincérité de leur démarche RSE**, les marques doivent non seulement se transformer de l'intérieur, mais aussi être capable de démontrer les résultats des actions mises en œuvre. **Cet effort de transparence est la condition sine qua non de l'adhésion des consommateurs**, et d'un **passage du « greenwashing » au « green marketing »**. Cela doit se matérialiser par une **mesure concrète et certifiée de leur impact**, et l'**obtention de labels reconnus** (comme l'Eco-Score ou le Label Numérique Responsable, que nous reverrons par la suite).



*L'accélération de ces changements s'est étendue aux régulateurs qui s'efforcent de **rendre plus transparent l'impact écologique de notre consommation**.*



*Tout comme les qualités nutritionnelles, qui sont évaluées avec un nutri-score, un « **score carbone** » devient la **nouvelle norme**, informant les consommateurs de l'impact environnemental des produits qu'ils achètent, et les aidant dans leur comportement d'achat.*

Vincent de Montalivet Sustainable AI Leader chez Capgemini

Au-delà des labels, **la SNCF fait le pari de la mesure de cet impact par les usagers eux-mêmes**. Son **comparateur de mobilité** permet d'évaluer l'empreinte carbone du train vis-à-vis des autres modes de transport, en laissant l'utilisateur entrer sa provenance et sa destination. Il peut ainsi **constater de lui-même que le train est plus « vert » que les autres**, chiffrage CO2 à l'appui. Résultat ? Les usagers habituels ont le sentiment de faire un geste pour la planète, et de nouveaux usagers sont sensibilisés à leur tour, avec pour objectif de leur faire privilégier le train. En outre, chaque réservation indique l'équivalent CO2 du voyage à venir, renforçant la démocratisation de la mesure d'impact.

Voyagez en train et limitez votre empreinte **carbone**



Notre comparateur de mobilité pour voyager responsable



Découvrez notre **nouvel outil** pour vous aider à préparer vos trajets. Comparez vos futurs trajets en un coup d'œil selon différents critères de choix. Vous en saurez plus sur votre **empreinte environnementale**, la **durée** et le **temps utile** de vos voyages.

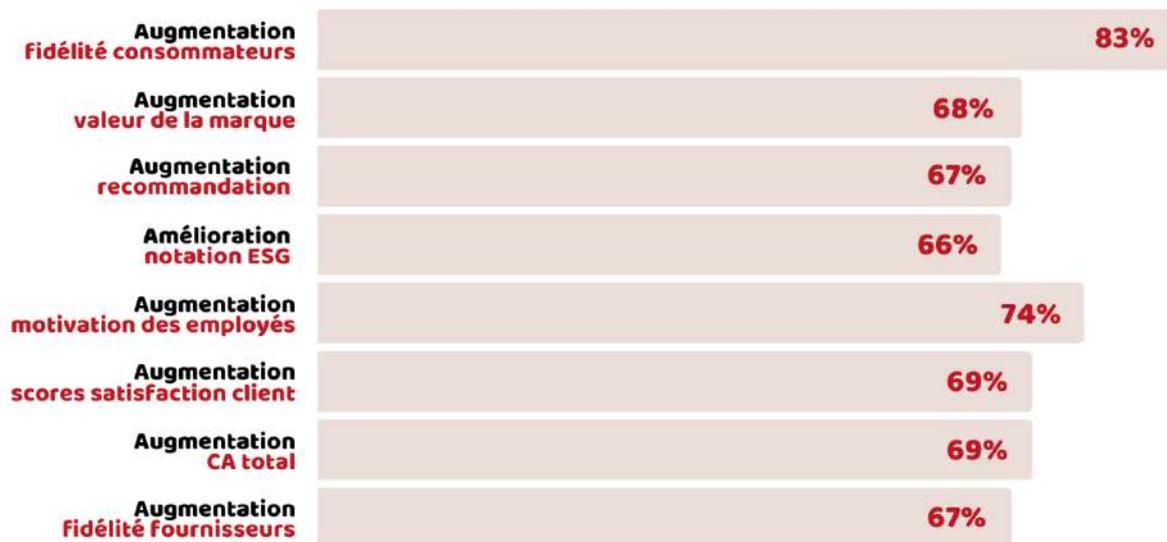
Reste à savoir si cette notion de « CO2 » reste suffisamment claire et parlante pour les **utilisateurs**, qui n'ont naturellement pas idée de ce que peut représenter 1 kg de CO2, ou de « l'équivalent CO2 » créé par le GIEC pour comparer et cumuler les impacts des divers gaz à effet de serre en terme de réchauffement climatique.

Dans certains secteurs, et certainement en B2B, « le carbone est devenu une devise [...] un KPI de plus en plus incontournable », selon Gwendal Bihan, CEO de Axionable, dans l'interview qu'il m'a accordé (p.234).

Mais, surtout en B2C, **trouver des équivalences doit faire partie de l'effort marketing afin que ces mesures soient comprises... et *in fine* impactantes.** En ce sens, dire qu'un trajet Paris-Caen a un impact écologique 16 fois moins important en train qu'en voiture, est sans doute plus efficace pour le grand public que de comparer des « kg CO2e », même s'il est bien sûr entendu que **la mesure scientifique doit, elle, rester rigoureuse, et établie quelque part derrière le message marketing.** Sans quoi les marques ne rempliraient pas leur devoir de transparence et de traçabilité.

LES BÉNÉFICES D'UN MARKETING RESPONSABLE SINCÈRE ET TRANSPARENT

Selon l'étude de Capgemini précédemment évoquée ¹²⁸, **les marques engagées de façon sincère dans une démarche éco-responsable en tirent de nombreux bénéfices :**



© Denis Lafont, inspiré par Capgemini Research Institute
"Sustainability in Consumer Products and Retail Survey", April-May 2020

Parmi les principaux bénéfices constatés, nous pouvons donc souligner :

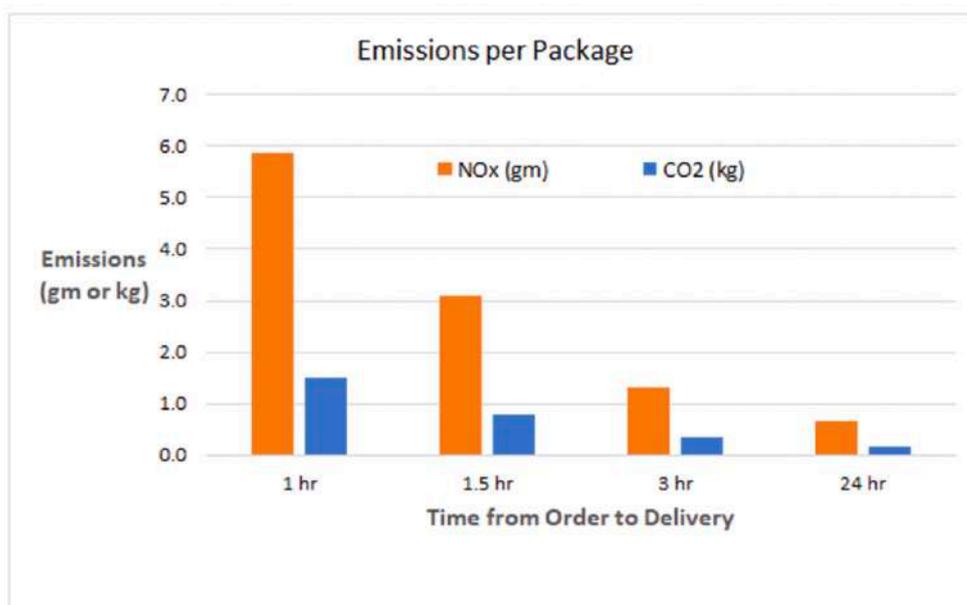
- Un renforcement de la **fidélité** ;
- Une augmentation des **ventes** et de la **satisfaction** ressentie en achetant des produits éco-responsables ;

¹²⁸ https://www.capgemini.com/fr-fr/wp-content/uploads/sites/2/2021/03/Synthèse_Sustainability-CPR_FR_Def1.pdf

- Une augmentation de la **notoriété** et du **referral** en véhiculant une image de marque impliquée et positive ;
- Une mobilisation des collaborateurs ainsi que des fournisseurs, renforçant la **culture d'entreprise** autour d'un projet porteur de sens.

Alors que **seulement 17% des dirigeants français pensent que leurs clients pourraient préférer des marques moins connues mais perçues comme plus responsables**, 60% des consommateurs ont déjà franchi le pas.

Cela démontre bien l'intérêt de questionner les attentes réelles de ses clients. Dans sa chronique « *Livraison le jour même, est-ce réellement ce que veulent vos clients ?* »¹²⁹, Frédéric Mirebeau **interroge la pertinence de la livraison en moins de 24h**. Car si cette promesse commerciale forte s'associe à une prouesse logistique, l'empreinte carbone s'en trouve démultipliée, comme le souligne ce graphique extrait du rapport « *Pour un développement durable du commerce en ligne* »¹³⁰. En effet, **livrer des colis isolés sans les regrouper accélère logiquement leurs émissions**.



Source : Institut for Transportation Studies, 2020, mission.

¹²⁹ <https://www.journaldunet.com/ebusiness/commerce/1499929-livraison-le-jour-meme-est-ce-reellement-ce-que-veulent-vos-clients/>

¹³⁰ https://www.economie.gouv.fr/files/files/PDF/2021/Rapport_Devt_durable_commerce_en_ligne.pdf

Si les Français appellent de leurs vœux des livraisons plus rapides, ils sont aussi près de **60% à vouloir des livraisons e-commerce plus écologiques**.¹³¹ 71% de cyber-acheteurs regroupent leurs commandes afin de limiter les livraisons, et 66% favorisent la livraison en points relais.¹³² Dès lors, **ne pensez-vous pas que d'autres seraient eux aussi prêts à adapter leurs habitudes d'achat s'ils avaient connaissance de la différence de l'impact d'un colis entre 1h et 2 jours ?**

Laisser le choix au consommateur en l'informant est très probablement la voie à suivre pour sensibiliser sans contraindre. En améliorant la livraison **non plus sur le critère de la rapidité mais sur celui de la précision**, avec des tranches horaires plus resserrées et plus fiables, ou d'autres alternatives proposées pour récupérer son colis, **vous agirez pour le bien de la planète tout en améliorant drastiquement votre relation client.**

En somme, **plus de la moitié des consommateurs français font évoluer leurs habitudes d'achat pour les aligner avec leurs préoccupations environnementales**. Les marques qui s'engagent dans cette voie, animés d'une motivation sincère, d'engagements concrets et d'une communication transparente, créent un **levier émotionnel puissant, susceptible de devenir un véritable atout compétitif**.

Pour accélérer leur transformation éco-responsable, et comme nous allons le voir dans la partie suivante, elles peuvent **compter sur l'IA pour optimiser chaque étape de la vie du produit, et notamment son marketing**.

Mais avant de parler de l'IA marketing au cœur du numérique responsable, voici **mon interview de Maurice Ndiaye, co-Président de l'Adetem** (association des professionnels du marketing), avec qui j'ai pu échanger sur la question du marketing responsable.

¹³¹ <https://journaldunet.digidip.net/visit?url=https://stuart.com/fr/blog/entrepots-urbains/&currurl=https://www.journaldunet.com/ebusiness/commerce/1499929-livraison-le-jour-meme-est-ce-reellement-ce-que-veulent-vos-clients/>

¹³² Etude KPMG x FEVAD « E-commerce et RSE : The Green Deal », 2020 : <https://www.fevad.com/etude-kpmg-x-fevad-e-commerce-et-rse-the-green-deal/>



Maurice Ndiaye
Co-Président de l'Adetem



L'Adetem, l'association des professionnels du marketing, vient de lancer le collectif « Responsables ! » pour s'emparer de la question du marketing responsable : pouvez-vous nous en dire plus ?

« Les deux constats qu'on fait, c'est qu'il y a un manque de définition, et un manque de formation. Notre objectif, c'est de venir essayer de résoudre ces deux problèmes, en apportant un peu plus de lisibilité sur ce que l'on entend par « marketing responsable ». Parce qu'on a tendance à le mélanger avec la dimension environnementale, qui est forte mais qui n'est pas la seule : parmi les 17 ODD de l'ONU il y a l'éthique, l'inclusion, etc.

Donc il y a un premier travail de définition. Et puis il y a une deuxième lame consistant à accompagner les professionnels, qui se sentent un peu perdus pour plein de raisons. Ne pas savoir précisément de quoi on parle, déjà. Pas mal d'entre eux sont aussi pris dans des modèles d'organisation qui, en fait, ne sont pas conçus nativement pour pouvoir laisser la place à ce genre de démarche. Ça rend très difficile le fait de se former tout en modifiant des chaînes de production complexes, qui engagent le sourcing des produits, le pricing, l'engagement des collaborateurs, la distribution etc. Refondre tout cela dans son intégralité, c'est compliqué. Enfin, les éléments d'évaluation ne sont pas clairs. On se fait rattraper par des indicateurs et mesures historiques de performance qui ne valorisent pas du tout les démarches responsables.

On se retrouve donc pris entre deux feux, entre d'un côté une intention, une ambition affichée assez noble du « il faut qu'on bouge », et de l'autre une réalité opérationnelle

RESPONSABLES!

by Adetem

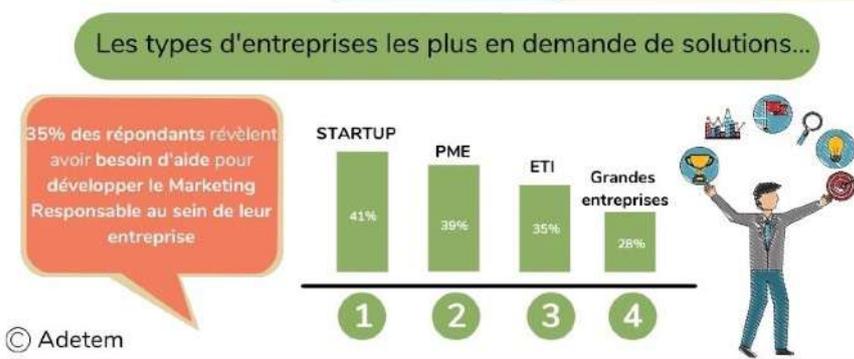
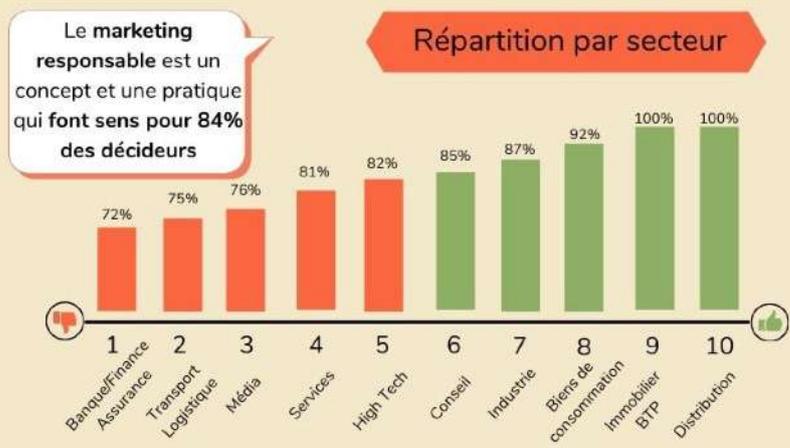
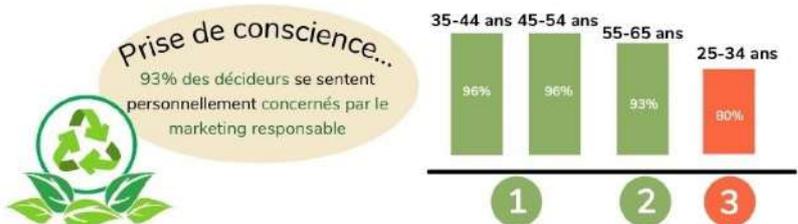
sur ce qui compte vraiment dans l'évaluation de la performance des entreprises et des salariés qui, en fait, ne met pas à l'honneur ou très peu les dimensions du responsable.

Comment s'inscrire dans une démarche éco-responsable ?

La première chose, c'est déjà d'**avoir une prise de conscience qu'on modifie les usages**. Et pour cela il faut éduquer. La majorité des gens n'ont pas conscience, par exemple, du coût environnemental considérable des emails inutiles ou des pièces jointes. Quand je fais l'expérience chez moi d'une ampoule restée allumée qui va me faire payer plus cher ma facture d'électricité, oui là je comprends et je me sens concerné. Donc il faut pouvoir faire comprendre facilement l'impact d'un changement, et **identifier par où commencer**.

La mesure de l'impact environnemental va-t-elle devenir pour les marques un argument marketing de plus en plus utilisé ?

Certaines marques s'en sont en effet déjà largement emparés. Malheureusement, il y a toujours assez vite des dérives, et dans les détails de ce genre de calculs, on se rend compte qu'il y a plein de zones d'ombre dans lesquelles on peut se faufiler. Par ailleurs, quand on calcule des logiques de neutralisation carbone, on est sur des échelles de temps qui ne sont pas du tout les mêmes. Si en tant qu'entreprise je fais une émission aujourd'hui et que j'essaie de le compenser dans le futur, je vais brûler plus vite que ce que je ne compense, et ça met en place une sorte de fuite en avant. C'est un peu le revers de cette métrique de « compensation carbone » qui consiste à planter des arbres ou autre a posteriori, mais qui ne prend pas le problème à la racine. Après, elle a quand même l'avantage d'être **très facile à comprendre, et pas si difficile que cela à calculer.** »



Cette infographie est le résultat d'une enquête du collectif RESPONSABLES ! de l'Adetem, en partenariat avec Stratégies, menée auprès de 286 professionnels du marketing en mars 2021.



- Principaux résultats :**
- 8 marketers /10 estiment que le marketing responsable fait sens, et 93% se sentent concernés.
 - 59% des décideurs révèlent que le marketing responsable est déjà une réalité dans leur entreprise.
 - 1/3 d'entre eux ont besoin d'aide pour avancer dans la mise en place pratique.

2. L'IA marketing au cœur du numérique responsable

Nous l'avons vu, **adopter une démarche éco-responsable consiste à être cohérent avec ses engagements jusque dans sa façon de communiquer.**

Monoprix a ainsi annoncé en 2019 l'arrêt définitif de ses catalogues papier, pourtant l'un de ses supports de communication phares, générant 10 à 15% de son apport de trafic en magasins.¹³³ Mais 30 millions de catalogues en moins, c'est 2400 tonnes épargnées de papier, non recyclables en totalité, alors que 90% sont jetés sans jamais avoir été lus. Le **catalogue digital** lui a été préféré, disponible sur son site et son application, pour un **impact écologique 14 fois inférieur**. D'autres enseignes lui ont emboîté le pas afin de réduire les 20 milliards de prospectus et flyers distribués chaque année en France.

Mais il faut **aussi prendre garde à l'empreinte carbone des supports et des leviers du marketing digital lui-même**, surtout à l'heure où la pandémie de Covid-19 a largement accéléré la transformation digitale. **L'IA et l'IA marketing vont vous y aider.**

2.1. Accélérer son éco-responsabilité grâce à l'IA

Au-delà d'être un outil au service de la protection de l'environnement et de la transition énergétique, **l'IA sert les ambitions du marketing éco-responsable grâce à ses pouvoirs de prédiction et de ciblage**. Ils permettent de **viser plus juste** et plus efficacement, pour **une plus grande frugalité des coûts économiques et écologiques de ses campagnes**.

En outre, en pouvant traiter des volumes de données gigantesques, **l'IA sert l'objectif de traçabilité et de transparence** recherchées par les consommateurs pour adhérer aux engagements durables des marques, et à ces dernières dans leur stratégie de notoriété.

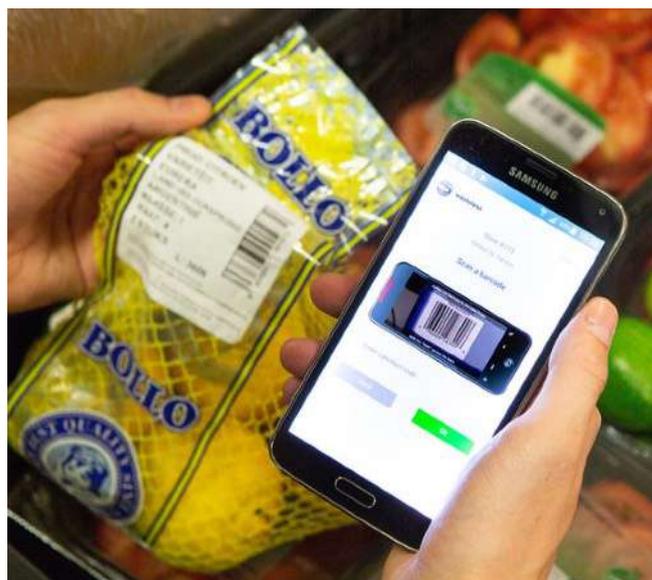
¹³³ <https://www.lesechos.fr/industrie-services/conso-distribution/0600491899323-regis-schultz-president-de-monoprix-je-ne-crois-pas-une-minute-aux-points-relais-leclerc-2235340.php>

L'IA POUR OPTIMISER LE CYCLE DE VIE DU PRODUIT ET SON MARKETING

Selon l'étude Capgemini *Climate AI*, depuis 2017 l'Intelligence Artificielle aurait déjà permis de réduire de 13% les émissions de GES liées aux produits de consommation et au commerce de détail. D'ici à 2030, grâce à l'IA, les marques pourraient atteindre 45% des objectifs de réduction carbone fixés par l'Accord de Paris. ¹³⁴

L'IA est en effet capable d'optimiser toutes les étapes du cycle de vie des produits, et de les aider à atteindre leurs objectifs de durabilité :

- **Conception** : en anticipant les **tendances émergentes et donc la demande** par du **social listening à grande échelle**, comme nous l'avons vu dans la 1^{ère} partie (p.70), permettant aussi d'identifier les réactions face à leurs produits, à leurs engagements et à leur positionnement, mais aussi à ceux de leurs concurrents.
- **Fabrication** : en améliorant la **durabilité des produits** (résistance...), et en limitant le nombre d'invendus, donc la production inutile, par la **prédiction des ventes**. Dans l'agro-alimentaire, l'IA peut ainsi remonter les données à même d'éviter le gaspillage alimentaire, grâce à l'imagerie visuelle identifiant l'état des fruits et légumes sur les étals, ou d'autres produits atteignant la date de péremption. Le détaillant peut alors les proposer en promotion. La startup américaine Wasteless expérimente son **IA anti-gaspillage** en Pologne auprès du grossiste Metro, avant de le déployer sur d'autres marchés. Résultat ? Une augmentation du CA sur les produits, déjà 40% de gaspillage en moins, et 80% visés avec l'amélioration de sa technologie. ¹³⁵



¹³⁴ <https://www.capgemini.com/wp-content/uploads/2020/12/Report-Climate-AI.pdf>

¹³⁵ <https://www.neoproducts.com/startups/la-start-up-wasteless-lance-sa-solution-anti-gaspi-pour-la-grande-distribution/>

- **Logistique** : pour de la **maintenance prédictive**, la **gestion affûtée des stocks**, et pas seulement : Amazon utilise par exemple des **algorithmes pour déterminer les emballages les plus adaptés aux produits**, limitant leur poids et leur place dans les camions, réduisant les coûts de livraison, leur impact carbone de bout en bout, et jusqu'à la quantité de cartons à recycler. Les algorithmes intègrent également les remontées clients sur les produits sur-emballés, et limitent les produits emballés individuellement **en optimisant l'envoi des commandes**. En 5 ans, Amazon a ainsi réussi à diminuer ses **besoins en emballage de -33%** (-1,6 milliard de cartons).¹³⁶
- **Distribution** : en identifiant en temps réel les trajets qui nécessitent **le moins de consommation de carburant** (embouteillages, etc.), et en optimisant l'**impact du « dernier kilomètre »**. Transmetrics et son IA de planification prédictive a ainsi permis au transporteur Speedy d'augmenter de 14% l'utilisation de sa flotte, grâce à une gestion fine des besoins d'acheminement selon le regroupement des colis par codes postaux, pour une réduction de son coût total de l'ordre de -9%.¹³⁷

Et bien sûr... le **Marketing** !

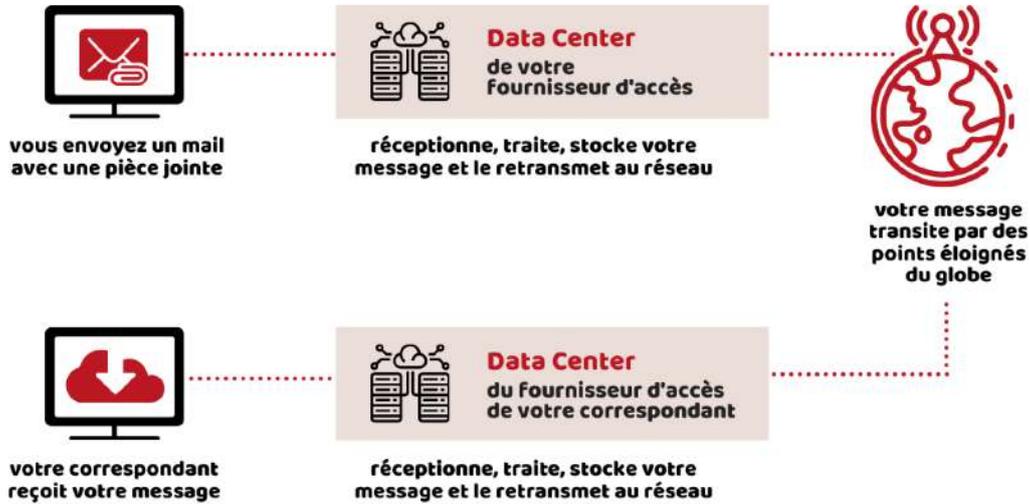
Comme nous l'avons vu dans le chapitre 1 (p.51), **l'analyse prédictive permet d'opérer un ciblage extrêmement précis** pour adresser les bonnes personnes au bon moment, sur le bon canal et avec le bon message. Les marques évitent ainsi d'envoyer leurs messages à l'aveugle tout azimut, en **améliorant leur expérience client, leurs coûts de diffusion... et leur empreinte carbone**.

En optimisant par l'IA la segmentation de votre base CRM, **vous évitez l'envoi inutile de milliers d'emails, qui chacun parcourent en moyenne 15 000 kilomètres pour parvenir jusqu'à leurs destinataires**, en passant par divers points du globe, en fonction de la localisation des Data Centers des fournisseurs d'accès respectifs.

¹³⁶ <https://www.triplepundit.com/story/2020/ai-amazon-sustainable-packaging/707721>

¹³⁷ <https://www.transmetrics.ai/fr/case-study/predictive-linehaul-planning-in-express-parcel-business/#cs1>

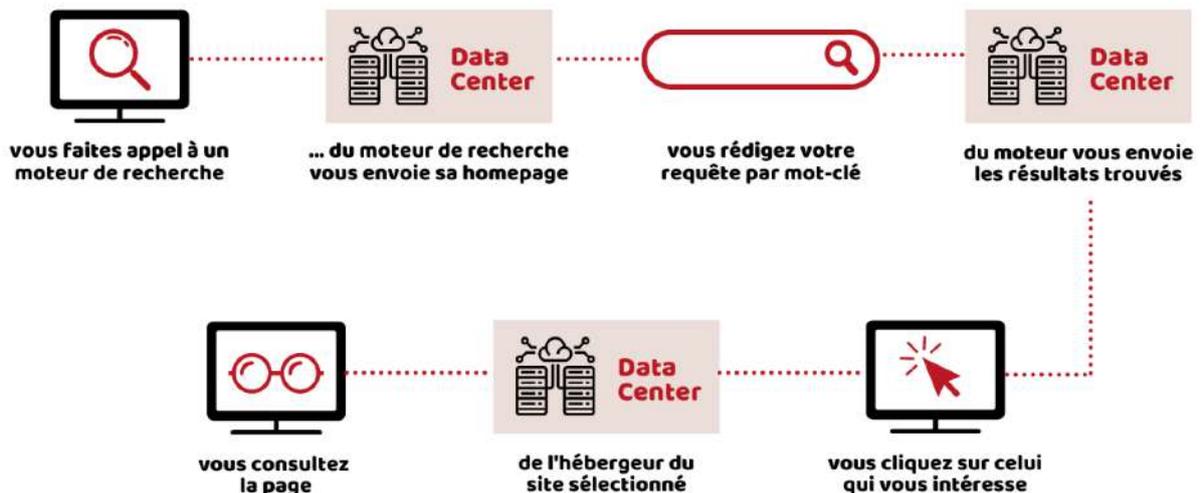
L'envoi d'un mail : comment ça marche ?



Les enjeux sont similaires en terme de **SEA, Display et Social Ads** : grâce à l'IA (p.41), visez mieux pour **augmenter votre ROAS tout en limitant l'empreinte liée à la diffusion de vos campagnes**.

Côté **SEO**, la multiplication des requêtes sur les moteurs de recherche nécessite des **parcours encore plus énergivores que ceux des emails**. Soignez votre référencement grâce à l'IA (p.32) pour que vos contenus atteignent plus rapidement vos cibles.

La requête web : comment ça marche ?



En outre, le **traitement massif d'informations par l'IA** permet de servir les objectifs du marketing éco-responsable, en **apportant la traçabilité et la transparence sollicitée par les consommateurs pour s'assurer de la réelle durabilité des marques**, sous le vernis des messages marketing. Après le nutri-score... place dès lors au « score carbone » !

L'IA, OUTIL MARKETING DE LA TRANSPARENCE : POUR UN « YUKA DU CO2 »

Il y a une dizaine d'années, Tesco, l'une des principales chaînes de supermarchés en Grande-Bretagne, a tenté de mettre en place un système de label mesurant l'impact carbone des produits vendus dans ses magasins. Mais la **complexité de la collecte de toutes les données nécessaires pour évaluer ce score**, à tous les niveaux de la chaîne d'approvisionnement, lui a finalement fait renoncer à ce projet.

Pourtant, en 2020, Unilever, l'un des leaders mondiaux sur les biens de consommation et 4^e acteur du secteur agro-alimentaire (400 marques dont Knorr ou Lipton), annonce vouloir mettre sur tous ses produits des étiquettes d'empreinte carbone.¹³⁸ Marc Engel, Chief Supply Chain Officer de Unilever, expliquait cette démarche :



*Nous constatons clairement que **les consommateurs veulent savoir comment les produits qu'ils achètent contribuent à leur propre empreinte carbone.** [...]*

*Nous mettrons en place un système permettant à nos fournisseurs de déclarer, sur chaque facture, l'empreinte carbone des biens et services fournis ; et nous créerons des partenariats avec d'autres entreprises pour **normaliser la collecte, le partage et la communication de données.***



Marc Engel Chief Supply Chain Officer de Unilever

Unilever est certes mieux placé que Tesco pour obtenir toutes les données nécessaires sur les produits de ses propres marques. Mais **l'IA est l'outil indispensable pour assurer**

¹³⁸ <https://www.unilever.com/news/press-releases/2020/unilever-sets-out-new-actions-to-fight-climate-change-and-protect-and-regenerate-nature-to-preserve-resources-for-future-generations.html>

la performance de leur analyse. En outre, **les Smart Contracts de la Blockchain** vont de plus en plus permettre d'assurer **leur totale traçabilité**. Grâce à eux, **tout le parcours du cycle de vie du produit** peut en effet être **tracé de façon fiable et transparente**. Ainsi, un « score carbone » bien plus précis que celui calculé grâce à l'ACV, l'Analyse de Cycle de Vie (outil traditionnel de mesure des impacts environnementaux des produits/services), peut être établi, **en remontant la piste de chaque fournisseur et de leur propre impact**. Ainsi, comme le souligne Vincent de Montalivet, Sustainable AI Leader chez Capgemini :



En ne valorisant pas les données environnementales pour rendre l'empreinte carbone de ses produits transparente, les entreprises risquent d'être notées sur des canaux non affiliés, au-delà du contrôle de l'entreprise.

*Il appartient donc **aux marques d'utiliser les dernières technologies pour activer les bons leviers et s'affirmer comme des entités résolument éco-responsables.***



Vincent de Montalivet Sustainable AI Leader chez Capgemini

La Blockchain et ses Smart Contracts permettraient donc d'**assurer la traçabilité à même de garantir la fiabilité des critères pris en compte par les labels**, créés pour aider les consommateurs à évaluer l'impact environnemental des produits et services.

Citons notamment, dans un secteur agro-alimentaire responsable à l'échelle mondiale de 28% des émissions de GES, **l'Éco-Score**, créé début 2021 par 9 acteurs dont Yuka, Marmiton ou encore FoodChéri. Il vise à inspirer le label gouvernemental de fin 2021 en s'inspirant lui-même du Nutri-Score, qui évalue pour sa part la valeur nutritionnelle des aliments. L'Éco-Score vise déjà à **compléter l'ACV en s'appuyant sur l'Open Data**, via la base de données collective sur les produits alimentaires Open Food Facts.



En somme, **prévenir le greenwashing passe par un marketing de la transparence, à travers des labels fiables et complets.** Pour ce faire, **s'appuyer sur le potentiel des nouvelles technologies sera gage de la confiance** accordée par les consommateurs envers les marques, et garant de la mise en place de pratiques réellement durables.

Cela inclut bien sûr la **mesure de l'impact de ces nouvelles technologies elles-mêmes**, qui se doivent d'être elles aussi éco-responsables. C'est l'enjeu d'une Blockchain plus écologique. C'est aussi et surtout **l'enjeu même de cette thèse professionnelle sur l'IA marketing durable...**

Mais plus globalement, **où en est-on actuellement du numérique responsable ?**

2.2. Green IT : où en est le numérique responsable ?

La **croissance exponentielle des usages liés au numérique** impose de **fixer un cadre en faveur du développement durable**.

Le « Numérique Responsable », ou « Sustainable IT », vise à sensibiliser les utilisateurs et les organisations pour leur **faire notamment adopter des éco-gestes dans leurs usages numériques**, et ainsi les aider à réduire leur empreinte carbone.

Le site « Qu'est-ce qu'on fait ? » explique de façon didactique comment tout un chacun peut y contribuer, au bureau ou chez soi, dans la gestion de ses emails et équipements, dans ses recherches Internet, etc. Je vous invite à consulter leur [infographie interactive](#) très bien réalisée, afin d'y découvrir **toutes les bonnes pratiques**.

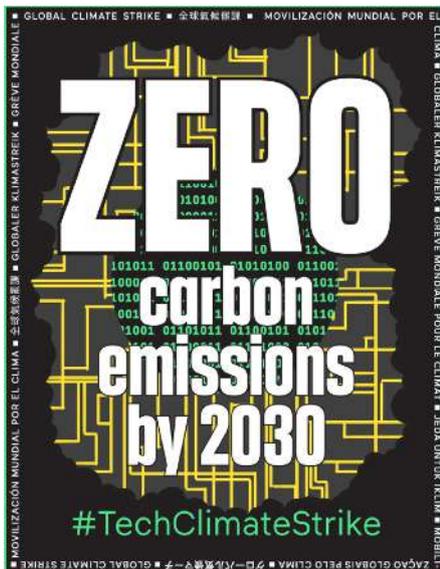
Mais où en sont aujourd'hui les entreprises quant à la mise en place de leur **démarche « Green IT »** ? **Où en sont en particulier les géants de la tech ?**

GAFAM : DES GÉANTS BIENTÔT VERTS ?

En 2019, des milliers d'employés de Amazon, Google, Facebook, Microsoft, Twitter et d'autres grandes entreprises technologiques se sont mobilisés dans le cadre de la grève mondiale contre le changement climatique, dirigée par Greta Thunberg.

Les **marcheurs de la Tech Workers Coalition** réclamaient que leurs employeurs prennent **des engagements clairs et réels contre la crise climatique**, en ramenant leurs émissions de **GES à zéro d'ici 2030**, et en arrêtant tout contrat avec des fournisseurs d'énergie combustible pour passer à **un réseau électrique alimenté à 100% par des énergies renouvelables**.

En outre, ils leur demandent d'être **plus transparents quant aux modalités de mesure de leur empreinte carbone**, pour les inciter à agir tout en informant à la fois leurs clients et leurs utilisateurs, mais aussi les régulateurs.



Les géants de la tech semblent avoir entendu le message, puisqu'ils ont annoncé tour à tour depuis 2019 la mise en place de **nouvelles politiques et initiatives vertes**. Certains des GAFAM étaient engagés dans la réduction de leurs émissions bien avant, à l'instar de Google. Mais **2019 marque un tournant dans l'accélération de leurs engagements** liés au mouvement du « *Tech for Good* ».

Faisons un **tour d'horizon des actions mises en œuvre par les Big Five**. Pour comparer leur degré d'engagement, nous emploierons les termes suivants, couramment utilisés, bien que minimisant l'impact réel des entreprises qui y prétendent. De ce fait, ils sont notamment dans le viseur de la loi Climat et Résilience sur le greenwashing en France :

- Une entreprise « **neutre en carbone** » ou « **net zéro** » se dit d'une entreprise qui à la fois réduit au maximum ses émissions de GES, mais aussi compense l'intégralité des émissions restantes en portant des projets de développement durable ;
- « **néгатif en carbone** » revient pour une entreprise à éliminer plus de carbone qu'elle n'en produit chaque année ;

Google

- Google a annoncé être la **1ère entreprise majeure à atteindre la neutralité carbone** depuis 2007, et avoir réussi depuis 2020 à **compenser toutes les émissions de GES générées depuis sa création.**

- Google vise une **alimentation de ses Data Centers et de ses campus à 100% en énergie sans carbone d'ici à 2030.**

- Comme nous l'avons vu, Google utilise aussi **l'IA via sa division DeepMind**, notamment **pour minimiser l'utilisation d'énergie et refroidir ses serveurs**, en parvenant ainsi à **réduire les besoins énergétiques de 35% environ.** Elle déplace également l'exécution des charges de travail

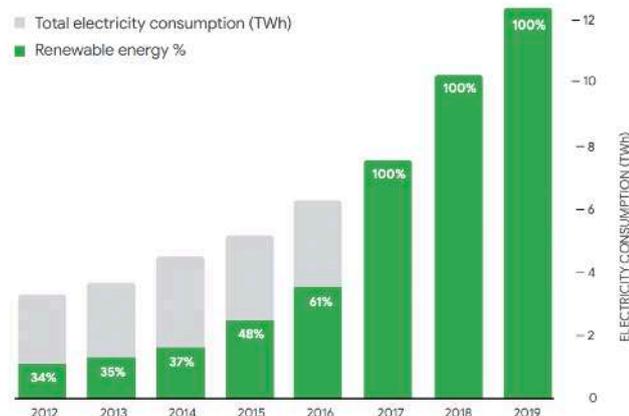
non urgentes dans ses Data Centers vers les moments où les sources d'énergie à faible teneur en carbone sont abondantes.

- Google a annoncé mettre 5,75 milliards de dollars sur la table pour financer des projets écologiquement et socialement responsables.
- Retrouvez tous les engagements durables de Google : <https://sustainability.google/commitments/>



Carbon neutral since 2007.
Carbon free by 2030.

Renewable energy purchasing compared with total electricity use (2012–2019)



Amazon

- L'exigence de transparence de la *Tech Workers Coalition* a **poussé Amazon à révéler pour la première fois son empreinte carbone.** En 2019, ce n'est donc pas moins de

51,17 millions de tonnes métriques de CO2 qui ont été émis par Amazon, le plaçant **parmi les 150 à 200 premiers émetteurs dans le monde.**

- Amazon a annoncé vouloir être **neutre en carbone d'ici 2040**, en alimentant ses Data Centers à **100% d'énergie renouvelable d'ici 2025**, et en convertissant **100% de sa flotte de livraison aux véhicules électriques.**
- Son objectif est d'atteindre des **expéditions sans carbone** en privilégiant processus durables et matériaux recyclés.
- Tous les engagements durables de Amazon : <https://sustainability.aboutamazon.com>



Amazon's Enterprise-Wide Carbon Footprint for the 2019 Fiscal Year

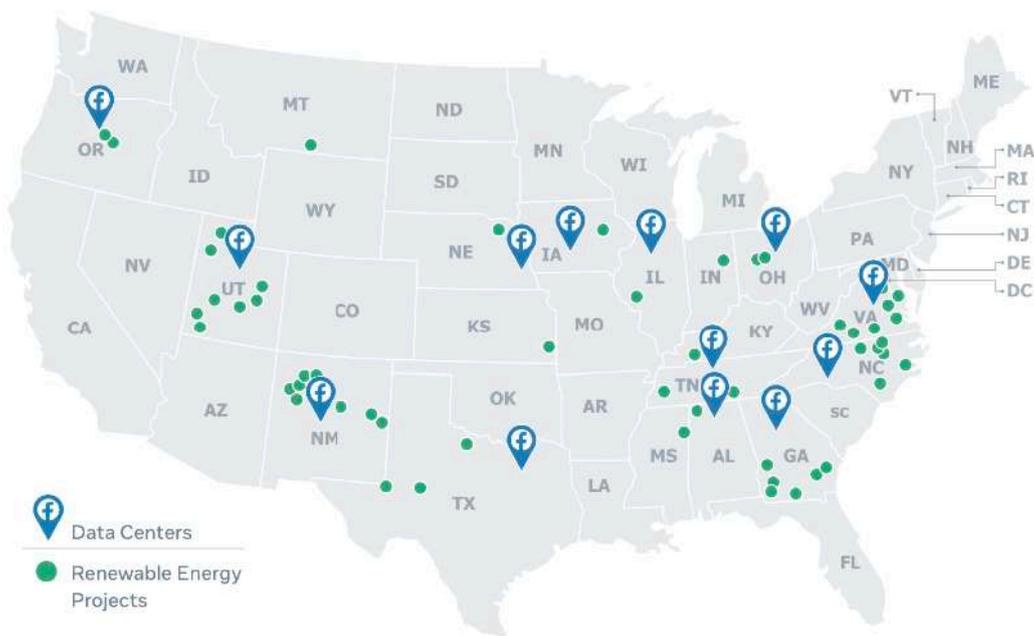
Category	MMT CO ₂ e
Emissions from Direct Operations (Scope 1)	5.76
Fossil fuels	5.57
Refrigerants	0.19
Emissions from Purchased Electricity (Scope 2)	5.50
Emissions from Indirect Sources (Scope 3)	39.91
Corporate purchases and Amazon-branded product emissions (e.g., operating expenses, business travel, and Amazon-branded product manufacturing, use phase, and end-of-life)	15.41
Capital goods (e.g. building construction, servers and other hardware, equipment, vehicles)	8.01
Other indirect emissions (e.g., third-party transportation, packaging, upstream energy related)	12.44
Lifecycle emissions from customer trips to Amazon's physical stores	4.05
Amazon's Total Footprint	51.17

Our carbon intensity metric, measured as grams of carbon dioxide equivalent (CO₂e) per dollar of Gross Merchandise Sales (GMS), is equal to 122.8 CO₂e per dollar (USD).

Facebook

- Dès 2012, Facebook figure **parmi les premiers à créer des Data Centers alimentés en énergie renouvelable**, après que Greenpeace l'ait épinglé, en 2010, sur son impact écologique. Depuis 2020, les opérations de Facebook sont soutenues par 100% d'énergie renouvelable, grâce à 63 projets dont 49 solaires et 14 éoliens déployés aux États-Unis, mais aussi en Irlande, Norvège, au Singapour ou en Inde.
- Facebook est **engagé depuis 2020 à atteindre la neutralité carbone d'ici 2030.**
- Facebook a aussi lancé le Centre d'information sur le Climat pour **informer sa communauté sur les enjeux et les solutions face au réchauffement climatique.**

- Tous les engagements durables de Facebook : <https://sustainability.fb.com/>



Apple

- Apple utilise **100% d'énergies renouvelables depuis 2018 pour alimenter ses Data Centers**, en ayant permis d'éviter 8M de tonnes métriques d'émissions de carbone.
- Apple utilise dorénavant de l'aluminium recyclé pour fabriquer ses appareils, et annonce avoir **réduit la consommation moyenne d'énergie de ses produits de plus de 70% en 12 ans**. Combiné aux meilleures performances de sa nouvelle puce M1, l'empreinte carbone du dernier iMac a pu être diminuée d'environ 20% par rapport à la génération précédente.

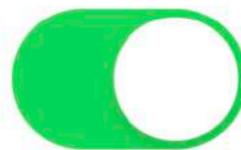
- Apple vise la **neutralité carbone d'ici 2030**.
- Tous les engagements durables de Apple : <https://www.apple.com/environment/>

**iMac brings
us closer
to net zero.**



Find out how

**Apple Inc.
Officially in
the green.**



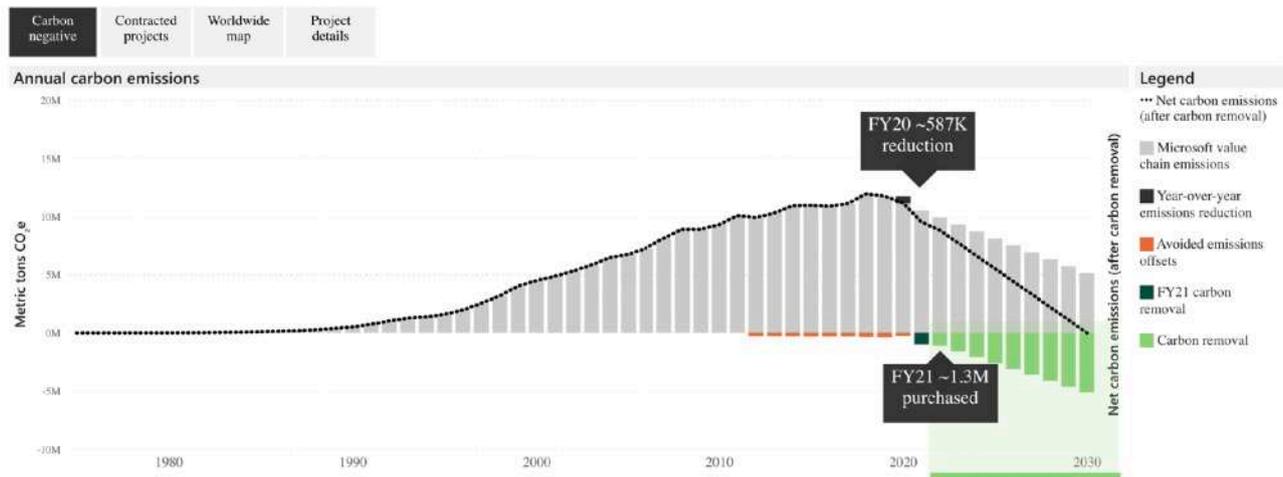
Since 2018, Apple's stores, offices, and data centers have run on 100% renewable energy. Today all our operations are carbon neutral, too.

Microsoft

- En 2020, Microsoft a annoncé vouloir être **néгатif en carbone d'ici 2030** en éliminant plus de carbone qu'il n'en émet chaque année. **1 an après, l'entreprise a déjà réduit ses émissions de 6%**, passant de 11,6 à 10,9 millions de tonnes métriques.
- Microsoft a réalisé le **plus gros achat annuel d'élimination de carbone auprès de 15 fournisseurs à travers 26 projets dans le monde**, pour un total de 1,3 millions de tonnes métriques de carbone compensées.
- D'ici 2023, il alimentera **70% de ses Data Centers en énergies renouvelables**. Son objectif est d'atteindre 100% d'ici 2025. Le projet Natick de **Data Center immergé** (p.164) s'inscrit dans la démarche de refroidissement naturel de ses serveurs.
- Microsoft s'est aussi engagé à éliminer d'ici 2050 toute l'empreinte carbone générée par la société depuis sa création en 1975.

Carbon Removal in Microsoft's Carbon Negative Pathway

In FY21 we purchased ~1.3M metric tons as a starting point in our sourcing strategy.

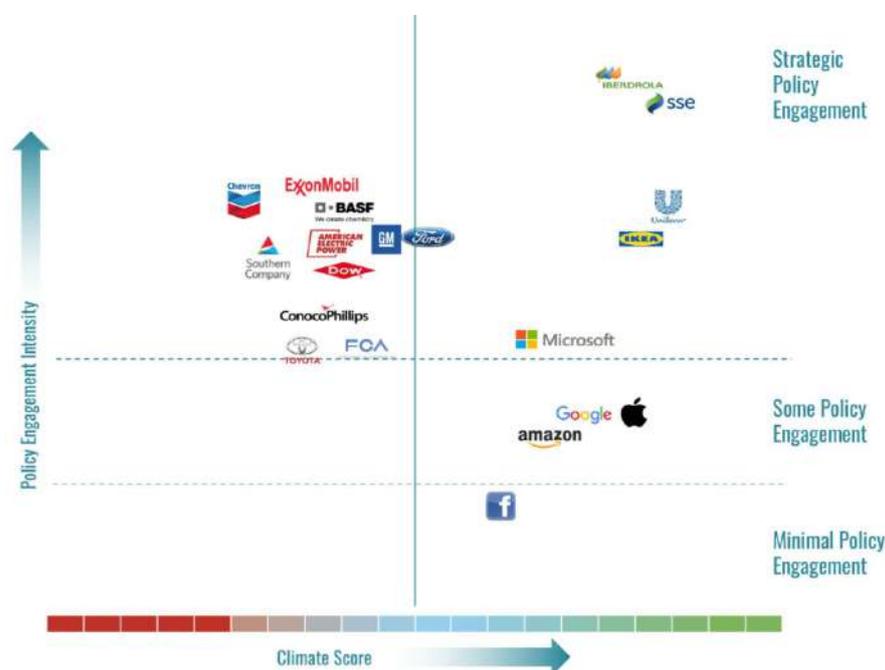


- Pour inculquer cette nouvelle valeur d'entreprise, **Microsoft taxe à hauteur de 15\$ par tonne métrique d'émissions de carbone ses départements pour les inciter à évoluer vers la durabilité.** Cet argent sert à compenser ses émissions. L'équipe de Xbox a ainsi développé une fonctionnalité qui réduit la puissance de 15W à moins de 2W lorsque l'appareil est veille.
- Microsoft publie un **rapport complet de ses engagements et actions durables**, dont les données font l'objet d'un examen indépendant du cabinet Deloitte.
- L'entreprise a enfin **nommé un Chief Environmental Officer**, Lucas Joppa, et créé le « **Planetary Computer** », une plateforme de collecte (images satellites et données utilisateurs) et d'analyse (à base de Machine Learning) de **la data environnementale**. L'objectif est d'aider les organisations à prendre des décisions en s'appuyant sur des données sur la faune et la flore, les risques d'inondation ou de séismes, etc.
- Microsoft a également lancé en 2017 le **fonds « AI for Earth »** visant à **fournir des ressources en Cloud et IA de sa solution Azure aux chercheurs travaillant sur des programmes de protection de l'environnement**. Deux ans après, 230 subventions avaient déjà été octroyées pour des projets dans 63 pays.¹³⁹

¹³⁹ <https://www.techrepublic.com/article/ai-for-earth-how-microsoft-is-using-azure-to-build-a-more-sustainable-world/>

- Tous les engagements durables de Microsoft : https://www.microsoft.com/en-us/corporate-responsibility/sustainability?activetab=pivot_1:primaryr3

Les GAFAM semblent donc s'être engagés sur de véritables plans d'action. Pourtant, dans son rapport *Big Tech and Climate Policy* de janvier 2021 ¹⁴⁰, InfluenceMap relève qu'ils n'utilisent **qu'une infime partie de leur pouvoir d'influence pour œuvrer contre le réchauffement climatique.** Alors que 38% du secteur de l'énergie a axé ses actions de lobbying en faveur du climat, les géants de la tech n'en ont **investi que 4% du fait de leurs liens avec de puissants groupes opposés aux mesures contraignantes liées aux objectifs de l'Accord de Paris,** à l'instar de l'accord de Cloud Computing de Microsoft avec le géant pétrolier Chevron ¹⁴¹, avec la société de forage Schlumberger, ou encore ExxonMobil.



Car **la tentation est forte pour les fournisseurs de Cloud** : selon BloombergNEF, les **dépenses en Cloud Computing des sociétés pétrolières** devraient passer de 2,5 Mds, en 2020, à 15,7 Mds de dollars d'ici à 2030. Cela s'explique par un **recours toujours plus intensif de l'IA** pour modéliser et améliorer l'exploration et la production.¹⁴²

¹⁴⁰ <https://influencemap.org/report/Big-Tech-and-Climate-Policy-afb476c56f217ea0ab351d79096df04a>

¹⁴¹ <https://news.microsoft.com/2019/09/17/schlumberger-chevron-and-microsoft-announce-collaboration-to-accelerate-digital-transformation/>

¹⁴² <https://about.bnef.com/blog/oil-sector-to-spend-16b-by-2030-on-cloud-analytics/>

Bien que Google ait signé en 2018 un accord avec Total sur son Cloud et ses solutions de Machine Learning ¹⁴³, elle est moins liée que Amazon ou Microsoft à ce secteur. En 2020, **Google a annoncé ne plus mettre son IA au service de l'énergie fossile.**

Comparaison des engagements climatiques					
	Empreinte carbone (MtCO2e, 2018)	Objectif neutralité carbone	Objectif sortie fossile	Réductions émissions logistique	IA pour industrie pétrolière
	44,4	2040			
	17,6	2030			
	15	2017			

En somme, certains enjeux économiques poussent les GAFAM à ne pas embrasser de politique climatique à la hauteur de leur puissance stratégique et financière. **S'ils ne sont pas encore pleinement des Géants Verts**, ils ont pris des engagements concrets, qui **permettent aux clients de leurs solutions Cloud de limiter par rebond leur propre empreinte carbone.**

Mais les PME ont-elles elles-mêmes rejoint, dans leur stratégie et leurs pratiques, le mouvement du numérique responsable ?

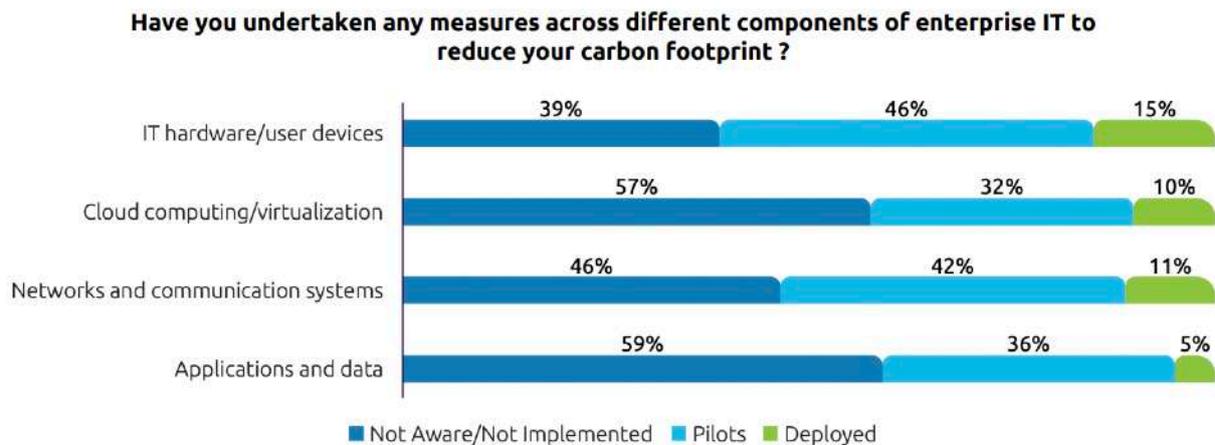
¹⁴³ <https://www.totalenergies.com/fr/medias/actualite/communiqués/total-va-developper-des-solutions-dintelligence-artificielle-avec-google-cloud>

GREEN IT EN ENTREPRISE : RÉCONCILIER ÉCOLOGIE ET ÉCONOMIE

Selon une étude de mai 2021, **61% des entreprises attendent des acteurs tech qu'ils les aident à mesurer l'impact environnemental de leur IT**. 23% seraient même allés jusqu'à rejeter les acteurs qui ne remplissent pas leurs critères RSE en la matière.¹⁴⁴

Pourtant, le « Green IT » n'est pas une réelle priorité pour la plupart de ces entreprises :

- 89% recyclent moins de 10% de leur parc informatique ;
- Seuls 43% des cadres se disent conscients de l'empreinte carbone de leur IT ;
- Bien que 50% des entreprises affirment avoir une stratégie d'informatique durable, seules 18% ont des objectifs et un rétroplanning clairs et bien définis.
- Au final, **seules 6% d'entre elles sont vraiment matures sur le sujet** du « Green IT ».



Source: Capgemini Research Institute, Sustainable IT survey, December 2020 – January 2021, N=1,000 organizations.

Au-delà des effets d'annonce, en réalité les 3/4 des entreprises débutent sur les actions concrètes liées au numérique responsable.

Et elles n'auront de **volonté réelle de le faire que si l'impact écologique impacte aussi positivement le compte de résultats**. Sans quoi les « projets pilotes » du graphique ci-

¹⁴⁴ <https://www.capgemini.com/fr-fr/wp-content/uploads/sites/2/2021/05/Final-Web-Report-Sustainable-IT.pdf>

dessus risquent de n'intéresser que la branche RSE des organisations, sans remonter au top management pour être concrètement déployés.

C'est d'ailleurs **tout aussi vrai pour le consommateur et l'impact sur son porte-monnaie**, comme vu précédemment (p.185) : un prix trop élevé ne permettra pas de faire adhérer massivement à des achats plus responsables, surtout auprès de ceux qui sont les moins sensibles à l'argument écologique. A contrario, cela ouvre de **nouvelles opportunités de business model pour les entreprises**, comme l'abonnement Darty Max à partir de 9,99€/mois qui propose la **réparabilité de ses appareils électroménagers en illimité**.

Les KPIs doivent donc être autant écologiques qu'économiques pour assurer la réussite de sa transformation digitale éco-responsable. Or, à la lumière des chiffres suivants ¹⁴⁵, il semble bien possible de **faire rimer « durable » avec « rentable »** :

- **20% à 50% d'économie en moyenne sur la facture électrique des ordinateurs**, qui représente à lui seul près de la moitié du coût d'un Système d'Information, avec la « veille automatique » qui réduit le besoin énergétique de 14%, ou le passage à des ordinateurs portables qui consomment jusqu'à 4 fois moins que des fixes.
- **30% sur les impressions papier (1 à 3% du budget d'une entreprise)** et jusqu'à 40% sur la mise en réseau d'imprimantes multifonctions (vs imprimantes individuelles).
- **L'optimisation de l'architecture des applications et serveurs par la virtualisation**, qui fait bénéficier de plusieurs machines virtuelles sur un seul serveur physique, limitant l'énergie qu'ils consomment (respectivement -19% et -11%), tout en réduisant leurs coûts d'achat, de stockage et de maintenance.
- **45% des entreprises** interrogées dans le cadre de l'étude ¹⁴⁶ seraient **prêtes à payer jusqu'à 5% de plus pour des offres « Green »** proposées par des entreprises tech.

¹⁴⁵ <https://www.greenit.fr/2011/10/13/green-it-queelles-economies-esperer/>

¹⁴⁶ <https://www.capgemini.com/fr-fr/wp-content/uploads/sites/2/2021/05/Final-Web-Report-Sustainable-IT.pdf>

Les gains se situent aussi à d'autres échelles que sur le plan économique : **satisfaction client, image de marque... et jusqu'à leur référencement**, puisque l'optimisation de la taille des images ou l'élimination des animations Flash, par exemple, améliorent le SEO.

L'obtention d'un Label « Numérique Responsable » permet enfin d'**afficher la crédibilité de sa démarche éco-responsable dans la création de son site, app mobile, logiciel** etc. Il existe divers labels de ce type, dont notamment ceux portés en France par :



- L'Institut du Numérique Responsable, anciennement « Club Green IT » et animé par des responsables RSE, Sustainable IT manager, etc. de grandes organisations (La Poste, SNCF, Engie, Société Générale, Pôle Emploi, WWF...), associé avec France IT pour son label de référence « NR » ;
- Green Code Lab pour un label d'éco-conception numérique ;
- Greenspector pour l'efficacité énergétique des solutions mobiles ;
- Wattvalue pour attester de l'alimentation d'un site en énergie renouvelable.

En somme, **le marketing éco-responsable consiste à adopter une démarche « Green de diffusion et de promotion**, en limitant au maximum son impact sur l'environnement.

Comment faire dès lors pour que l'IA marketing, nouvel outil phare du marketing dont il faut toutefois prendre garde à l'empreinte carbone, puisse montrer à son tour « patte verte » ? **Comment rendre l'IA marketing durable ?**

Comment rendre l'IA marketing durable ?



Plus que jamais, l'impact de l'IA sur les émissions de CO2 doit être un élément clé de votre processus décisionnel.



Glenn Gow Ancien CEO et expert en IA

À l'heure de la prise de conscience collective sur la réalité du réchauffement climatique, **l'engagement des dirigeants en faveur de la durabilité environnementale devient de plus en plus critique. Même le secteur bancaire**, qui par ses investissements dans les industries fossiles fait figure de premier pollueur, **se met à la « finance durable »** : la grande banque américaine Goldman Sachs s'est ainsi engagée à investir 750 Mds de dollars d'ici 2030 pour financer des projets en faveur de la transition écologique.¹⁴⁷

Il devrait en être de même pour accompagner, dans les entreprises, le déploiement de l'IA à grande échelle. Mais selon une étude du BCG¹⁴⁸, **à ce jour moins de la moitié des organisations qui ont mis en œuvre un projet d'IA l'ont fait avec une IA « responsable »**.

Par « IA responsable », on entend une IA **qui ne soit pas une boîte noire**, mais qui au contraire soit **à la fois explicable, digne de confiance, dénuée de biais, au service de l'humain et frugale en énergie**. Le critère environnemental n'est donc pas le seul, bien qu'il en représente une part grandissante, que ma thèse s'efforce de mettre en lumière.

L'étude du BCG va plus loin, en révélant que la majorité des entreprises qui disent avoir déployé des projets d'IA responsables **surestiment en réalité leur maturité sur le sujet**. Un effort de formation interne des équipes sur ce sujet devient donc indispensable.

¹⁴⁷ <https://www.goldmansachs.com/our-commitments/sustainability/sustainable-finance/>

¹⁴⁸ <https://www.bcg.com/fr-fr/press/31march2021-artificial-intelligence-responsible-new-imperative-for-companies>

Car bien que seulement 6% des entreprises souhaitant déployer une IA responsable le feront pour des enjeux de responsabilité sociale, et à **42% pour des raisons business**, loin devant d'autres critères tels que la satisfaction client (20%), la gestion des risques (16%) ou la contrainte réglementaire (14%), **les bénéfiques offrent une palette plus large :**



Aujourd'hui, les organisations les plus en pointe en matière d'IA responsable [...] la perçoivent comme un véritable avantage compétitif, que ce soit pour différencier leur marque ou bien accélérer leurs recrutements et développer la fidélisation de leurs employés. Implémenter une IA responsable est la preuve d'une culture d'innovation responsable.



Sylvain Duranton Directeur monde de BCG-GAMMA

Après avoir exposé pourquoi il était à présent nécessaire d'adopter une approche d'IA durable, je vous propose d'aborder dans ma dernière partie la **manière concrète de le faire, sous le prisme de mon cas d'étude sur l'IA marketing.**

Pour mener à bien votre projet d'IA marketing durable, la méthodologie que je vous ai concoctés permet, étape par étape, d'interroger vos besoins pour faire les bons choix d'optimisation, en suivant un certain nombre de recommandations pratiques :

- **Évaluation de vos objectifs marketing**, et des **besoins en IA correspondants** ;
- Définition de l'**écosystème IT nécessaire pour atteindre vos objectifs de la façon la plus durable** possible ;
- **Périmètre des données utiles / futiles** que vous allez ou non collecter et analyser ;
- **Mesure de vos performances marketing incluant votre KPI environnemental**, sur lequel vous allez pouvoir communiquer.

Alors **comment tirer partie des super-pouvoirs de l'Intelligence Artificielle dans votre stratégie marketing... sans pour autant détruire notre planète ? Suivez le guide !**

1. Évaluez les modalités de recours à l'IA

Selon vos objectifs, mais aussi selon vos ressources humaines et financières, vous allez déjà pouvoir **déterminer votre capacité à faire de l'IA selon 3 niveaux de complexité** :

1. **En créant votre propre modèle *from scratch*** à partir des infrastructures techniques de Amazon ([AWS](#)), Microsoft ([Azure](#)) ou Google ([Google Cloud](#)) principalement. Cette approche sur-mesure est la plus coûteuse, car elle nécessite de déployer une très forte capacité de calcul et une expertise pointue portée par des Data Scientists.
2. **En créant votre propre modèle, mais en vous appuyant pour cela sur les modèles déjà existants des GAFAMI** (en y ajoutant le « I » de IBM et son service d'IA [Watson](#)). L'avantage de cette approche est d'avoir accès via des API à des briques d'IA pour accéder à moindres coûts à leurs capacités de calcul.
3. **En souscrivant à une offre en mode SaaS clé en main**, accessible à travers le Cloud de l'éditeur de la solution, ou installée sur vos propres serveurs (« on-premise »). C'est le plus simple et le moins coûteux, mais aussi le moins personnalisable pour vos besoins spécifiques.

Nous allons décrire les étapes successives d'un framework en IA marketing durable, en regroupant ces niveaux de complexité **selon 2 cas de figure** : celui où vous développez votre propre modèle d'IA (niveaux 1 et 2), et celui où vous adoptez une solution d'IA prête à l'emploi (niveau 3).

Si vous avez déjà mis en place un projet d'IA marketing au sein de votre entreprise, ce mode d'emploi pratique a vocation à vous faire poser les bonnes questions, et à activer des leviers pour réduire votre empreinte.

1.1. Cas n°1 : vous développez votre propre modèle d'IA

Si le développement d'un projet d'IA spécifique se révèle plus exigeant, les gains que vous pouvez en tirer sont aussi proportionnellement bien plus importants. Mais il reste nécessaire de **vous interroger sur vos besoins réels, et la façon d'atteindre vos objectifs sans faire tourner des machines pour le plaisir de les faire tourner.**



Une grande partie des dernières recherches sur l'IA néglige l'efficacité énergétique, car de très grands réseaux de neurones se sont avérés utiles pour une variété de tâches, et les entreprises et les institutions qui ont un accès abondant aux ressources informatiques peuvent en tirer parti pour obtenir un avantage concurrentiel.



Carlos Gómez-Rodríguez Chercheur au Max Planck Institute For Software Systems

Comme le rappelle Thomas Solignac, CEO de Golem.ai, dans notre interview (p.247), OpenAI a par exemple doublé l'énergie déployée pour faire passer le taux de fiabilité de GPT-3 de seulement 70% à 72%, en doublant son jeu de données déjà gigantesque. C'est ce que Thomas Solignac appelle « **le plafond de verre du Deep Learning** » : **plus on s'approche d'un certain taux de fiabilité, plus l'augmenter devient difficile.**

Les géants du numérique cherchent à **faire sur-performer leurs algorithmes**, en prenant soin de ne pas communiquer sur le nombre de processeurs graphiques qu'ils utilisent, et mettre plutôt en avant leur alimentation en énergie renouvelable.

Mais à l'échelle de votre entreprise, vous laisser appâter par la tentation du pourcentage supplémentaire n'est **pas un « calcul » judicieux, ni économiquement, ni pour mener à bien votre stratégie de marketing éco-responsable.** Pour pouvoir communiquer sur une empreinte carbone raisonnée, évaluez en premier lieu les moyens à mettre en œuvre pour opérer vos apprentissages. Même si, en tant que marketeur, vous déléguerez leur mise en place à vos équipes IT, ayez conscience de ce que vous mettez sous le capot !

QUESTIONNEZ L'UTILISATION DES MODÈLES D'APPRENTISSAGE SELON VOS BESOINS

Le point le plus important est de **systématiquement interroger, pour le développement de vos projets d'IA marketing, la pertinence du recours au Deep Learning.**

Il serait d'ailleurs même sage d'interroger **jusqu'à la pertinence du recours à l'IA tout cours.** Car comme nous l'explique Gwendal Bihan, CEO de Axionable (p.234), l'audit préalable à la mise en place d'un projet d'IA dans une entreprise permet, parfois, de mettre en lumière de fausses bonnes idées. On peut alors se rendre compte qu'en fait, le besoin est un simple Dashboard ou un fichier Excel. Si l'IA est un mot galvaudé, à la mode, prenez le recul critique nécessaire pour **évaluer sa plus-value réelle pour votre projet : en avez-vous vraiment besoin, et si oui de quelle IA ?**

Le Deep Learning est aujourd'hui le modèle privilégié, utilisé pour la reconnaissance d'images et le NLP notamment. Pourtant, nous l'avons dans la seconde partie, le Deep Learning est **particulièrement énergivore** : là où le cerveau humain ne requiert que 20 watts pour fonctionner, soit autant que la consommation d'une ampoule électrique, les réseaux de neurones artificiels eux en nécessitent pour leur part pas loin de 1/2 million !

Ces réseaux de neurones créent **une architecture d'autant plus complexe sur de grands volumes de données non structurées,** où il leur faut identifier des images, des sons ou des textes. Le Machine Learning, dans sa version moins « profonde », s'appuie pour sa part sur des données structurées, clairement identifiables, en cherchant avant tout des corrélations entre elles (variables ou correspondances), et vous permettre par exemple d'identifier des clusters d'audience marketing.

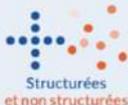
Il s'agit dès lors de **se poser les bonnes questions pour choisir votre modèle :**

- Devez-vous traiter des **données structurées ou non structurées ?**
- Si vous devez traiter des données non structurées, quel est le **niveau de complexité** du réseau de neurones nécessaire ? De quelle **puissance de calcul** avez-vous besoin

pour que le modèle créé puisse répondre à votre objectif ? Pouvez-vous **fixer un 1^{er} sous-objectif avant de déployer un modèle de plus grande envergure ?**

- Quels sont les résultats que vous attendez ? Le **gain de performances** sera-t-il suffisant par rapport au coût à la fois économique et énergétique qu'il entraîne ?

Le tableau ci-dessous expose les **avantages et inconvénients des différentes techniques de Machine et de Deep Learning**. C'est une **matrice de choix pour votre projet d'IA** :

Tableau comparatif des modes d'apprentissage						
	Supervisé	Non supervisé	Renforcé	Profond - Supervisé	Profond - Non Supervisé	Profond - Renforcé
Types de données en entrée	 Structurées	 Structurées	 Structurées	 Non structurées	 Non structurées	 Structurées et non structurées
Apprentissage en continu	 Non	 Non	 Oui	 Non	 Non	 Oui
Transfert d'apprentissage*	 Non	 Non	 Non	 Oui	 Oui	 Oui
Complexité	 Faible à moyenne	 Moyenne	 Élevée	 Très élevée	 Extrêmement élevée	 Extrêmement élevée
Durée de l'apprentissage	 Secondes à heures	 Minutes à jours	 En continu	 Minutes à jours	 Minutes à jours	 En continu
Transparence et explicabilité	 Très bonne	 Bonne	 Faible	 Très faible	 Nulle	 Nulle
Risque de biais	 Fort	 Faible à Moyen	 Faible à Moyen	 Fort	 Moyen	 Moyen
Type de ressources	CPU	CPU	CPU	CPU, TPU ou GPU	CPU, TPU ou GPU	CPU, TPU ou GPU

Voici les **éléments clés à retenir de ce comparatif**, pour éclairer vos orientations :

- Plus l'architecture d'un réseau de neurones est complexe, moins il sera transparent et explicable, s'éloignant d'autant des pré-requis d'une IA responsable. Or, **privilégier des algorithmes auditables permet de mieux les optimiser, et de limiter leur impact.**
- **Le risque de biais est accentué si l'apprentissage se fait de façon supervisée**, par la labellisation (« étiquetage ») des données avant de les entraîner. Mais **le recours au Deep Learning est un facteur de risque supplémentaire, du fait de son effet « boîte noire »**. Des erreurs ou incohérences dans le dataset initial risquent en effet d'altérer vos résultats en créant des « faux positifs », et de créer des déviations (biais racistes, sexistes...) potentiellement désastreuses en terme d'image de marque. En outre, si l'apprentissage supervisé reste à ce jour le modèle le plus utilisé, il exige des coûts d'annotation et de correction non négligeables, ainsi qu'un jeu de données suffisant pour que les algorithmes puissent s'entraîner de façon pertinente. Or, **le manque de données qualifiées pose souvent problème** au sein des entreprises, comme dans le retail où les acteurs de la grande distribution n'ont **pas toujours un CRM multicanal**.
- **L'apprentissage renforcé reste le plus énergivore, surtout en Deep Learning**. Car il apprend en continu des données clients pour, par exemple, appréhender le churn, identifier quels produits leur recommander, adapter le prix en temps réel, etc. Il se forme donc en permanence, et sollicite des ressources en conséquence.
- Nous reviendrons sur les types de ressources et le recours possible au « Transfer Learning », dans le cas où le Deep Learning reste nécessaire pour votre projet.

En somme, **l'apprentissage non supervisé en Machine Learning semble être le meilleur compromis**. Pour rappel, consiste à **ne pas faire d'étiquetage pour laisser le modèle chercher de lui-même** dans les données des éléments qui les rapprochent, et en faire des « clusters ». Vous pouvez également opter pour un **apprentissage semi-supervisé** qui permet, une fois qu'un modèle supervisé est abouti, d'envoyer des données non étiquetées pour mettre le modèle à l'épreuve, et le perfectionner tout en l'utilisant.

Ce qu'il faut retenir, c'est que **pour encadrer un projet d'IA marketing à finalité durable, il faut commencer par définir vos cas d'usage en fonction de vos données disponibles et de vos objectifs.**

Sur cette base, vous allez pouvoir choisir le mode d'apprentissage le plus adapté pour la conduite de votre projet d'IA. La métaphore employée par Laurent Cervoni, Directeur de la Recherche et l'Innovation chez Talan, lors de notre interview, est imparable :



*Il faut, en fonction des problèmes posés, **trouver le bon algorithme d'IA et pas nécessairement chercher à assommer un moustique avec une enclume.***

Le Data Scientist ou Data Analyst doit être capable d'analyser les données et de se dire « là je n'ai pas besoin de faire un Deep Learning à 18 couches, ça va tenir avec un algorithme à 2 couches ».

L'intelligence humaine est aussi de choisir le bon algorithme.



Laurent Cervoni Directeur de la Recherche et de l'Innovation chez Talan

ÉVALUEZ L'ENVIRONNEMENT IT LE PLUS SUSTAINABLE

L'état de l'art se poursuit dans cette seconde étape, qui vise à définir un environnement de traitement le plus « green » possible, pour faire tourner vos algorithmes.

Dans une démarche d'IA marketing durable, il vous faut **scrupuleusement sélectionner l'infrastructure IT la plus à même de minimiser votre impact environnemental, et ce sur l'ensemble du cycle du projet** : du matériel utilisé jusqu'aux méthodes pour former l'IA, puis l'exécuter en production.

Adoptez une solution d'hébergement « Green »

Si vos serveurs sont installés physiquement au sein de votre entreprise (« on-premise »), **privilégiez plutôt une offre de Cloud Computing**, qui permet des gains économiques et énergétiques conséquents en mutualisant les infrastructures :



Passer de l'informatique sur site au Cloud, **c'est comme passer de la voiture individuelle aux transports en commun.**



Stanislas de Rémur PDG de Oodrive, et l'un des auteurs de "Pour un Cloud européen" (Le Cherche Midi, 2020)

Pour autant, la consommation liée aux services Cloud aurait malgré tout été, en Europe, multipliée par 3,5 entre 2010 et 2018. ¹⁴⁹ Et, nous l'avons vu sur l'étude Capgemini citée p.214, seules 10% des entreprises ont mis en œuvre des actions concrètes pour limiter l'impact carbone liée à leur utilisation du Cloud Computing.

Il s'agit donc avant tout d'**opter pour une solution d'hébergement chez un fournisseur engagé dans la réduction de l'empreinte carbone de ses Data Centers.** À l'instar des GAFAM qui, nous l'avons vu, les alimentent **en énergies renouvelables**, et compensent leurs émissions pour viser la neutralité carbone : Google Cloud Platform, Amazon Web Services ou Microsoft Azure. La palme du **meilleur mix énergétique revenant à Google**, avec 56% d'énergies renouvelables utilisées, contre 17% pour AWS. ¹⁵⁰

Vous pouvez aussi opter pour des fournisseurs de colocation comme Switch, qui tourne depuis 2016 à l'énergie verte.

	Renouvelable	Gaz	Charbon	Nucléaire
Amazon - AWS	17%	24%	30%	26%
Google	56%	14%	15%	10%
Microsoft	32%	23%	31%	10%

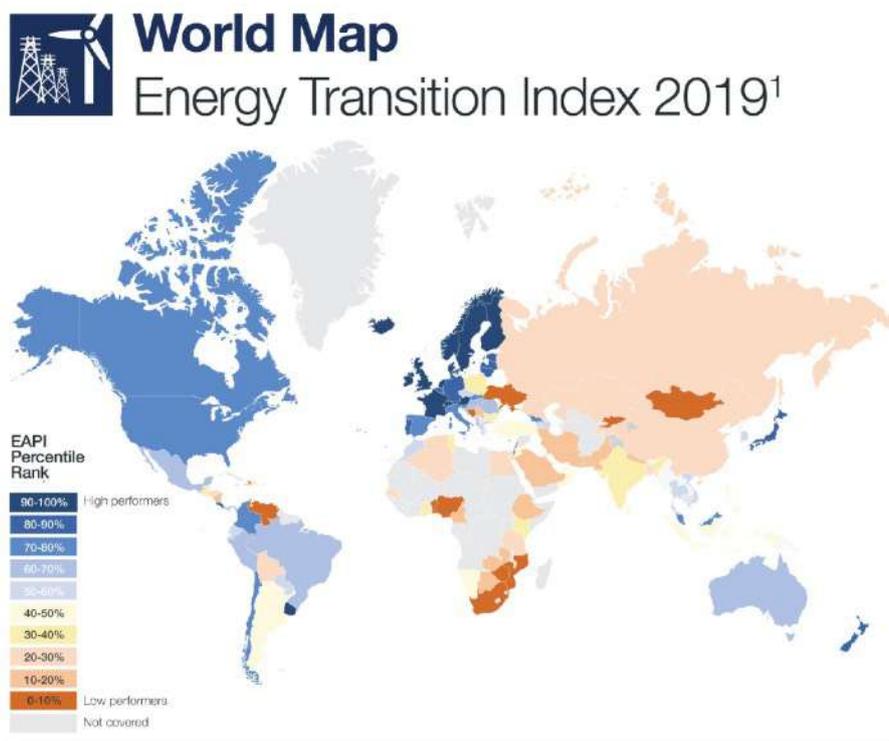
¹⁴⁹ Étude de la Commission Européenne : [https://cdn.nextinact.com/data-next/file-uploads/FINALSTUDYEnglishKK-03-20-210-EN-N13072020pdf%20\(4\).pdf](https://cdn.nextinact.com/data-next/file-uploads/FINALSTUDYEnglishKK-03-20-210-EN-N13072020pdf%20(4).pdf)

¹⁵⁰ Energy and Policy Considerations for Deep Learning in NLP, 2019 : <https://arxiv.org/pdf/1906.02243.pdf>

Reste un problème de **transmission sur la partie réseaux, qui fait transiter les données d'un bout à l'autre de la planète** pour les envoyer puis les récupérer vers et depuis des Data Centers, et ce en continu.

Face à cette problématique, **3 options se présentent à vous, consistant chacune à faire rapprocher la donnée de son lieu de calcul :**

- **Hébergez vos données dans des Data Centers localisés en France.** La France est en effet classée au 8^e rang mondial de l'indice de transition énergétique, avec un **mix fortement décarboné** comprenant beaucoup de nucléaire et très peu de charbon. Exaion, filiale d'EDF, va même jusqu'à proposer **une offre de Cloud Computing éco-responsible** : alimentation en énergies renouvelables, traçabilité assurée par la Blockchain, réutilisation de la chaleur des serveurs pour chauffer des bâtiments, et seconde vie donnée aux super-calculateurs... de quoi faire gagner des points à votre propre éco-responsabilité ! **Les pays scandinaves font aussi figure de premiers de la classe**, et bénéficient pour leurs Data Centers de l'opportunité du « free cooling », qui permet de refroidir naturellement les serveurs grâce à l'eau et l'air extérieurs.



Source: Fostering Effective Energy Transition Report 2019, World Economic Forum

- Sur des données récoltées par des capteurs IoT, par exemple pour mesurer le taux de fréquentation dans votre magasin physique, **privilégiez le *Edge Computing* qui permet d'en traiter tout ou partie directement dans le périphérique**. Ce qui réduira la consommation d'énergie en **ne les renvoyant pas vers un Data Center**.
- **Faites vos calculs sur vos données là où elles sont susceptibles de produire de la chaleur utile**, comme nous l'avons vu avec les radiateurs ordinateurs de Qarnot Computing qui permettent de chauffer des HLM (p.173). Votre impact sociétal et environnemental sera ainsi en phase avec vos objectifs (éco-)responsables.

Embrassez la logique Open Source

Nous l'avons constaté, seuls les géants de la tech possèdent les ressources suffisantes pour travailler sur l'élaboration des plus grands modèles. **Vous n'allez pas repartir de zéro, mais bien au contraire vous appuyer sur des outils open source qui facilitent et accélèrent l'entraînement et l'exécution de modèles déjà entraînés.**

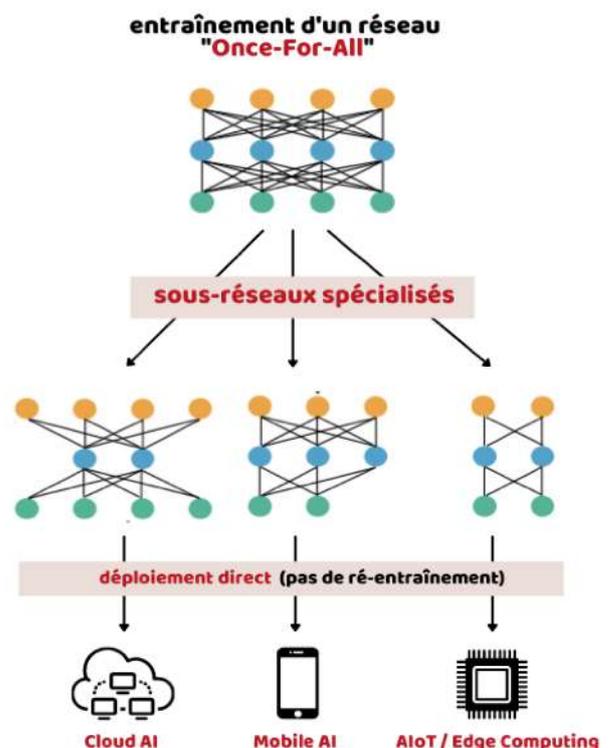
Les GAFAMI jouent le jeu de l'open-source en les mettant gratuitement à disposition de la communauté des chercheurs et des développeurs, pour ensuite pouvoir vendre des offres et services connexes au sein de leur écosystème. C'est comme cela que AWS rapporte à Amazon la majeure partie de ses revenus (35 Mds de dollars en 2019) avec un taux de croissance annuel de 37%.¹⁵¹

C'est de bonne guerre, et surtout, **la réutilisation de modèles d'apprentissage déjà formés** (comme le « Transfer Learning » sur lequel nous reviendrons) **réduit l'impact environnemental au regard de logiques propriétaires**, ce qui ne peut qu'être salué et encouragé. Le cycle de vie d'un modèle, coûteux d'un point de vue énergétique et économique pour son entraînement initial, sera d'une certaine façon « rentabilisé ».

¹⁵¹ <https://www.zdnet.fr/actualites/aws-rapporte-pres-de-10-milliards-de-dollars-de-ventes-pour-amazon-au-quatrieme-trimestre-39898351.htm>

En outre, l'open source permet aux chercheurs de s'appuyer sur l'existant en vue de travailler à des modèles d'IA toujours plus économes en énergie. Voici 2 exemples :

- Comme l'objectif de l'IA est de reproduire le fonctionnement du cerveau humain, des chercheurs tentent de **reproduire sa sobriété énergétique à partir du mode de création des connexions des neurones au cours des premières années de notre vie**. Car plus de 1 million de connexions se forment chaque seconde dans le cerveau d'un enfant, avant qu'un processus nommé « élagage » ne viennent supprimer celles qui ne sont finalement pas utilisées.¹⁵² Cet élagage peut se faire de la même façon sur les réseaux de neurones artificiels, pour réduire le poids des modèles avant de les déployer, et à performance égale.¹⁵³
- **Le MIT** a quant à lui travaillé sur un réseau de neurones nommé OFA pour « Once-For-All » : il fonctionne **sur le système des poupées russes, avec des sous-réseaux plus petits adaptés à la taille de l'appareil sur lequel sont opérés les calculs** (IoT, smartphones...) et donc à ses capacités matérielles (mémoire, processeur, batterie). Grâce à cette approche, **l'architecture de réseau la plus adaptée et optimale est automatiquement choisie**, tout en bénéficiant d'une formation commune partagée entre tous les sous-réseaux. En n'ayant pas à reentraîner chacun de ces sous-réseaux indépendamment et plusieurs fois, l'OFA offre un niveau élevé de précision tout en réduisant de manière significative leur coût énergétique et économique.¹⁵⁴



¹⁵² <https://developingchild.harvard.edu/science/key-concepts/brain-architecture/>

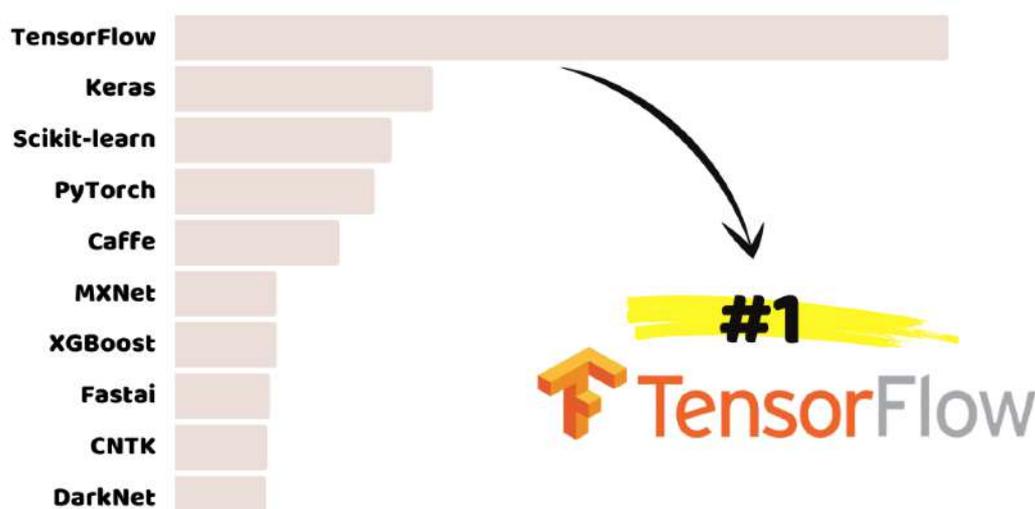
¹⁵³ <https://openreview.net/pdf?id=S1gSj0NKvB>

¹⁵⁴ <https://arxiv.org/pdf/1908.09791.pdf>

La Recherche avance, mais **revenons à présent aux aspects opérationnels** :

- **Des bibliothèques de modèles sont accessibles en open source**, dont notamment TensorFlow de Google qui est le plus couramment utilisé, mais également PyTorch développé à l'origine par Facebook, Apache MXNet utilisé sur AWS, etc.

Bibliothèques de modèles ML les plus populaires



- On trouve **dans ces bibliothèques des modèles déjà créés pour vos apprentissages en Machine et en Deep Learning**. D'autres ressources comme ModelZoo permettent de les référencer de façon transverse, en consultant **tous les modèles disponibles** : par exemple tous ceux présents sur TensorFlow, avec les liens Github vers chacun des modèles, **ou alors selon vos besoins**. Ainsi, tous les modèles qui permettent de faire du traitement du langage (NLP) pour un chatbot, de la traduction automatique, de l'analyse de sentiment, etc. sont accessibles sur cette page.
- Modèles prédictifs, modèles de clustering pour faire de la classification de données, de la recommandation de produits, de la reconnaissance d'images... **tous les usages de l'IA sont couverts par ces modèles open source, que vous allez ensuite alimenter avec vos propres données** pour répondre à votre objectif métier.

- **Reste à sélectionner les algorithmes les plus frugaux possibles**, dans le cadre de votre démarche d'IA durable. Par opposition à « l'IA rouge » (*Red AI*), qui cherche à maximiser la fiabilité des résultats par des puissances de calcul toujours plus élevées, « l'IA verte » (*Green AI*) a le souci de l'efficacité en terme de temps d'entraînement et d'émissions CO2, à précision égale ou quasi égale des modèles. ¹⁵⁵
- **Pour éclairer votre choix, comparez les performances des modèles** dans des outils mis à disposition sous licence open-source. Par exemple, le « **Responsible Machine Learning** » de Axionable, 1^{er} spécialiste français de l'IA à finalité durable, (cf. mon interview p.234), permet via sa **calculatrice d'empreinte carbone GreenOps** de vous remonter l'ensemble de ces estimations. Dans l'exemple ci-dessous, sur un objectif de scoring pour savoir à quels clients attribuer un crédit, le modèle de Régression Logistique semble le plus efficace à tous les niveaux. Le « RML » offre divers outils pour évaluer et corriger la qualité de votre jeu de données, mettre en lumière des optimisations des modèles... tout cela en proposant une sorte de compromis entre performance et impact.

The screenshot shows the 'Run Analysis' section of the Axionable Responsible Machine Learning (Demo Version) interface. It displays a table titled 'Training history per dataset' for 'German credit scoring'. The table compares three models: Random Forest, Logistic Regression, and XGBoost. The metrics shown are Precision, Recall, Time, CO2 (g), and Estimated sea ice lost (mm²). Logistic Regression is highlighted as the most efficient model.

Id	Model	Precision	Recall	Time	CO2 (g)	Estimated sea ice lost (mm²)
1	Random Forest	98%	2%	2h12	120	360
2	Logistic Regression	99%	2%	1h02	61	183
3	XGBoost	97%	3%	2h32	300	900

¹⁵⁵ « Green AI », 2019 : <https://arxiv.org/pdf/1907.10597.pdf>

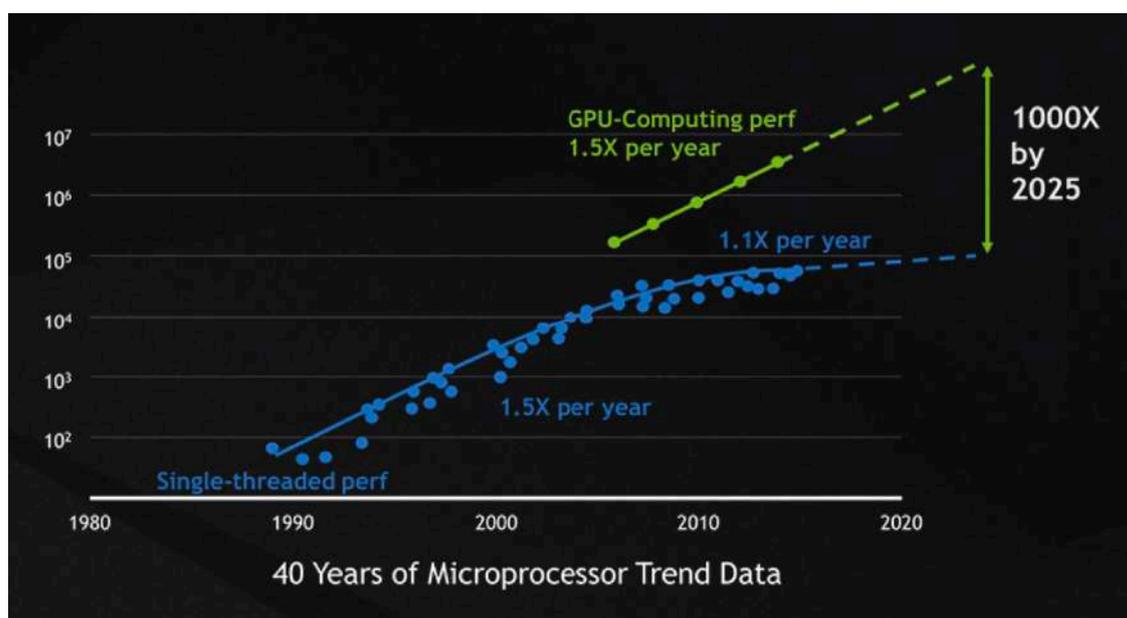
Nous vous proposerons d'autres outils de mesure d'empreinte carbone des algorithmes par la suite. Pour l'heure, après avoir sélectionné votre hébergement et votre modèle, **complétez votre architecture de traitement Green en privilégiant des capacités de calcul bénéficiant des meilleurs rendements énergétiques.**

Optez pour des puissances de calcul plus vertes

Les modèles de Deep Learning ne sont pas entraînés sur des unités centrales classiques dont les microprocesseurs CPU n'offrent pas de puissance de calcul suffisante. Le CPU peut toutefois être utilisé pour effectuer des apprentissages moins profonds.

Pour réaliser des calculs beaucoup plus intensifs, votre carte mère a du se doter de processeurs graphiques GPU, ou *Graphic Processing Unit*. Les GPU ont été créés en 1999 par Nvidia afin d'accompagner l'essor de la 3D, de la réalité virtuelle, de l'IA, et jusqu'au minage de Bitcoin.

Les GPU ont donc **pris le relais des CPU pour poursuivre la Loi de Moore** (p.105), en continuant d'assurer une augmentation des **puissances de calcul de l'ordre de +40% chaque année**. D'ici 2025, le traitement GPU devrait être d'après Nvidia **jusqu'à 1000 fois plus rapide qu'avec le CPU.**



Mais si vous avez déjà joué à un jeu vidéo avec des graphismes hyper-réalistes sur un PC boosté au GPU, ou si vous avez déjà entendu le vacarme créé par les ordinateurs et leur ventilation dans les fermes de minage de Bitcoin (comme ici en Chine), vous vous êtes rapidement rendus compte de l'énergie consommée par ces cartes graphiques. Il **faut à la fois refroidir des ordinateurs sur-sollicités, et les remplacer régulièrement. Or, le hardware est ce qui coûte finalement le plus en terme d'impact environnemental.**

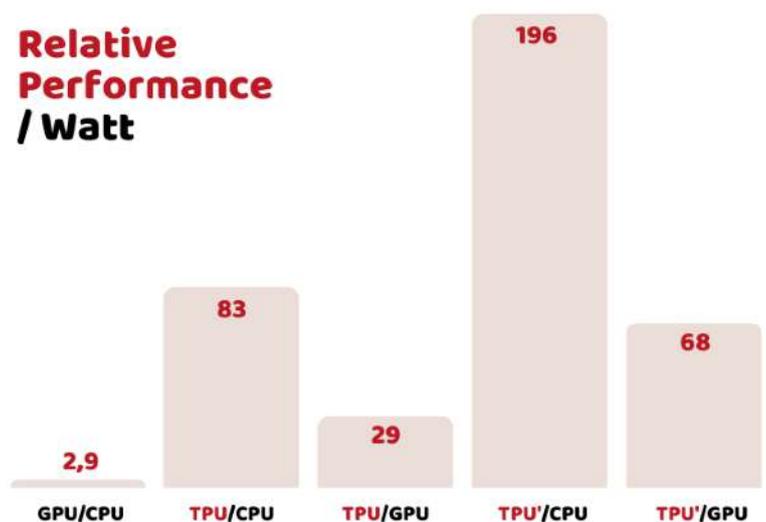
Le développement de modèles d'IA toujours plus sophistiqués requiert des puissances de calcul toujours plus importantes. Il est donc clé, pour mener à bien votre projet d'IA marketing durable, de choisir des processeurs rendant les entraînements plus efficaces.

Deux types de puces sont dès lors à considérer :

- Le **TPU (*Tensor Processing Unit*)** utilisé par Google depuis 2015, qui serait **15 à 30 fois plus rapide que les GPU tout en étant 30 à 80 fois plus efficace d'un point de vue énergétique !** Le TPU permet ainsi de réaliser pour chaque watt consommé un volume de calcul bien plus conséquent, comme le montre les ratios ci-dessous (le TPU' étant une version encore plus optimisée). Google utilise sa puce TPU sur tous ses services : moteur de recherche, Google Image/Photos/Cloud Vision/Translate, DeepMind, etc. De son propre aveu ¹⁵⁶, **Google n'a pas eu d'autre choix que de créer lui-même une puce plus efficace**, après s'être rendu compte vers 2011 que

son modèle de Deep Learning pour la reconnaissance vocale, à raison d'une utilisation de seulement 3 minutes /jour de recherche vocale, impliquerait de doubler ses Data Centers !

Le TPU est commercialisé via son offre Cloud Computing.



¹⁵⁶ <https://cloud.google.com/blog/products/gcp/quantifying-the-performance-of-the-tpu-our-first-machine-learning-chip>

- Le **NPU (Neural Processing Unit)** est lui aussi particulièrement adapté au traitement en IA, puisqu'il **s'inspire du fonctionnement du cerveau humain** pour tendre vers son efficacité. Cette **puce « neuromorphique »** est très utilisée pour la téléphonie mobile, les tablettes et l'IoT. Elle fonctionne par traitement local en évitant le transfert auprès de Data Centers. **L'IA entre sur les smartphones grâce au NPU**, pour identifier votre voix, votre visage, pré-régler vos applications selon vos comportements, ou encore fermer celles que vous n'utilisez pas. Le NPU est l'objet d'innovations constantes par des acteurs technologiques en quête de **performance accrue à un coût économique et énergétique toujours plus attractif**. Comme notamment Intel mais aussi ARM, qui conçoit les puces de Apple, Samsung, Qualcomm, et qui a été racheté en 2020 pour \$40 Mds par Nvidia, lui redonnant un avantage compétitif clé dans la course sur le marché de l'IA. Les dernières puces ARM permettent en particulier de réaliser des **calculs économes et haute performance en Edge Computing à destination de l'AIoT (Artificial Intelligence of Things)**. Enfin, il faut relever les annonces successives de **IBM**, engagé dans un objectif de neutralité carbone d'ici 2030. **Sa puce accélératrice d'IA à faible consommation d'énergie est capable d'accélérer ou décélérer lors de ses phases de calcul pour gérer elle-même l'énergie consommée**. Elle maximise elle aussi les performances tout en restant économe.¹⁵⁷ Cette puce accélératrice passait de février à mai 2021 de 7 à seulement 2 nanomètres, une amélioration de 45% de performances, et une baisse de sa consommation d'énergie de... 75% ! En terme d'usage concret, elle permettrait par exemple de multiplier par 4 la durée de vie de la batterie d'un smartphone. Si elle n'en est qu'au prototype, **les innovations sur les puces frugales s'enchaînent à vitesse grand « V »**. Et il faudra aussi très certainement compter sur **l'informatique quantique**, qui réalise des calculs extrêmement rapides avec très peu d'énergie : l'ordinateur quantique de la pépite française Pasqal ne consomme que 7 à 8 kilowatts... contre 20 pour une ampoule.¹⁵⁸

¹⁵⁷ <https://www.zdnet.fr/actualites/ia-ibm-presente-une-nouvelle-puce-a-faible-consommation-d-energie-39918185.htm>

¹⁵⁸ https://www.francetvinfo.fr/replay-radio/nouveau-monde/nouveau-monde-la-startup-francaise-pasqal-presente-son-ordinateur-quantique_4268535.html



Gwendal Bihan CEO de Axionable



Comment aidez-vous concrètement vos clients à réduire l'impact environnemental des projets d'IA sur lesquels vous les accompagnez ?

« La toute première étape, c'est d'**objectiver leur point de départ**. Photographier là où ils en sont en terme d'empreinte. On pense surtout au CO2 émis par la consommation électrique, mais souvent on oublie que la majorité de l'empreinte vient en fait des équipements eux-mêmes. Pour faire cet état des lieux, on s'appuie sur ce qui fait référence aujourd'hui, notamment la méthodologie « Bilan carbone » de l'ADEME.

De là, on peut vite **se mettre dans une logique 80/20 voire 70/30** : si on se rend compte qu'optimiser un algorithme ne ferait gagner que 0,1%, alors il n'y aurait pas de ROI et ça n'en vaut pas la peine. A contrario, un fort impact peut résulter simplement du recours au Wifi plutôt qu'à la 4G. Quand on lance un audit, la règle qu'on se fixe c'est toujours de factueliser. Car on a parfois d'instinct de fausses bonnes idées, alors qu'en objectivant on se rend compte des vraies marges de manœuvre. Si leur code n'est pas propre, que les requêtes sur leurs flux de données se fait toutes les heures alors que, pour le *use case* choisi, il pourrait ne se faire qu'une fois par semaine, alors on peut agir pour diminuer à la fois la consommation énergétique et le coût Cloud associé.

Ce **volet de rentabilité économique** est essentiel, d'autant plus depuis la crise de la Covid-19 qui a certes accéléré la prise de conscience des Comex sur les enjeux environnementaux, mais qui a aussi rendu les dirigeants encore plus ROIstes.



Impacts
business
positif



Défis sociaux
et
environnementaux

Entreprise



Certifiée



Pour ce qui est de la conception de nos IA, on essaie aussi au maximum d'**utiliser des modèles déjà entraînés**. Car au-delà de l'aspect environnemental, ça va de toute manière beaucoup plus vite de partir de tout ce qui a déjà été fait dans les Labs, notamment des GAFAM, et de pouvoir être juste sur l'ajout de quelques couches, de quelques hyper-paramètres, **sans avoir à tout re-développer**. Le monde de l'open source est quand même assez prolifique là dessus.

Quelque chose qu'on fait aussi, c'est de **passer beaucoup de temps sur le *problem statement***. Si on a des approches trop rapides sur la conceptualisation du problème, alors on risque souvent de se dire "ok on va faire du Deep Learning ou un Random Forest." Mais en fait, en faisant ça, on ne se pose pas vraiment la question de comment répondre au problème. Des fois on peut avoir des surprises, un simple Dashboard ou un Excel peut rester la meilleure solution.

C'est **une approche *frugal by design*** surtout parce qu'on passe vraiment du temps à comprendre le *pain point*, en mode *design thinking*. Du coup, on identifie bien l'irritant, on décompose comment on va entraîner, mesurer, etc. et on peut réussir à être plus frugal dans des phases très en amont de conception scientifique des modèles.

Après, il arrive qu'on ait quand même besoin au final de X modèles qui doivent s'enchaîner, et que ça restera costaud et complexe. Mais on met toujours un point d'honneur à passer du temps en conception, pour **ne pas aller trop vite sur la solution**.

Pouvez-vous nous parler de votre outil open-source « RML » visant la réalisation de projet de Machine Learning Responsable ?

En effet, on a benchmarké les outils qui existent, en mappant ce qui manquait pour **développer une IA de confiance sur tous les aspects de sobriété, mais aussi d'éthique, de robustesse etc.** On a alors réalisé nos propres développements pour tenter de combler ces manques, à travers notre outil « Responsible Machine Learning ».

Le RML intègre notamment une **calculatrice d'empreinte carbone, « GreenOps »**. Car on s'est rendu compte qu'il n'existait pas de fonction standard permettant de calculer facilement l'empreinte carbone d'un Data Lake. Et donc, ce qu'on a fait c'est qu'on a utilisé les API disponibles, comme chez Google Cloud Platform, pour récupérer la consommation au plus proche, en identifiant aussi là où sont basés les serveurs. Car entre la Belgique et la France par exemple, c'est un rapport de 1 à 10 en terme de facteurs d'émission. **Le RML permet de retranscrire via un simple dashboard ce qui consomme le plus.** Avec cette clé de lecture rapide, un DSI ou un responsable RSE pourra appliquer notre fameux 80/20 pour arbitrer ses priorités : est-ce que ce qui consomme c'est du *compute* ou bien du *storage* ? Si c'est du *compute*, est-ce que ce modèle qui consomme le plus ne peut pas être re-designer sans trop dégrader la performance ? Est-ce que ce cas d'usage est toujours utile à traiter ? etc.

Pouvez-vous nous donner vos cas d'application de l'IA durable dans le secteur du marketing ?

Nous venons d'annoncer une collaboration avec le Groupe M6 et Dentsu, pour les aider à mesurer et réduire l'impact carbone de leurs campagnes publicitaires. Pour l'instant il n'y a pas vraiment d'IA, mais il devrait y avoir, sur la phase d'industrialisation, **un sujet prédictif lié au marketing mix modeling**. Car on était jusqu'à présent sur des indicateurs purement médias/financiers : quel est mon coût contact ? Quel est mon *Gross Rating Point* ? [nombre moyen de contacts publicitaires obtenus sur 100 personnes de la cible

visée]. À présent, les annonceurs veulent **introduire une double compatibilité entre leur plan media et l'impact carbone qu'il peut avoir, pour trouver le bon compromis.**

Typiquement la TV, notamment mobile, est bien moins frugale que la radio. Et certains annonceurs commencent à se dire « est-ce que je ne devrais pas réorienter quelques investissements vers des médias à plus faible impact ? ». De l'autre côté, des médias comme le Groupe M6, qui par exemple possède RTL, peut ré-enchanter la radio, plutôt en déclin, en valorisant la sobriété du media. Ça vaut bien sûr aussi pour le podcast qui est en plein essor, dans une stratégie de media planning global pour le Groupe.

En se basant sur le **carbone comme unité de mesure**, on a un **référentiel commun pour tout le monde**, ce qui n'est pas le cas sur plein d'indicateurs où chacun a sa définition, et où ce n'est pas harmonisé. On le voit auprès des acteurs de la finance avec lesquels on travaille : chacun a son référentiel, sauf sur le carbone sur lequel il y a une convergence. Du coup, les cas d'usages qui font bouger les filières sont quand même sur des sujets carbone. **Le carbone, c'est devenu une devise.** On parle de taxe carbone, d'équivalent carbone. Il y a des méthodes de mesure qui sont connues et reconnues. Si on compare un bilan carbone, que ce soit nous qui le faisons, EcoAct ou Carbon4, globalement on va retrouver les mêmes ordres de grandeur. C'est donc **un KPI qui devient de plus en plus incontournable**, et le grand public lui-même commence à y être sensibilisé. »

The logo for 6play, featuring a stylized '6' followed by the word 'play' in a bold, lowercase sans-serif font.

finalise l'audit de son bilan carbone
avec

axionable

**et propose aux annonceurs
de mesurer l'impact
de leurs campagnes**



1.2. Cas n°2 : vous adoptez une solution IA clé en main

Vous n'avez **pas nécessairement besoin de créer votre propre modèle d'IA marketing**, y compris si vous êtes une grande entreprise. Certes, le faire en interne permet de mieux le maîtriser, et de l'adapter à ses besoins spécifiques. Cela dépend aussi bien sûr des capacités humaines et financières de votre entreprise.

Mais **si votre cas d'usage est bien cadré et qu'une solution sur le marché propose déjà d'y répondre, testez-la en priorité**. Dans la majorité des cas, elle devrait vous suffire. Si ce n'est pas le cas, **enrichissez-vous déjà de ce qu'elles proposent, avant de chercher à construire votre propre modèle**. C'est ce que fait **par exemple la SNCF** : après avoir testé une solution de chatbot tierce, la centaine de spécialistes en IA et Data recrutés par ses soins a travaillé à internaliser son développement, tant l'outil se rapprochait de leur *core business*. Comme nous l'explique Matthieu Ruault, Directeur Marketing Data & Performance chez Oui.SNCF, dans notre interview (p.258), leur approche n'en reste pas moins pragmatique, en restant également adossés à des solutions tierces (à l'image de Kameleoon ou Tinyclues, décrits dans le chapitre 1) pour des cas d'usages sur lesquels leurs équipes internes n'obtiennent pas d'aussi bons résultats.

Ainsi, bien que les capacités de calcul soient à présent bien moins coûteuses et bien plus accessibles, la plupart des projets d'IA ne seront pas construites par les entreprises, mais bien achetées et appliquées.

L'explosion des start-ups dans le secteur de l'Intelligence Artificielle offre pléthore de solutions SaaS (et on-premise) **couvrant l'ensemble des cas d'usage du marketing**. Tour d'horizon pour savoir identifier ceux qui correspondront le mieux à vos besoins, en vous assurant de leur propre éco-responsabilité.

DÉFINISSEZ VOS OBJECTIFS MARKETING POUR CADRER VOTRE BESOIN EN IA

Fred Cavazza a dressé un **panorama des solutions d'IA pour le marketing**. L'écosystème se nourrit en permanence de nouvelles start-ups spécialisée. Aussi faut-il considérer ce panorama comme non exhaustif (certaines évoquées au chapitre 1 n'y figurent pas).



Toujours est-il que vous pouvez déjà vous appuyer sur cette classification pour **piocher celles susceptibles de vous intéresser, selon votre besoin**, et en comparant leurs offres.

Alors, que souhaitez-vous faire ?

- **Avez-vous un objectif d'acquisition ?** pour segmenter votre audience en vous basant sur l'analyse des données chaudes, et scorer en temps réel vos leads et clients pour les regrouper selon leurs appétences ?

- **Avez-vous un objectif de conversion ?** pour cibler finement vos segments en hyper-personnalisant leur expérience de navigation et/ou en leur envoyant les bons stimuli marketing, sur le bon canal (emailing, Ads, pop-in...) et au bon moment ?
- **Avez-vous un objectif de fidélisation ?** pour automatiser votre SAV, et détecter les signaux faibles pour agir en temps réel sur le churn ; ou au contraire favoriser la rétention en anticipant les préférences et probabilités d'achat ?

Chaque partie du parcours client regorge de solutions dédiées à même de répondre à votre cas d'usage. Certaines d'entre elles sont bien sûr transverses, et peuvent faciliter leur intégration au sein de votre propre écosystème.

Reste que ces solutions ne pourront **pas faire de magie sans un volume et/ou un flux de données suffisant, rendus disponibles dans une CDP (Customer Data Platform), et porté par une culture d'entreprise ouverte aux nouveaux enjeux de la Data.**

Au-delà de la technologie, le véritable défi est en effet celui de la transformation digitale de votre entreprise et de vos équipes, qui doivent accepter d'intégrer et d'évaluer ces solutions de façon opérationnelle. Sans ce pré-requis, pas d'intelligence « augmentée » : l'IA restera sans support ni résultats, et ne pourra pas être efficacement déployée.

ÉVALUEZ L'ÉCO-RESPONSABILITÉ DES SOLUTIONS D'IA SUR LE MARCHÉ

Pour identifier le bon prestataire pour votre projet d'IA marketing durable, vous devez **placer comme l'un des principaux critères de choix leur éco-responsabilité**, qui sera à même, en rebond, d'assurer la vôtre. Gardez en effet à l'esprit que **le calcul annuel de votre bilan carbone se fera aussi à l'aune de celui de vos partenaires.**

Exigez donc des solutions d'IA qu'elles soient **transparentes sur la mesure de l'impact carbone de vos campagnes**, comme d'un KPI essentiel à votre stratégie marketing.

Si les Data Centers font à présent de leur démarche « green » un argument commercial central, il n’y a pas de raison que les solutions d’IA clés en mains n’acceptent pas de le faire à leur tour. **Privilégiez celles qui jouent le jeu**, et qui sont en phase avec le nouvel impératif d’éco-responsabilité comme moteur de l’économie de demain.

Des outils pour évaluer l’empreinte carbone de l’IA

Êtes-vous dans le vert quant à l’empreinte CO2 de vos actions en IA marketing ? Vos prestataires doivent pouvoir vous aider à le dire. La mesure de l’impact, au passage, devrait leur permettre d’améliorer leurs modèles d’IA vers plus de frugalité, et *in fine* vers une meilleure rentabilité.

Comme on n’améliore que ce que l’on mesure... commençons déjà par mesurer. Puis nous envisagerons **dans la partie suivante des pistes concrètes d’optimisation.**

Selon The Shift Project, **l’énergie totale consommée équivaut au cumul entre « l’énergie grise »** (à savoir celle de la production des équipements et leurs déchets), **et « l’énergie de fonctionnement »** (celle du temps d’utilisation).¹⁵⁹

Ce deuxième volet est celui qui permet d’estimer plus précisément votre propre impact, dans la mesure où le premier est réparti entre l’ensemble des clients du prestataire.

3 approches de calcul sont possibles, et doivent idéalement être croisées :

- **Calcul de l’impact a priori**
 - Le **FPO** (*Floating Point Operations*) permet d’anticiper l’évaluation des calculs et de la quantité d’énergie nécessaires, sur une tâche et un objectif donné.¹⁶⁰

¹⁵⁹ <https://theshiftproject.org/article/deployer-la-sobriete-numerique-rapport-shift/>

¹⁶⁰ <https://cacm.acm.org/magazines/2020/12/248800-green-ai/fulltext#R13>

- Son calcul n'est pas corrélé au mix énergétique ni au hardware utilisé, ce qui permet de le comparer plus facilement avec d'autres modèles, et surtout **d'évaluer son efficacité énergétique avant de le déployer.**
- Pour utiliser le FPO, votre prestataire peut implémenter sur son modèle des analyseurs comme torchstat sur PyTorch, ou keras-flops sur TensorFlow.
- **Calcul de l'impact a posteriori**
 - **ML CO2 Impact** : cet outil, à tester ici, permet d'évaluer la quantité de CO2 émise selon 4 types de données à entrer dans le calculateur :
 - Le modèle de processeur graphique utilisé ;
 - Le temps de fonctionnement du modèle de Machine Learning ;
 - Le fournisseur Cloud (Google Cloud Platform, AWS, Microsoft Azure ou une infrastructure privée) ;
 - La région où se trouve le Data Center.

ML CO2 Impact Compute Publish Learn Act About

Machine Learning Emissions Calculator

Choose your hardware, runtime and cloud provider to estimate the carbon impact of your research.

This calculator will give you 2 numbers: the **raw** carbon emissions produced and the approximate **offset** carbon emissions. The latter number depends on the grid used by the cloud provider and we are open to update our estimates if anything looks inaccurate or outdated.

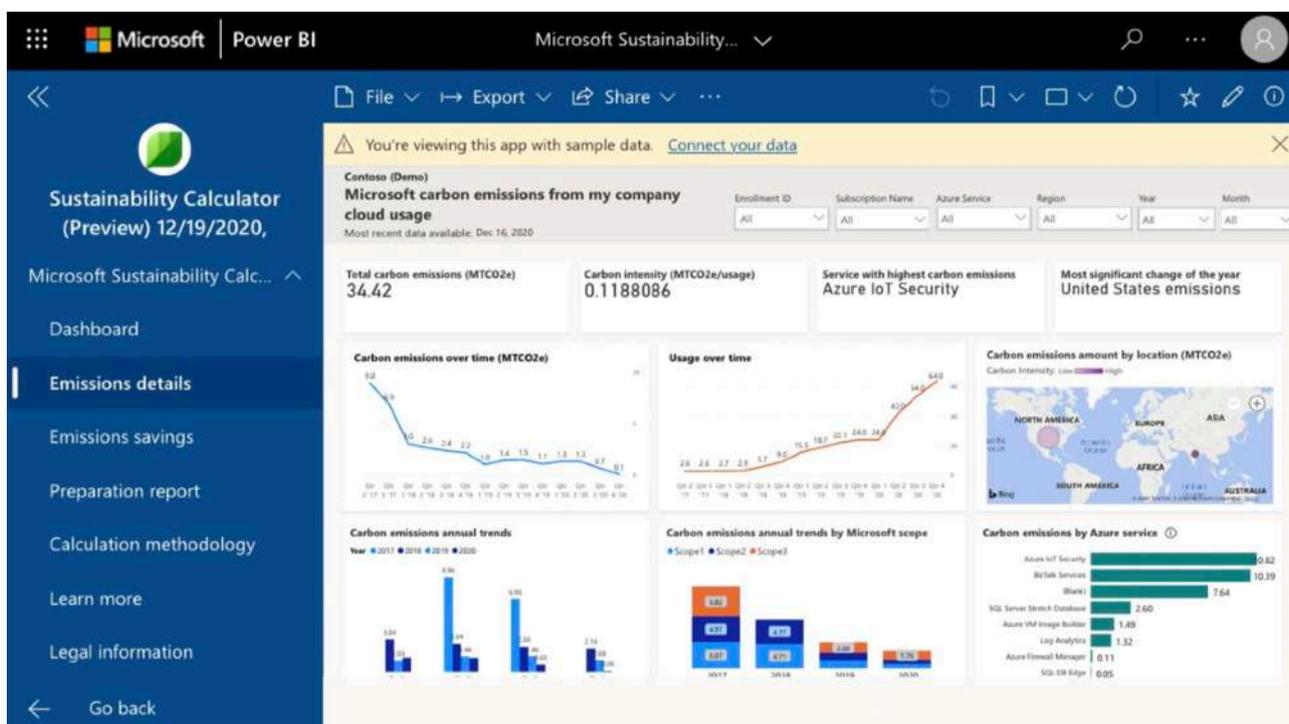
Hardware type: TPUv3 Chip
Hours Used: 36
Provider: Google Cloud Pla
Region of Compute: europe-north1

COMPUTE

CARBON EMITTED: **2.14**
CARBON ALREADY OFFSET BY PROVIDER: **2.14**
kg CO₂ eq. ?

PUBLISH THIS!

- **Microsoft Sustainability Calculator**, [à tester ici](#), a été développé par Microsoft pour **aider ses clients à comprendre l’empreinte carbone liée à leur usage de Azure Cloud** (outil CRM *Dynamics 365* etc.). Cet outil a été créé dans le cadre de son propre objectif de neutralité carbone d’ici à 2030.
 - Des informations détaillées dans un tableau de bord sur Power BI vous permettront d’évaluer vos émissions liées à votre usage du Cloud, puis **activer des leviers recommandés par l’outil afin de les réduire**.
 - Microsoft Sustainability Calculator permet en outre de **tenir à jour sa comptabilité carbone** pour vous permettre de **communiquer de façon fiable et transparente sur votre mesure d’impact**.



- **Calcul de l’impact à la volée**

- **Carbontracker**

- Des recherches récentes (*Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning*¹⁶¹) ont permis de

¹⁶¹ <https://arxiv.org/abs/2002.05651>

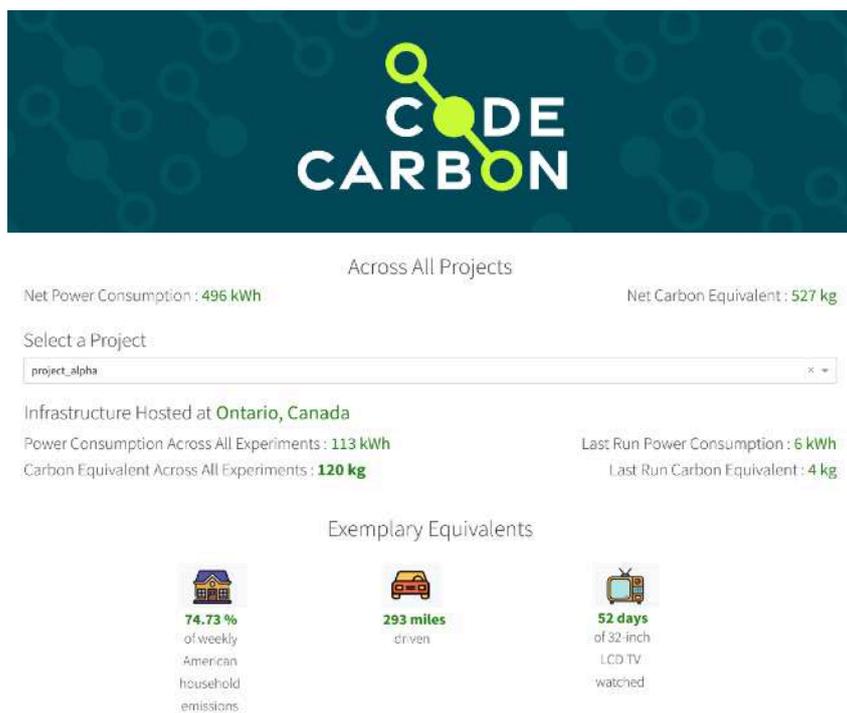
créer Carbontracker, un **outil de suivi et de prévision des émissions de carbone des modèles de Deep Learning**.

- Créé par des étudiants du département d'informatique de l'Université de Copenhague, ce programme gratuit et accessible ici présente deux avantages : il permet d'**établir un « diagnostic carbone » en évaluant la consommation CO2 du modèle**, mais il permet aussi de **placer une zone d'alerte au-dessus de laquelle la formation ou le réglage d'un modèle d'IA pour une tâche particulière peut être arrêtée**. C'est donc une fonction phare pour **rationaliser la puissance de calcul**. Arbitrer ce choix permet de responsabiliser les chercheurs et développeurs, là où ils n'avaient pas nécessairement conscience de ces impacts jusque-là.

- **Code Carbon**

- Créé notamment par Yoshua Bengio, Prix Turing 2018, Code Carbon est un outil open source qui peut être téléchargé gratuitement ici.
- Il permet lui aussi de **mesurer les émissions CO2 des algorithmes**, en estimant les ressources informatiques sollicitées à chaque exécution du code. L'outil prend en compte 3 facteurs : l'endroit où sont réalisés

les calculs, le type d'énergie qui est sollicitée, ainsi que la quantité consommée. Il livre alors des **équivalences compréhensibles**, comme on peut le voir ci-contre. Puis il fait des recommandations (privilégiez le Cloud alimenté en énergie renouvelable, etc.) pour **vous aider à diviser jusqu'à 5 la facture écologique**.



Ces outils sont des clés de compréhension et d'action incontournables pour assurer la durabilité environnementale des modèles d'IA au sens large, et de votre projet d'IA marketing en particulier. Exigez de vos prestataires qu'ils les utilisent, ou utilisez-les vous-mêmes dans le cadre des projets d'IA développés en interne.

Car dans tous les cas, **l'évaluation systématique de l'impact des entreprises sur le réchauffement climatique, et la façon dont elles œuvrent pour le limiter, devient le nouveau *Green Deal* auquel vous allez devoir répondre.**

Adoptez une solution éco-responsable : l'exemple de Golem.ai

Et si vous adoptiez **une solution qui vous garantit directement l'éco-responsabilité de sa technologie** ? C'est la promesse de [Golem.ai](#), une startup française qui développe en SaaS une **IA sans Machine Learning pour du NLP**.

Comme nous l'explique son CEO et co-fondateur Thomas Solignac, dans l'interview qu'il m'a accordée (dans les pages suivantes), **Golem.ai s'appuie sur une IA symbolique et non pas connexionniste**. Cela signifie qu'au lieu de chercher à établir des corrélations entre les données (les mots) jusqu'à obtenir un résultat attendu (le sens d'une phrase), Golem.ai se base **sur la structure existante et logique du langage**. Elle reproduit en cela le raisonnement humain, qui crée des relations à la fois grammaticales, syntaxiques et sémantiques entre les mots.

Très concrètement, l'IA symbolique de Golem.ai est **capable de détecter, dans un appel d'offres marketing par exemple, tous les critères de choix fixés pour décider ou non d'y répondre** (en allant chercher les informations relatives au délai de réponse, aux budgets etc.). L'audit, l'analyse de tout type de documents et le reporting permettent d'accélérer la prise de décision des entreprises. Avec au bout du compte une très forte fiabilité, et un véritable gain de productivité : **jusqu'à 75% de temps économisé sur l'analyse d'un dossier**, selon Golem.ai.

Parmi les autres cas d'application marketing, **l'outil facilite la gestion de la relation client, en automatisant l'analyse des messages entrants** depuis n'importe quel canal (email, SMS, réseaux sociaux...) **pour identifier l'intention de l'utilisateur**, puis au choix :

- rediriger le message au bon service de l'entreprise ;
- répondre automatiquement à l'utilisateur ;
- proposer à l'entreprise des éléments de réponse.

Au-delà d'apporter une transparence totale sur le mode de fonctionnement de son IA face à sa dimension « boîte noire », la dimension éthique se situe aussi sur l'absence de données personnelles requises, éliminant au passage toute problématique RGPD.

Grâce à l'IA symbolique, plus besoin de jeux de données ni d'entraînements, plus d'apprentissage supervisé : c'est une IA « frugal by design ».

De quoi garantir la durabilité de votre projet d'IA marketing... pour peu qu'il soit lié au traitement du langage naturel (qui couvre bon nombre de cas d'usage).

Je vous invite à en découvrir davantage sur l'approche de Golem.ai dans les prochaines pages, à travers l'interview que Thomas Solignac, son CEO et co-fondateur, a bien voulu m'accorder.



Thomas Solignac CEO de Golem.ai



Vous développez une technologie d'IA sans Machine Learning pour l'analyse automatique du langage. Quelles sont les caractéristiques de votre approche sur cette IA « symbolique » ?

« Il existe en effet deux grandes familles dans l'IA : le connexionnisme et le symbolisme, qui sont opposés sur la manière de penser l'intelligence artificielle, mais qui peuvent être complémentaires dans la pratique.

Avec le connexionnisme, lorsqu'on écrit l'algorithme on n'a pas trop d'idée de comment le problème va se résoudre dans les faits, on a plutôt une idée de comment on peut créer un algorithme qui va explorer ses options. Et après on va le nourrir de données pour le tester. L'approche la plus connue du connexionnisme, c'est le réseau de neurones artificiels, le Deep Learning. Depuis une dizaine d'années, c'est l'approche de l'IA de très loin la plus à la mode. **La plupart des gens n'arrivent pas à penser l'IA sans Deep Learning.** Entre autres parce qu'on a beaucoup plus de données disponibles que nous n'avons d'experts en IA.

À l'opposé on a l'IA symbolique, à qui on laisse moins d'autonomie pour inventer ses propres raisonnements. Mais du coup, on a des raisonnements qui sont aussi **beaucoup plus élégants, transparents, propres, traçables, et réutilisables.**

On est dans une approche qui capitalise davantage et qui est aussi souvent exprimée en consommation de données d'électricité, et en terme de frugalité. Elle est effectivement beaucoup plus légère.



Si vous privilégiez l'analyse de la structure logique du langage sans labellisation de données, comment gérez-vous les fautes de syntaxe, d'orthographe et de grammaire ?

La mécanique de compréhension de Golem.ai n'est pas basée sur l'idée que les gens vont utiliser leur langue maternelle de manière parfaitement correcte. Au contraire, l'objectif est surtout de s'appuyer sur la façon dont ils ont l'habitude de s'exprimer.

Notre algorithme est « souple », il part du principe que les gens vont oublier des mots, faire des fautes, voire mélanger des langages. Sur les sites de luxe par exemple, certains pays asiatiques mélangent leur langue maternelle avec de l'anglais. C'est très courant.

Mais pour nous peu importe : notre algorithme est basé sur des règles fondamentales qui laissent une grande liberté pour interpréter la façon dont les gens s'expriment. S'ils parlent correctement, tant mieux. Sinon ce n'est pas un problème. Ce qui importe le plus est que notre algorithme **possède une mécanique de compréhension universelle**, qui marche dans toutes les langues. Ça permet d'avoir une fois pour toutes et sans ré-entraînement, une compréhension de ce qui est dit.

Car ce qu'il faut comprendre, c'est qu'il n'y a **pas du tout d'apprentissage : aucun besoin de réinjecter des textes à notre IA, une fois qu'elle est codée elle sait « lire » toutes les langues**. En revanche, ce dont elle a besoin, c'est d'un référentiel de catégorisation.

Prenons notre cas d'usage le plus récurrent : le **traitement automatique de la relation client**. Plus on répond bien et rapidement, plus le client est content. Mais pour aller vite, l'IA doit se référer à une grille qui lui dit : un client qui demande à changer son adresse de facturation c'est l'action numéro 33, celui qui demande à acheter un de nos produits,

c'est l'action numéro 122, avec tel identifiant pour telle catégorie de produits. Quand on fait du traitement du langage en IA on veut transformer de l'information non structurée en information structurée. Et pour automatiser cela il faut lui transmettre un référentiel.

Sur quels canaux gérez-vous la relation client grâce à votre IA ?

On ne fait pas de chatbot, mais on automatise par exemple la **gestion de volume d'emails importants pour améliorer le flux de la relation client**. Avec plusieurs options : soit l'IA va savoir répondre directement et de façon automatique, soit elle va créer un brouillon de réponse à valider ou corriger par le pôle de la relation client, soit elle va rediriger l'email vers le bon service de l'entreprise. L'avantage d'une IA est de permettre **des niveaux de personnalisation élevés tout en déchargeant les équipes**. Dans le retail, on va jusqu'à gérer les bons de commande au sein même des emails.

Cas client

5 SECONDES MAX

de temps de réponse pour traiter les **4000 emails /jour** par Manutan (spécialiste équipement industriel et bureaux B2B) grâce à Golem.ai

90%

de taux de fiabilité sur les réponses grâce à une catégorisation automatique simplifiée et multi-canal (email, SMS, réseaux, CRM...).

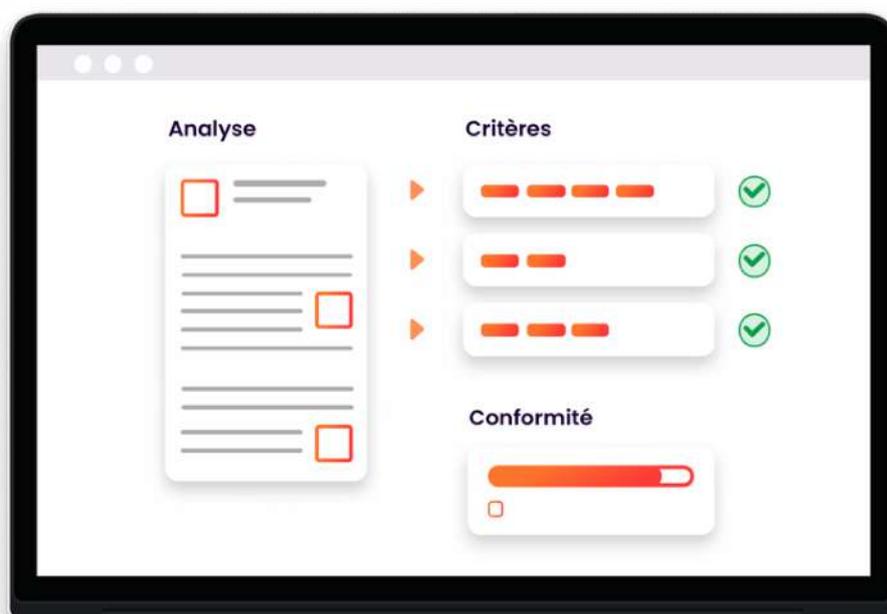


On est aussi en train d'industrialiser un autre service, qui permet de **trouver le meilleur produit dans une base de données**. Ce qui peut paraître fou, parce qu'aujourd'hui on a des moteurs de recherche sur tous les sites pour trouver des produits. Mais en fait le taux de réussite de ces moteurs n'est pas très bon. On s'imagine que les catalogues de produits sont très bien structurés, mais dans les faits ils ne le sont pas. Ou plutôt ils le sont selon une logique qui n'est pas la logique de requête linguistique. C'est intéressant parce qu'il y a une vraie divergence entre ce qu'écrivent les clients, et la catégorisation marketing d'un catalogue. Mais les termes utilisés ne seront pas les mêmes, et donc le résultat ne va pas toujours être pertinent. Golem.ai sait faire **des moteurs spécialisés sur des domaines, y compris avec énormément de références**, et augmenter drastiquement son taux de fiabilité. On sait ré-indexer une base de données produits de A à Z pour lui donner plus de sens, et faire du matching avec des requêtes de façon bien plus efficace.

Sur quels autres cas d'usage travaillez-vous ?

On travaille avec le secteur public et avec n'importe quelle entreprise qui a **un volume de textes conséquent à traiter** : banque, assurance, retail, énergie... Par exemple, on traite automatiquement les appels d'offres de Engie. Et on arrive à **réduire leur temps de traitement par 4 !** Un appel

d'offres c'est 500 à 1000 pages de Docx, de Excel, de PDF... Notre IA va « manger » tous ces contenus en 2 minutes environ, et en sortir les 30 à 80 informations les plus importantes pour aider Engie à prendre une décision rapide, en appliquant des règles métier pour valider la conformité par exemple. »



2. Rationalisez vos méthodes dans une démarche ROIste

Maintenant que vous avez pu déterminer, selon vos objectifs, le socle de votre projet d'IA marketing durable (environnement IT sustainable, choix de solutions SaaS éco-responsables...), rationalisez sa mise en production.

Pour cela, **visez l'efficacité maximale du modèle adopté en intégrant un critère de coût énergétique qui garantira un coût économique optimisé.**

2.1. Faites les bons choix en terme d'optimisation

Selon Vincent de Montalivet, Sustainable AI Leader chez Capgemini, diminuer le coût du stockage, de la bande passante, de la maintenance, etc. passe par **4 grands piliers d'un framework « Green AI » appliqué par Capgemini pour ses clients :**

- **L'infrastructure**, comme nous l'avons vu, en s'interrogeant sur la configuration la plus efficace en terme de ressources utilisées selon l'application souhaitée de l'IA ;
- **La data**, dont la collecte et l'exploitation doivent être raisonnés, comme nous allons le voir dans la partie suivante ;
- **Les algorithmes**, en privilégiant dès leur conception les solutions d'IA ne nécessitant pas, pour un même objectif de performance, un seuil de consommation énergétique 10 fois plus important, au risque de répliquer cette sur-consommation inutile lors de la mise en production ;
- **L'exécution**, pour s'assurer de corriger des requêtes mal développées au départ, et qui solliciteraient trop d'énergie.

Infrastructure

Managing the energy cost of your data ready infrastructure

Sustainable architecture, network optimization, storage, VM and cloud efficiency

Algorithm

Designing a low carbon algorithm

Models, parameters, trained strategy



Data

To preprocess your data efficiently and make data storage more resource efficient

Data queries, data modeling, data storage

Execution

A clean coding execution in a front application or dataviz tools

Queries, notifications, customized parameters, efficient applications

Opérer des optimisations sur ces 4 piliers va vous permettre de **gagner en efficacité énergétique dans une démarche primaire de retour sur investissement**, sans laquelle, nous l'avons vu, vous n'auriez que peu de chance de convaincre votre hiérarchie.

Exploiter l'existant disponible en open-source, commencer avec peu avant de déployer à plus grande échelle, placer le curseur d'un niveau d'erreur acceptable sur les résultats du modèle... autant de « **quick-wins** » **budgétaires qui sont aussi des opportunités de réduire l'empreinte carbone** de votre projet d'IA marketing.

Entrons plus dans le détail de ce que vous pouvez mettre en œuvre concrètement.

CAS N°1 : OPTEZ POUR L'ÉCO-CONCEPTION ET LE RECYCLING DE L'APPRENTISSAGE

Être frugal, c'est simplement utiliser au juste niveau les ressources dont on a besoin.

Cette quête de frugalité se joue dès la conception du modèle, pour assurer sa fiabilité et sa rentabilité.

Mettons en lumière **2 techniques d'optimisation intelligente permettant d'être « frugal by design » en Machine Learning comme en Deep Learning**.

Machine Learning : faites mieux avec moins, pour commencer

Didier Gaultier, Directeur Data Science & AI chez Business & Decision, nous donne un *business case* concret d'une **itération frugale dans le processus d'apprentissage** :



*Nous avons travaillé pour une Mutuelle qui souhaitait développer un moteur de détection de fraudes. Elle **ne disposait pas d'une base d'exemples de fraudes qualifiés**. Du coup, **nous avons dû nous rabattre dans un premier temps sur une approche non-supervisée**.*



*Ce qui nous a permis de dégager une typologie comptant quelque 1000 dossiers susceptibles de contenir des fraudes parmi des dizaines de millions. Une équipe les a ensuite épluchés et y a découvert une centaine de fraudes avérées. **Partant de là, nous avons pu créer un data set labellisé sur lequel lancer l'apprentissage supervisé**.*

Didier Gaultier Directeur Data Science & AI chez Business & Decision

Certes, la frugalité est ici liée initialement à un volume de données labellisées insuffisant pour former le modèle. Mais ce cas de figure se retrouve à vrai dire dans **une majorité d'entreprises**, qui ont comme **principale contrainte de ne pas bénéficier d'assez de données disponibles et/ou suffisamment qualifiées**.

Or, **labelliser des données coûte cher, comme solliciter des bases de données tierces**, qui ne seront d'ailleurs pas toujours assez représentatives ou qualifiées vis-à-vis de votre propre activité.

Face à cette problématique, **commencer avec peu pour passer d'un apprentissage non-supervisé à un apprentissage supervisé**, comme dans l'exemple cité, est une approche très intéressante. Cela permet de **ne pas partir sur une usine à gaz**, et de bien maîtriser son dataset initial pour garantir en sortie les meilleurs résultats.

D'autant que, comme vu p.222, **l'apprentissage non-supervisé en Machine Learning est le plus durable**, tant d'un point de vue énergétique que dans la gestion des biais.

Deep Learning : privilégiez le Transfer Learning

Si votre projet nécessite que vous développiez un modèle en Deep Learning, vous serez gagnant à tous les niveaux en adoptant **l'apprentissage par transfert**.

Car pourquoi repartir de zéro, quand des modèles ont déjà été entraînés sur des cas d'usage similaires ou suffisamment proches des vôtres ?

Le principe du « *Transfer Learning* » est de **s'appuyer sur un réseau de neurones pré-entraîné, puis y ajouter des paramètres spécifiques en vue de l'adapter à votre propre besoin**. Ainsi, vous pouvez personnaliser le modèle sans devoir le bâtir à partir de rien. Vous ne pourrez certes pas en modifier la structure, puisque comme nous l'avons vu, il vous sera gracieusement « offert » en open source, sans que ne soit révélé le secret de ses entraînements. Mais peu importe : cette base servira déjà largement votre besoin.

En outre, le modèle d'IA sera **d'autant plus performant qu'il aura déjà été éprouvé**. Il vous fera donc gagner **un temps et des ressources considérables**.

Ainsi, le Transfer Learning permet d'**empiler des connaissances** et de **les appliquer à de nouveaux champs d'application**, comme le fait naturellement notre cerveau humain. De nombreux cas d'usages sont couverts, notamment pour le traitement du langage naturel et la reconnaissance d'images.

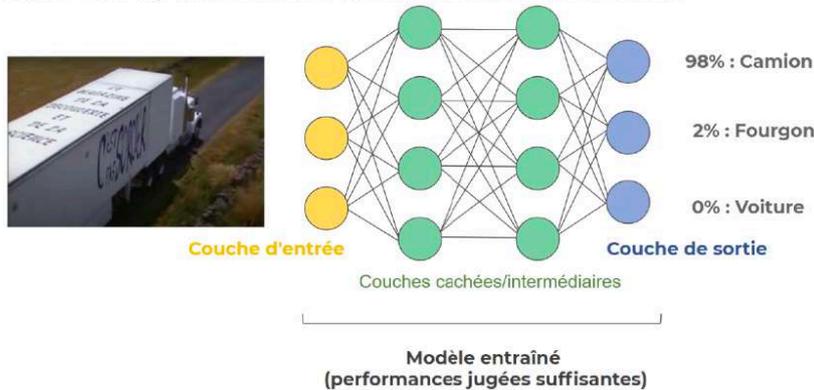
Dans l'**exemple ci-dessous**, emprunté à Laurent Cervoni de Talan (cf. interview p.133), le nouveau modèle apprend d'un modèle antérieur, qui avait été formé pour distinguer différents moyens de transport routiers. **En y ajoutant un jeu d'images ferroviaires, le modèle s'appuie sur ses connaissances existantes pour accélérer la classification** des images de trains, et éliminer celles qui représentent des véhicules routiers.

Transfert Learning

Objectif : limiter l'entraînement en utilisant un modèle déjà entraîné

But (pour cet exemple) : Entraîner un modèle permettant la classification d'images de transport ferroviaire

À disposition : Modèle déjà entraîné servant à la classification d'images de véhicules routiers

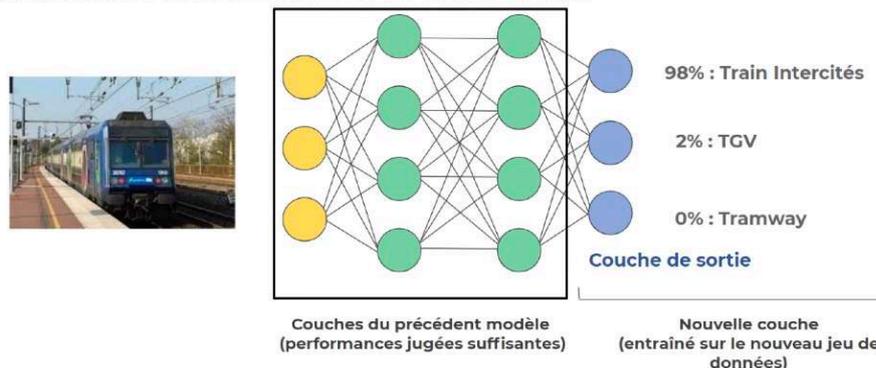


Transfert Learning

Objectif : limiter l'entraînement en utilisant un modèle déjà entraîné

Méthode

- Récupération des couches d'entrée & cachées du précédent modèle (pré-entraîné)
- Entraînement de la ou des couches de sortie(s) avec un "jeu" d'images ferroviaires



Cette approche permet d'**identifier facilement des variables dans les données, et s'en servir comme support** pour adapter rapidement le modèle à votre propre besoin, avec vos propres données.

Si vous ne disposez pas de suffisamment de données labellisées pour vos nouvelles couches d'apprentissage, vous pouvez utiliser des jeux de données accessibles là aussi en open source, comme notamment sur [VisualData](#), [Google Dataset Search](#), ou encore [UCI Machine Learning](#). Sur ce dernier, vous trouverez par exemple des jeux de données potentiellement très utiles pour des projets d'IA marketing et e-commerce, comme :

- **Dataset sur l'intention d'achat sur un site e-commerce**, [accessible ici](#), et regroupant les caractéristiques propres à plus de 12 000 visiteurs uniques sur 1 an, classés entre 15,5% d'acheteurs et 84,5% de non acheteurs.
- **Dataset sur les données transactionnelles liées à des achats en ligne**, [accessible ici](#), avec plus de 50 000 données (hors données personnelles) relatives à la quantité de chaque article par transaction, aux transactions annulées, etc.
- **Dataset sur les caractéristiques relatives à la popularité d'articles en ligne**, visant à prédire le nombre de partages sur les réseaux sociaux : basé sur les articles publiés pendant 2 ans sur le site Mashable, ce jeu de données est [accessible ici](#).

En somme, le recycling des modèles, **la « seconde main » de l'Intelligence Artificielle, constitue une ressource clé de votre projet d'IA durable.** Elle permet d'amortir le coût carbone consommé par les géants de la tech, grâce à votre réemploi de leurs modèles et des jeux de données gracieusement mis à disposition.

CAS N°2 : CIBLEZ LE TRAVAIL DE L'IA POUR MIEUX RENTABILISER VOS CAMPAGNES

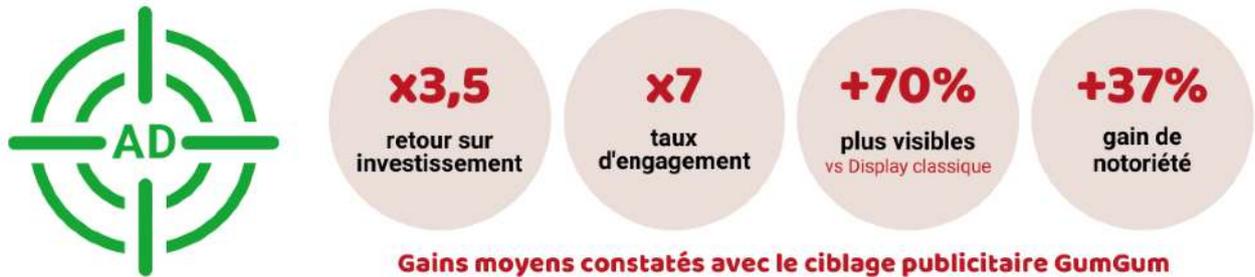
L'Intelligence Artificielle vous aide à optimiser le ROI de vos campagnes marketing en ciblant bien mieux vos leads et clients, et en diminuant sur eux la pression commerciale. Viser juste vous permet de mieux rentabiliser vos budgets, sur tout le parcours client.

Comme nous l'avons vu dans le 1^{er} chapitre, l'analyse prédictive vous promet à la fois **de meilleurs taux de transformation mais aussi une meilleure stratégie d'acquisition**, grâce à une connaissance beaucoup plus fine et fiable de vos clusters à adresser, ainsi que de la façon de les adresser.

Selon Quantcast ¹⁶², **52% des responsables marketing utilisant des technologies d'IA marketing ont ainsi vu leurs ventes augmenter**, et 51% ont vu une augmentation de la

¹⁶² <http://topcom.fr/zoom/comment-les-directeurs-marketing-apprehendent-lia/>

fidélité de leurs clients. En outre, 80% d'entre eux estiment que l'IA leur permet de se concentrer sur leur stratégie et la créativité de leur campagnes, en supprimant d'autres tâches à faible valeur ajoutée.



Si, en tant que spécialiste marketing, vous ne pouvez que difficilement être un expert pointu de l'IA et de la Data, il vous faut toutefois en comprendre les rouages, et vous familiariser avec ces nouveaux champs de compétences. Vous serez **mieux armé dans vos prises de décisions en vous basant sur les informations tirées de vos données.**

Pour cela, **sélectionnez les prestataires de la « Martech » les plus à même de mettre en application vos objectifs, chacun selon leur spécialité** : Syllabs par exemple pour de la génération automatique de textes, Kameleoon pour l'hyper-personnalisation de votre parcours client, GumGum pour vos publicités contextuelles, etc. **Ils vous aideront à cibler le travail de leur IA pour vous garantir le meilleur ROI possible, sur la base d'insights business actionnables** récoltés auprès de leurs autres clients.

Le site Oui.SNCF, nous l'avons dit, arbitre ses choix par une logique purement ROIste, en comparant pour ses projets d'IA marketing les meilleurs résultats obtenus, soit par des solutions tierces, soit par ses propres POC développés en interne. Une approche pragmatique qui le pousse à adopter celui présentant les meilleures performances.

Matthieu Ruault, son Directeur Marketing Data & Performance, nous explique dans les pages suivantes ce mode de fonctionnement. En outre, il nous expose **comment s'y déploie la dynamique Green** chère au fleuron du rail français.



Matthieu Ruault *Directeur Marketing Data & Performance chez Oui.SNCF*



Pouvez-vous nous dresser un rapide panorama des principaux cas d'usage de l'IA marketing aujourd'hui chez OUI Sncf ?

« On exploite les capacités de l'IA dans un certain nombre de cas de figure, **à la fois via des solutions du marché, et en développant certaines solutions par nous-mêmes**. On a une approche pragmatique, en se demandant toujours si c'est plus pertinent d'acheter ou de construire. Puis de mettre régulièrement en « concurrence » la solution achetée ou construite pour la challenger. **Plus on se rapproche de notre core business, et plus il y a un enjeu pour nous à maîtriser la technologie en faisant du « sur-mesure »** [comme sur le OUI bot, internalisé]. Soit parce que ça n'existe pas sur le marché et qu'il faut bien le faire, soit parce qu'on considère qu'on sera en mesure de faire mieux. C'est ce qui nous permet derrière d'apporter un avantage concurrentiel et de nous distinguer.

Au niveau SEA, on avait initialement adopté un outil du marché. Puis on a décidé de se lancer sur **notre propre outil de bidding**. On a d'abord fait un POC pour comparer les performances entre leur outil et le nôtre sur un petit groupe de mots-clés. On a constaté qu'on avait une performance similaire, alors que notre algorithme n'était pas forcément très sophistiqué. On s'est dit qu'en faisant des optimisations et en exploitant surtout la donnée qui nous est propre, on allait être meilleur, ce qui a été avéré par la suite sur ce qu'on a mesuré. Quelque temps plus tard, on a comparé notre solution interne avec un autre outil du marché, qui s'est révélé plus performant que le nôtre. On a dès lors mis le nôtre de côté, pour **encore améliorer notre gestion des enchères grâce à leur IA**.

Ce qui est intéressant, c'est que dans chaque cas de figure, **ce qui fait la différence, ce n'est en fait pas tant l'algorithme que la donnée mise à disposition de l'outil**. Le moteur de bidding du marché qu'on utilise dorénavant a accès à de la donnée à laquelle nous n'avons pas accès, ni aucun de ses concurrents. C'est finalement cette donnée qui rend un outil nettement plus performant, et pas tant l'algorithme.

Notre SEO est lui aussi plus intelligent. On a mis en place une solution propriétaire qui permet de faire un **maillage interne dynamique de notre site**, pour driver le maximum de trafic. On génère également des **contenus automatiquement**, pour toujours plus de pertinence en terme de SEO. Enfin, dans une dimension de marketing plus « produit », on a mis en place des solutions tierces exploitant l'Intelligence Artificielle pour proposer à nos clients la meilleure expérience possible, sur des usages soit de **personnalisation**, ou en fonction de ce que la personne est en train de faire sur le site à un instant T.

100

spécialistes Data et IA en interne pour « devenir une société d'intelligence artificielle au service de la mobilité » (Alexandre Viros, ancien DG).



+12%

de revenus par clic en upsale (réservation hôtel, voiture... en plus du billet) grâce à la solution de personnalisation Kameleoon.

+115%

du CA des campagnes email grâce au ciblage de TinyClues (avec -11% d'emails envoyés et +157% de taux d'ouverture)

On peut aussi à présent **réserver un billet de train par une interface vocale**. Après, il y a une question d'usage. Je pense qu'aujourd'hui, sur la partie recherche d'informations, clairement les gens ont pris le pas et se sont lancés. Mais de là à aller jusqu'à acheter un billet de A à Z par des échanges vocaux, nous n'y sommes pas encore. Une proposition de valeur de service à destination des clients, avec un très bon niveau d'expérience, devrait être le facteur qui déclenchera derrière les usages. C'est comme pour le mobile : chaque année était l'année du mobile, jusqu'à ce que Apple sorte l'iPhone, et que ça finisse par le devenir. Je pense que c'est aussi une question de génération, les jeunes étant toujours plus nombreux à n'échanger quasiment que par des messages vocaux.

Comme la SNCF est engagée dans une dynamique Green, est-ce que vous développez vos modèles d'IA marketing en essayant de limiter leur empreinte écologique ?

Dans nos pratiques de développement, il y a une **démarche « Green Tech » en interne**, pour faire en sorte que toutes les nouvelles solutions mises en œuvre, avec ou sans IA, intègrent bien une dimension Green dès leur conception. **Nous faisons en sorte que la consommation d'électricité, ou de mémoire, et tout ce qui consomme de l'énergie soit limité au strict minimum.** Ça va jusqu'à une démarche autour de l'expérience client, où nous analysons par quel moyen nos interfaces pourraient être moins consommatrices.

On a aussi **basculé dans le Cloud**. Nos équipes de production avant cela maintenaient des milliers de serveurs, avec un dimensionnement pour adresser des pics de charge,

**GRANDE
VITESSE,
BAS CARBONE.**



alors que ça n'arrivait que quelques fois dans l'année. Donc **on ne consomme plus que ce dont on a besoin**. Nous ne sommes pas non plus des philanthropes : il y a bien sûr un moindre coût économique, qui offre un bénéfice indirect en économies sur le plan de la consommation énergétique. Le meilleur moyen de faire avancer le sujet Green, c'est de corréliser les deux. Peut-être qu'avec les changements environnementaux qui s'aggravent le côté écologique pourrait se suffire à lui-même. Mais ce n'est pas encore le cas. **Le ROI reste le premier driver.**

Pour ce qui est de la gestion des données, historiquement on a essayé de corréliser leur collecte à des usages, que ce soit des usages de pilotage ou d'activation. *In fine*, il y a **assez peu de données qu'on collecte sans forcément les exploiter derrière.**

L'éco-responsabilité des solutions tierces avec lesquelles vous travaillez pourrait-elle devenir un nouveau critère de choix ?

Aujourd'hui ça ne l'est pas, mais je pense en effet que **ce sera un nouveau critère de choix**. Comme je vous l'ai dit, il y a chez nous une démarche « Green Tech » qui a été lancée. On l'applique déjà en interne. En général ce qui se passe c'est qu'ensuite on **élargit à notre environnement externe, donc à nos fournisseurs**. Il est complètement envisageable et souhaitable que ça devienne un critère de choix à l'avenir.

On a bien conscience, et on sent que **l'écologie fait partie maintenant des attentes des consommateurs**. On sent une augmentation du questionnement autour de ces sujets là. Ce n'est pas encore le premier critère de choix, mais c'est devenu un critère de choix, ce qui n'était pas le cas il y a encore 4 ou 5 ans. Nous accompagnons donc ce changement.

L'IA y participe par exemple dans la gestion du ciblage. Un email envoyé génère une dette écologique. Grâce à un outil du marché utilisé pour notre marketing relationnel, on cible mieux, ce qui permet aussi de générer moins de dépenses énergétiques. »

2.2. Réduisez vos jeux de données

La construction réfléchiée et efficiente de votre jeu de données participe pour beaucoup à la frugalité de votre projet d'IA marketing, que vous construisiez vous-mêmes votre modèle d'IA ou que vous adoptiez une solution clé en main.

Ainsi Marc Benioff, le CEO de Salesforce, estime que si la finalité d'une entreprise est de contribuer à améliorer l'état du monde, elle devrait déjà commencer par s'occuper du traitement de ses données.

Car **la Data, c'est le nouvel or noir**. C'est **le pétrole qu'on extrait, qu'on raffine**, et qui peut se transformer en **fantastiques gisements de profits**. Mais il faut pour cela **forer toujours plus profondément l'entraînement des algorithmes**. Le coût carbone du Deep Learning est en cela métaphoriquement comparable à celui de l'industrie pétrolière.

Pour limiter l'impact environnemental de son projet d'IA, il convient donc, comme nous l'avons vu, de **revoir ses méthodes de forage**. Mais aussi de **mieux identifier ce que l'on cherche, et ce qui aura de la vraie valeur sur votre tamis d'orpailleur...**

PASSEZ DE LA BIG DATA À LA SMART DATA

La tendance de ces 10 dernières années est de faire ingurgiter à l'IA toujours plus de données pour lui faire « cracher » les meilleurs résultats.

Mais attention à l'indigestion.

Plutôt que de laisser l'IA s'empiffrer de données sans goût, sans intérêt, périssables ou même avariées, réfléchissez à ce qu'il serait pertinent de mettre dans son assiette. Quel serait le menu équilibré de son dataset ?

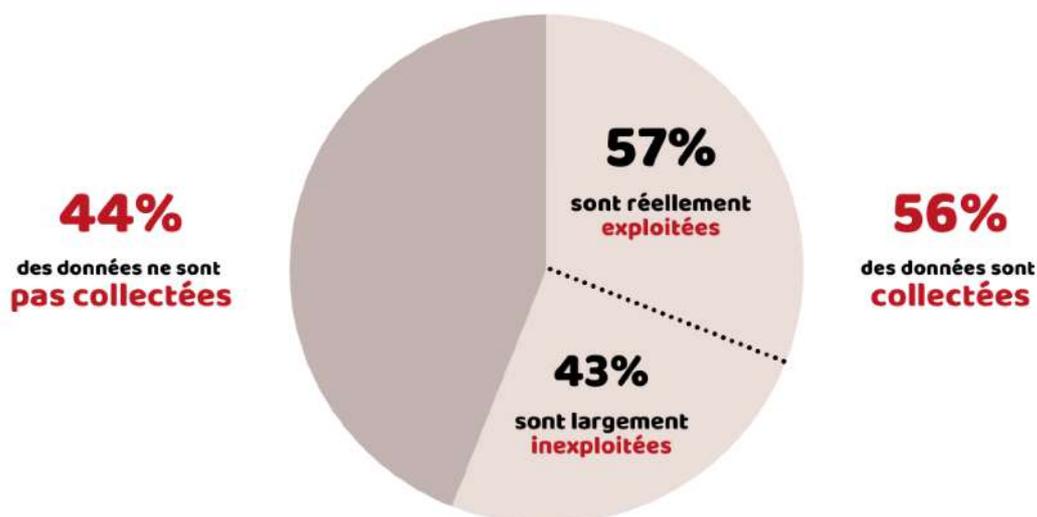
Pour mener à bien votre projet d'IA durable, **ne pensez plus « Big Data » : pensez plutôt « Smart Data ».**

La Smart Data, **ce n'est pas collecter moins : c'est collecter mieux**. Ne pas amasser de la donnée dont vous n'avez pas besoin, et vous assurez que vous collecterez bien celles qui vous seront réellement utiles. Pour cela, il faut **bien cibler vos besoins et réfléchir aux usages de votre projet d'IA marketing**.

Cette phase préliminaire est d'autant plus importante que dans les années à venir, les entreprises devraient de plus en plus être amenées à « **sortir leurs données du frigo** », selon la formule de Mick Levy.¹⁶³ **Seules 32% des données des entreprises seraient à ce jour exploitées**, d'après une étude IDC : ainsi, parmi les 56% de données collectées, seules 57% sont réellement exploitées.¹⁶⁴

Avec le déploiement de l'IA dans toutes les industries, **les entreprises vont prendre de plus en plus conscience de la nécessité de traiter leurs données**. Pour accompagner ce passage à grande échelle, il s'agit donc d'adopter les bonnes pratiques en la matière.

32% **seulement des données collectées par les entreprises sont exploitées**



© Denis Lefebvre inspiré par IDC, The Seagate Rethink Survey, 2020

¹⁶³ <https://www.dunod.com/entreprise-economie/sortez-vos-donnees-du-frigo-une-entreprise-performante-avec-data-et-ia>

¹⁶⁴ https://www.seagate.com/files/www-content/our-story/rethink-data/files/Rethink_Data_Report_2020.pdf

L'enjeu est tout aussi important pour ceux qui collectent déjà des données, mais qui les collectent mal. Selon IBM, **83% des entreprises américaines reconnaissent que leurs données collectées sont incomplètes ou inexactes.** ¹⁶⁵

Le manque à gagner en milliards de dollars chaque année se voit d'autant plus alourdi par l'allocation de **80% des tâches d'un Data Scientist sur le « Data cleaning »**. C'est une étape pourtant essentielle afin de **rendre les jeux de données exploitables, et assurer la fiabilité des résultats**. Mais elle pourrait être évitée par une collecte initiale plus propre. Les Data Scientists, dont les salaires sont estimés jusqu'à 60k€ en Europe et 100k€ aux Etats-Unis ¹⁶⁶, seraient ainsi bien mieux rentabilisés...

Plus important encore : **certes, votre stratégie marketing doit être *data-driven***, pour toujours mieux connaître vos clients, et les adresser de façon toujours plus pertinente grâce à l'IA marketing. **Mais collecter vos données avec discernement vous permettra d'accroître très nettement votre relation client.**

Alors concrètement, comment faire pour passer de la *Big Data* à la *Smart Data* ?

Recentrez votre stratégie autour de la First Party Data

- **Privilégier vos données propriétaires plutôt que les données tierces** permet à la fois de répondre à une contrainte technologique (la fin annoncée des cookies tiers), un cadre législatif (le RGPD en Europe, qui pourrait inspirer les Etats-Unis comme déjà avec le California Consumer Privacy Act ¹⁶⁷), et un attachement de plus en plus grand des internautes quant au respect de leur vie privée.
- Choisissez les données *first party* permettant de remplir vos objectifs marketing.
Envisagez des cas d'usage pertinents sans tomber dans le piège du « collectons

¹⁶⁵ http://mitiq.mit.edu/ERIQ/2008/MIT2008IQIS_Proceedings.pdf

¹⁶⁶ <https://datascientest.com/salaire-data-scientist>

¹⁶⁷ <https://www.latribune.fr/economie/international/le-rgpd-californien-une-loi-modele-exportable-au-reste-des-etats-unis-840240.html>

au cas où ». Si votre *use case* ne nécessite pas de récolter de données relatives à la géolocalisation par exemple... pourquoi les collecter ? Vous pourrez toujours le faire ultérieurement si vous finissez par en avoir besoin, en l'activant seulement au moment où vous en aurez besoin. Car de toute façon, **vos clients actifs vous assureront des données fraîches, et non périmées comme le seraient celles de vos clients perdus.** De plus, **grâce à l'IA frugale, vous n'aurez pas besoin d'un jeu de données historiques trop important** pour entraîner votre modèle (cf. ci-après).

- Pour déterminer les données utiles et pertinentes à récolter, **appuyez-vous sur votre *customer journey* idéal**, en déterminant selon vos divers points de contact les informations qui vous permettront de mieux connaître le profil de vos clients, mieux comprendre et appréhender leur comportement. Car la personnalisation passe, nous l'avons vu dans le chapitre 1, par la segmentation, le clustering de votre base selon des signaux faibles. Ces **signaux faibles sont les données clés dont vous avez besoin, les indicateurs à même de vous permettre d'activer vos leviers marketing** pour convertir vos groupes d'utilisateurs ainsi segmentés.

Oubliez la data futile pour améliorer votre relation client

- La frugalité de la collecte améliore l'expérience client **en limitant la quantité d'informations requises**, par exemple lors de l'inscription sur votre site ou votre application. Ainsi, à titre d'exemple **ne demandez pas ET l'âge ET l'année de naissance d'un nouvel utilisateur.** En faisant cela, **vous multipliez inutilement les champs à lui faire remplir, et vous risquez de le perdre en route.** Alors que vous pouvez déduire son âge en ayant déjà recueilli son année de naissance (qui reste plus utile pour s'assurer qu'il a bien l'âge minimum requis, lui adresser un mailing d'anniversaire...). **L'utilisateur sera d'autant plus ouvert à partager ses données avec vous s'il constate que la collecte sert avant tout à lui offrir une expérience optimisée.**

- Vous n'aurez de toute façon pas les bons insights sur vos utilisateurs en voulant tout savoir à leur sujet. **L'infobésité accroît surtout la complexité du *Data Mining***, à savoir l'exploration de vos données disponibles :



Florian Thiebaut Consulting Director chez Artefact

La pertinence prime sur la quantité, et la quantité de données est décorrélée de la qualité de la relation client. Plus il y a de données, plus il est compliqué de les gérer.



Chaque octet de données inutile est un frein à l'automatisation du programme du marketing en matière de sélection des informations pertinentes.

ORGANISEZ ET VIDANGEZ VOTRE DATA LAKE

Réconciliez vos données dispersées

- En recentrant votre stratégie sur la *First Party Data*, vous **évitez d'éparpiller vos données transmises jusque-là par des sources variées sur des environnements tout aussi variés**, qui limitent à la fois la qualité de la donnée et l'accès même à certaines données qui pourtant pourraient être clés.
- Pour commencer, il vous faut donc **réunir vos données sur une seule plateforme sur le Cloud**, afin d'avoir une vision unique de votre base.
- **Constituez votre *Data Lake* en l'alimentant**, selon leur utilité pour vos différents cas d'usage, **aussi bien de données structurées de votre CRM multicanal et ERP, mais aussi des données anonymes de votre DMP** (*Data Management Platform* où sont stockées les données d'utilisateurs non logués, les interactions enregistrées avec vos campagnes displays, etc.), et d'autres non structurées (emails reçus sur votre support client, etc.) représentant 90% des données aujourd'hui collectées.

- **L'identifiant CRM du client** (aussi appelé RCU pour « Référentiel Client Unique ») **joue un rôle de clé de réconciliation entre toutes ces sources**. Vous pourrez alors **appliquer un modèle de marketing prédictif dans votre Data Lake pour calculer des scores d'appétence** d'un lead identifié en fonction de ses interactions sur la campagne publicitaire que vous venez de mettre en place, par exemple.
- Pour autant, **ne rapatriez pas systématiquement toutes vos données du Cloud pour opérer ces calculs** : pour optimiser le temps de calcul des requêtes, créez des tuyaux d'accès spécifiques selon les cas d'usage que vous voulez traiter. Ce qui implique, encore une fois, de bien avoir réfléchi au cas d'usage en amont, et de quelles données doivent être sollicitées pour y répondre.
- En bref, tous ces outils de collecte de la Data parfaitement configurés vont vous permettre de **significativement accroître votre connaissance client**. Vous pourrez ainsi les suivre et les cibler de façon rapide et pertinente.

Faites la chasse aux Dirty Datas

- Nous avons évoqué dans le chapitre 2 les « **Dark Datas** », ces données présentes dans votre Data Lake qui ne sont **jamais utilisées ni analysées**, et représentent **plus de 50% de vos données non structurées**.
- Certaines de ces données peuvent être **utiles à un certain moment** : pensez aux archives de vos contrats, ou vos réponses à des appels d'offre, au sein desquelles vous pourriez venir piocher efficacement, grâce à des outils à base d'IA tels que celui de Golem.ai (cf. interview p.247). D'autres solutions offrent également des services de « **Dark Analytics** », à l'instar de Invenis, Alfresco, Splunk ou encore M-Files, pour se mettre en quête de données utiles au fin fond de vos Dark Datas... de vrais chercheurs d'or !
- Mais **une majorité de ces données ne seront in fine jamais « réveillées »**. Si elles sont stockées sans raison, il serait peut-être temps de faire le ménage...

- Au rang des « Dirty Datas » à évacuer, il y a bien sûr aussi toutes celles qui, dans le processus de réconciliation dont nous venons de parler, se révèlent **obsolètes, erronées, et/ou dupliquées. Nettoyer et dédoubler sa base est une étape clé pour la préparer à une analyse et une exploitation efficace.** Ainsi, vos clients ne risqueront plus d'être sur-sollicités par des messages marketing qui leur seraient adressés plus de fois que nécessaire.
- Pour réaliser cette **étape de nettoyage, appelée « Data Cleansing »**, privilégiez des outils de spécialistes en *Data Quality Management* tels que Winpure, Data Ladder, TIBCO Clarity, OpenRefine (open source) et bien d'autres encore ¹⁶⁸, qui automatiseront le processus. Ils vous aideront à auditer régulièrement votre base pour identifier de nouveaux Dirty Datas à faire disparaître.

Réduire votre jeu de données en l'organisant efficacement grâce à des outils intégrés, des objectifs réfléchis et segmentés, et une base régulièrement nettoyée, permettra de tirer le meilleur parti de vos données pour les mettre au service de votre relation client.

Bien sûr, toute cette étape **permet aussi de libérer de l'espace Cloud inutile, les appels qui en découlent, et donc de limiter votre coût économique et écologique.**

Bref, la bonne gestion de votre jeu de données est clé pour mener à bien votre projet d'IA marketing durable, et d'en tirer les meilleurs résultats.

¹⁶⁸ <https://www.lebigdata.fr/data-cleansing-top-outils>

3. Analysez vos résultats, Communiquez sur votre impact

Selon l'étude *Sustainable IT* de Capgemini, **seules 23% des entreprises mesurent leurs émissions de GES dans le cadre de l'évaluation de leurs initiatives en IT durable.**¹⁶⁹

En France, les entreprises de plus de 500 salariés sont déjà contraintes légalement de rendre compte de l'empreinte carbone de leurs activités au global, en établissant pour cela un « Bilan GES réglementaire ». ¹⁷⁰

Pour les autres entreprises, établir un Bilan Carbone® par le biais de l'ADEME, l'Agence de l'Environnement et de la Maîtrise de l'Énergie, permet de baser son intention, et les actions à mener, dans un framework clair, ordonné et certifié.

Pour ce qui est de l'IA durable plus précisément, l'Institut du Numérique Responsable, qui a déjà créé un label « Numérique Responsable » (p.216), se penche actuellement sur les questions relatives à l'IA durable, sous l'axe à la fois éthique et environnemental. Je contribue à leur groupe de travail, qui a nourri mes réflexions dans le cadre de cette thèse. **Si, à ma connaissance, il n'existe pas encore à ce jour de framework officiel « IA durable », mon modeste « mode d'emploi » dédié à l'IA marketing durable pourrait, je l'espère, y être utile.** J'ai en tout cas bon espoir que l'INR fasse émerger de son travail actuel un framework de ce type, tout en créant **pour l'IA durable un label certifié.**

Car les entreprises seraient prêtes à le suivre, mais il existe un vrai besoin de formation : **alors que 46% des répondants de l'enquête de CapGemini souhaiteraient diminuer l'impact carbone des algorithmes qu'ils utilisent, seulement un tiers sont conscients du coût CO2 représenté par l'IA.** Or, on ne peut améliorer que ce qu'on sait mesurer...

¹⁶⁹ <https://www.capgemini.com/fr-fr/wp-content/uploads/sites/2/2021/05/Final-Web-Report-Sustainable-IT.pdf>

¹⁷⁰ <https://www.bilans-ges.ademe.fr/fr/accueil/contenu/index/page/art75/siGras/0>

Comprendre la problématique, puis **définir un objectif** clair de réduction d'impact, et y **associer des KPIs pour évaluer la manière d'y parvenir** : tel sera le chemin vers une « IA marketing durable ». Vous pourrez ensuite **valoriser vos résultats**, en interne comme en externe, au service de votre relation client et de votre rentabilité.

Ajoutez des KPIs environnementaux à votre tableau de bord

Nous avons listé des outils pour calculer l'empreinte carbone du modèle d'IA lui-même, mais **c'est à vrai dire tout le cycle de vie de votre projet d'IA qui doit être évalué.**

Le recours à la Blockchain sera très certainement, dans les années à venir, une solution de choix pour **assurer la traçabilité de la comptabilité du coût carbone de votre IA de bout en bout**, sur l'ensemble de la chaîne, incluant l'empreinte de vos partenaires (Data Centers, solutions SaaS, API...). Le meilleur moyen de ne pas être taxé de *greenwashing* restera en effet de pouvoir assurer la **pleine transparence de la mesure de votre impact**. Cette traçabilité sûre et transparente sera gage de votre crédibilité.

Toujours est-il que pour mener à bien une conduite du changement, il vous faut dès à présent savoir quoi mesurer. **Partez des critères listés dans le mode d'emploi que je vous ai proposé** (cf. synthèse p.276) **pour savoir sur quels pain points agir en priorité.**

Il ne s'agit pas forcément de passer tout de suite dans le « vert » : l'important reste de **définir et fixer dans le temps des objectifs quantifiables**, pour agir petit à petit sur les différents leviers que nous avons évoqués :

- Sélectionnez des prestataires engagés dans une démarche durable, en localisant par exemple vos serveurs dans des Data Centers alimentés en énergie renouvelable ;
- Ayez recours à l'open source et au Transfer Learning, pour assurer l'éco-conception de votre modèle ;
- Optimiser la structuration de votre Data Lake, des données récoltées et de la façon de les exploiter ;
- etc.

Communiquez en interne comme en externe

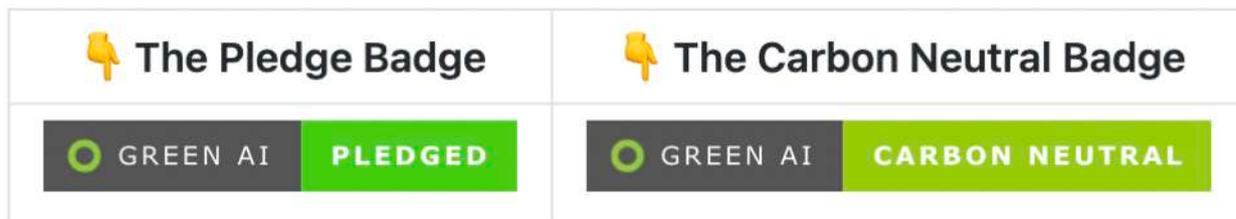
Pour assurer le financement et le bon fonctionnement opérationnel de votre projet d'IA marketing durable, il faut **mobiliser autour de vous vos équipes et votre hiérarchie**.

Pour cela, prenez soin de **bien valoriser le double gain financier et extra-financier**, étant donné que le KPI environnemental sera considéré comme non prioritaire s'il n'a pas de réel impact économique derrière.

A contrario, si ce sont les décideurs qui sont aux manettes d'une initiative « Green » à déployer en interne, ils pourront **s'inspirer de l'initiative de Microsoft**, que nous avons précédemment citée : ce dernier a en effet mis en place une **taxe carbone appliquée au budget annuel de ses unités opérationnelles dépassent les objectifs d'émissions de CO2 globales fixés par l'entreprise**. Les équipes internes doivent donc **rendre compte du coût carbone induit par leurs pratiques et leurs prestataires**. Grâce à cette approche, elles se sensibilisent bien plus rapidement à la nécessité de limiter leur impact, tout en adoptant les bons gestes pour le faire.

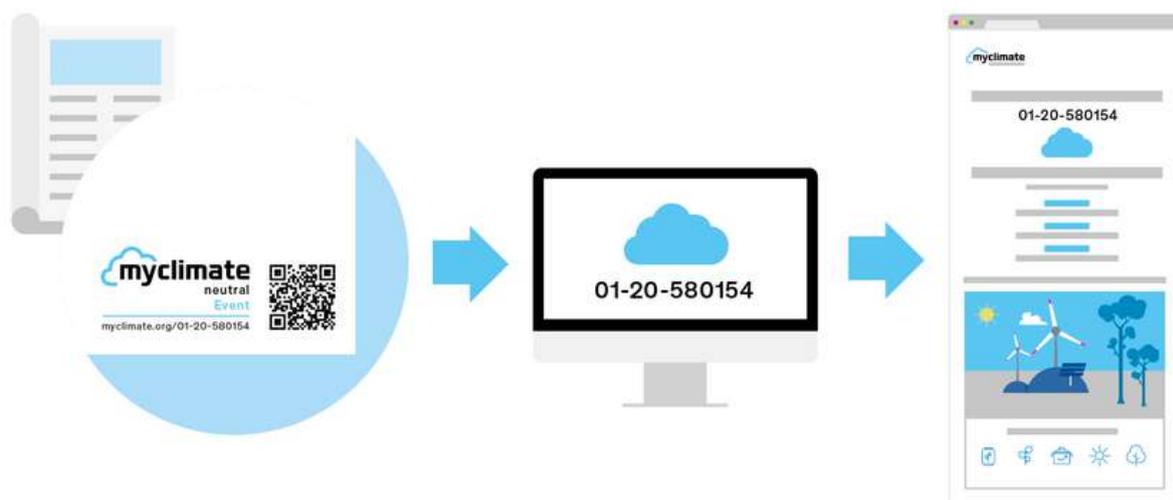
Attendez d'avoir engagé une **vraie politique de durabilité dans votre projet d'IA, et de premiers résultats, pour commencer à communiquer publiquement dessus**. Plusieurs actions pourront être mises en œuvre :

- **Obtenez un label officiel**, tel que celui que l'Institut du Numérique Responsable va probablement mettre en place dans la lignée de son label *Numérique Responsable*, et du travail de réflexion sur l'IA durable auquel je contribue ;
- Le *Green AI Standard*, créé par David Dao, par ailleurs fondateur de *GainForest*, qui utilise le Machine Learning et les Smart Contracts de la Blockchain pour lutter contre la déforestation, vise à développer une norme sur l'IA verte. Le *Green AI Standard* a ainsi vocation à sensibiliser les entreprises sur les bonnes pratiques en matière d'IA durable, et à attribuer des badges à celles qui y sont « engagées » (*pledged*).



Au-delà de limiter votre impact environnemental, compensez les émissions ne pouvant pas être évitées, en faisant des donations à des initiatives écologiques : par exemple sur de la reforestation, la préservation de la biodiversité, etc.). Nous l'avons vu au cours du chapitre 2, c'est à vrai dire la pratique la plus couramment utilisée par les entreprises et les marques pour afficher un engagement « Green », au risque qu'elle ne soit pas à la hauteur des émissions réelles... et que cet « arbre qui cache la forêt » puisse alors être taxé de *greenwashing* (p.188). Cette démarche doit donc s'inscrire en complément de vos efforts réels et sincères pour réduire au maximum votre impact énergétique à la racine, et pas seulement *a posteriori*.

Ces actions de compensation peuvent être menées **via des organismes comme Gold Standard, ActForest ou MyClimate**. Votre engagement doit pouvoir être prouvé, non seulement en redirigeant vers le certificat officiel du *Green AI Standard*, mais aussi en exposant les méthodes de calcul de vos compensations carbone.



Conclusion

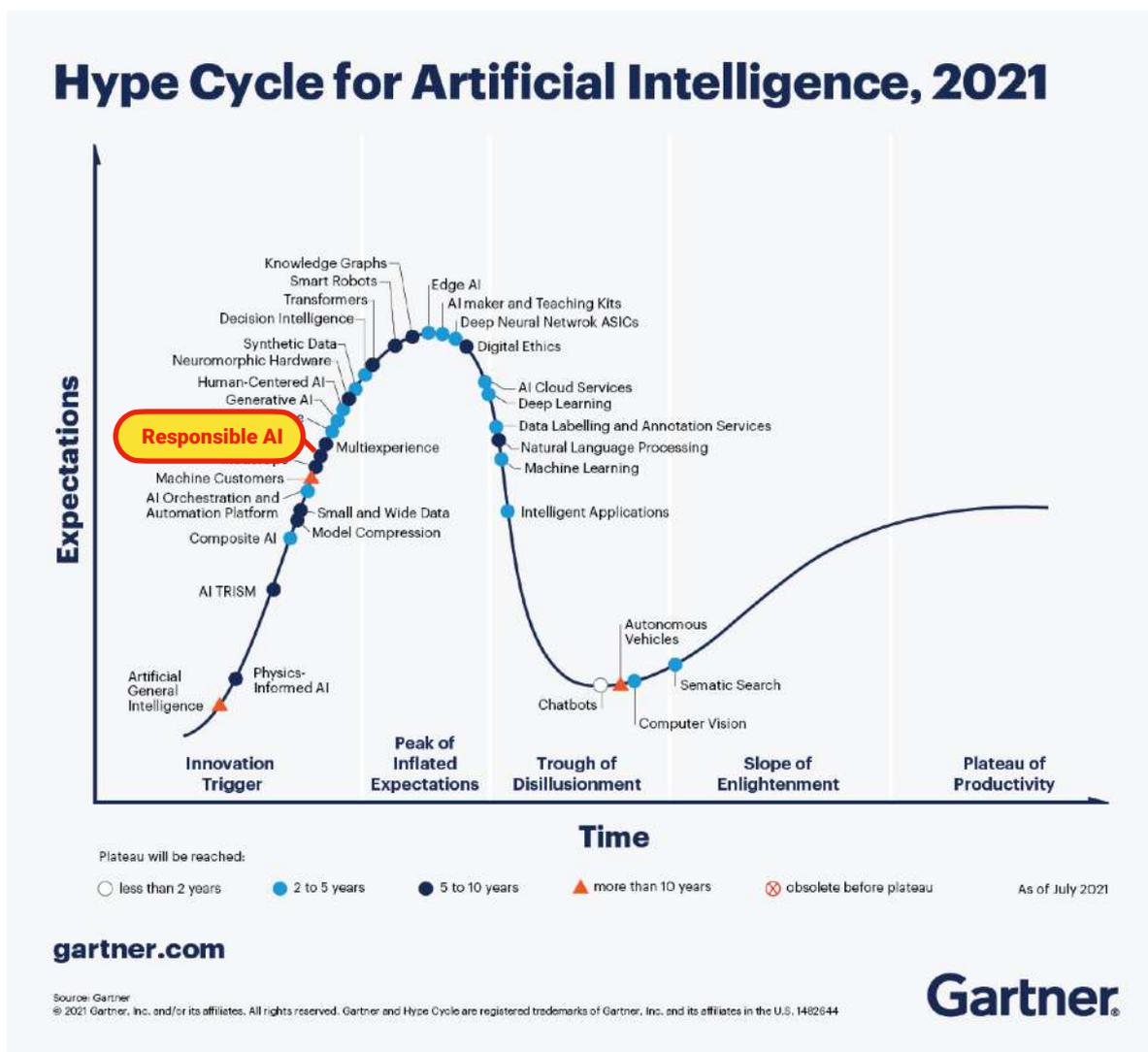
Comme nous l'avons vu dans le 1^{er} chapitre, en listant plus de 40 outils opérationnels, **l'Intelligence Artificielle devient un allié incontournable de votre stratégie marketing**, et en particulier **de votre marketing digital**, en vous permettant de :

- **Collecter et analyser efficacement vos données**, bien plus rapidement que vous ne pourriez le faire manuellement, grâce aux données chaudes captées en temps réel.
- **Sur-optimiser vos leviers d'acquisition**, en content marketing, SEO, SEA, Display... grâce à des outils d'aide à la création et à la diffusion intelligente de vos contenus.
- **Segmenter finement votre audience pour mieux l'adresser**, selon un ensemble de « signaux faibles » détectés dans vos données, révélateurs de la probabilité d'achat, de l'appétence pour une catégorie de produits, etc.
- **Hyper-personnaliser l'expérience de navigation** en adaptant automatiquement le parcours sur votre site aux segments dans lesquels sont placés vos visiteurs, et en vous rendant disponible 24/24 7/7 grâce à un assistant virtuel hybride.
- **Limiter la pression commerciale** en adressant le bon message à la bonne personne, au bon moment et sur le bon canal.
- **Prédire le comportement de vos cibles** pour convertir l'intention d'achat, prévenir le risque d'abandon, anticiper la demande en adaptant vos produits ou vos services.

En somme, l'IA vous permet de mieux connaître vos clients, améliorer la précision de votre ciblage et de vos campagnes, et fluidifier votre parcours client de bout en bout. Grâce à quoi **vous optimisez significativement vos coûts d'acquisition, vos conversions et votre satisfaction client.**

Les super-pouvoirs de l'Intelligence Artificielle n'ont pas fini de vous surprendre, pour une stratégie *data driven* toujours plus performante. Mais, vous le savez, **de grands pouvoirs impliquent de grandes responsabilités...**

Or, le perfectionnement des modèles d'IA se heurte à certaines problématiques, liées à l'effet « boîte noire » du Deep Learning, qui engendre un manque de transparence et un risque de biais. Mais aussi, depuis peu, sur la **consommation énergétique générée par ses apprentissages, et par l'hébergement des données** dans les Data Centers. Or, pour devenir une IA de confiance industrialisable à grande échelle, le « **Green AI** » sera une étape clé du « *Responsible AI* », ciblé comme l'un de ses enjeux majeurs selon le Hype Cycle 2021 de Gartner :



Le **bilan carbone de l'IA**, nous l'avons vu dans le chapitre 2, est en partie compensé par ses capacités prédictives à même de **nous aider dans la lutte contre le réchauffement climatique**, pour la transition énergétique et la préservation de la biodiversité.

Mais pour qu'elle reste notre meilleur allié dans ce défi majeur du XXI^e siècle, et qu'elle n'en alourdisse pas la facture, il lui faut **appliquer à elle-même la voie de la frugalité**. Le risque d'une consommation d'énergie galopante proviendra en effet de la multiplication des applications de l'IA dans toutes les industries.

Pour que l'Intelligence Artificielle soit un vrai accélérateur de la révolution écologique, il faut veiller à rationaliser ses méthodes d'apprentissage ainsi que les ressources utilisées. Mais il faut aussi **inscrire l'IA dans une démarche de développement à la fois durable et rentable**, en prenant soin d'**aligner enjeux environnementaux et intérêts business** pour assurer la mobilisation des entreprises.

Les technologies d'IA au service du marketing doivent **d'autant plus y répondre que les consommateurs attendent à présent des marques des pratiques éco-responsables**. La prise de conscience écologique est exacerbée par l'intensification des catastrophes naturelles, qui sensibilise d'autant plus l'opinion depuis la crise sanitaire de la Covid-19.

Dans le troisième chapitre, nous avons cherché à délimiter les contours d'un marketing éco-responsable, et la façon dont l'IA marketing pouvait y contribuer.

Je vous ai aussi et surtout livré **un mode d'emploi pratique pour vous aider à vous poser les bonnes questions avant de déployer ou d'utiliser une solution d'IA marketing**. Et ce, que vous souhaitiez développer votre propre modèle d'IA, et/ou que vous fassiez appel à une solution SaaS clé en main. L'objectif dans les deux cas est de **viser l'efficacité** pour atteindre vos résultats, **tout en restant économe écologiquement et financièrement**.

Vous trouverez dans la page suivante une **version synthétisée de ma recommandation**, sous forme d'infographie, en 5 étapes clés et en 3 axes majeurs : évaluez, rationalisez... puis communiquez !

Comment rendre l'IA marketing durable ?

Si vous développez votre propre modèle d'IA

Si vous adoptez une solution IA clé en main

Évaluez selon vos objectifs marketing

➤ **le mode d'IA adapté à votre besoin**

en remettant toujours en question l'utilisation du Deep Learning (le plus énergivore)

➤ **un environnement IT "Green"**

hébergement Cloud alimenté en énergie renouvelable, recours à l'open source...

➤ **une solution SaaS dimensionnée**

en fonction de vos données disponibles (quali/ quanti) et vos cas d'usage ciblés

➤ **l'éco-responsabilité de la solution**

en estimant l'impact carbone de son propre environnement IT (cf. liste outils proposés)

Rationalisez dans une démarche ROIste

➤ **développez une IA *frugal by design***

commencez petit avant de déployer à l'échelle

➤ **recyclez les apprentissages**

en adoptant le *Transfer Learning*



➤ **ciblez le travail de l'IA pour une performance tant économique que énergétique**

viser juste = budget d'acquisition optimisée + meilleures conversions



➤ **réduisez vos jeux de données**

- collectez uniquement la data utile à vos cas d'usage, en vous concentrant sur vos *first party*
- nettoyez et dédoublonnez votre base de données afin de pouvoir l'exploiter efficacement



Communiquez sur votre éco-responsabilité

➤ **en interne** sur le double gain économique et écologique pour mobiliser toutes vos équipes

➤ **en externe** en vous appuyant sur une mesure d'impact fiable, certifiée par un label officiel en valorisant la compensation des émissions ne pouvant être évitées (dons)



Boîte à outils (non exhaustive)

IA marketing

 Kameleoon  tinyclues  GetQuantity
Predictive B2B Scoring

Ciblage prédictif

 SITECORE  Marketo  TARGET SELL  CHEETAH DIGITAL

Personnalisation

 AB Tasty  botmind  naci  mindsay  dydu

 topic  frase  gumgum  pencil  AntVoice

Pub & Référencement

 syllabs  LABSENSE  G  Looka

Content Marketing

 narrative science  grammarly

 copysmith  phrasee

 articoolo

 AI-WRITER



 Golem.ai

 axionable

 QARNOT COMPUTING

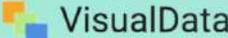
 TensorFlow

 PyTorch

 exaion EDF GROUP



Model Zoo

 VisualData



Keras

 mxnet

 Google Cloud

Open Source & Green Cloud

 fast.ai

 sklearn

 XGBoost

 Caffe2

Dark & Dirty Data Analytics/Cleansing

 OpenRefine

 Data Ladder
Get the most out of your data



 TIBCO



 WinPure™

 splunk

 invenis

 M-Files

 TRIFACTA Wrangler

 Alfresco

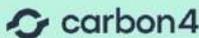
Mesure d'impact CO2

 CODE CARBON

 Carbon Tracker
Initiative

 Microsoft Sustainability Calculator



 carbon4

 ML CO2 Impact

 ecoact
atlassian.com

IA durable

Annexes



Bibliographie

LIVRES BLANCS

- AI for sustainability and the sustainability of AI. Springer, février 2021.
- Climate AI - How artificial intelligence can power your climate action strategy. Capgemini Research Institute, 2020.
- Data éthique / IA éthique : les 2 visages d'un futur responsable, Business & Décision, décembre 2019.
- E-Commerce et RSE : The Green Deal. KPMG x FEVAD, novembre 2020.
- Grande consommation et Distribution. Le développement durable, facteur déterminant de préférence d'achat. Capgemini Research Institute, 2021.
- How AI can enable a Sustainable Future. Microsoft et PwC, 2020.
- IA durable & Résilience - Accélérer la sortie de crise. Axionable et EBG, 2020.
- Intelligence Artificielle, Restez maître de votre futur ! Business & Décision, mars 2021.
- L'Intelligence Artificielle responsable : le nouvel impératif des entreprises. BCG Gamma (Boston Consulting Group), mars 2021.
- Personnalisation IA - Adoptez l'approche prédictive pour créer des expériences uniques et maximiser vos conversions. Kameleoon, 2019.
- Personalize & Convert - 10 grandes marques témoignent. Kameleoon, 2019.
- Sustainable IT - Why it's time for a Green revolution for your organization's IT. Capgemini Research Institute, 2021.

WEBINARS

- IA frugale : une IA qui consomme moins de data et d'énergie, mythe ou réalité ? Talan et Catalix, avril 2021.
- Intelligence artificielle : comment construire une ville durable grâce à la data. Paris Zéro Carbone, décembre 2020.
- Intelligence Artificielle et consommation énergétique. Planet Tech' Care, février 2021.
- Le futur du Marketing Automation : l'hyper-personnalisation, Marketing Week, mai 2021 (nombreux webinars sur l'IA au service de l'expérience client).
- L'intelligence artificielle pour comprendre l'Océan. Fondation Tera Océan, mai 2020.
- Marketing prédictif & IA : matchez vos produits avec vos clients, et inversement ! Avanci, TinyClues et BeautySuccess, février 2021.
- Responsible Machine Learning. Axionable x Explor'IA, mai 2020.

ÉTUDES SCIENTIFIQUES

- CAI Han, GAN Chuang, WANG Tianzhe, ZHANG Zhekai, HAN Song. Massachusetts Institute of Technology, MIT-IBM Watson AI Lab. *Once-for-all: Train One Network and Specialize it for Efficient Deployment*. Avril 2020.
- DEVLIN Jacob, CHANG Ming-Wei, LEE Kenton, TOUTANOVA Kristina. Google AI Language, *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. Mai 2019.
- HENDERSON Peter, HU Jieru, ROMOFF Joshua, BRUNSKILL Emma, JURAFSKY Dan, PINEAU Joelle. Stanford University, Facebook, McGill University. *Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning*. Février 2020.
- SCHWARTZ Roy, DODGE Jesse, SMITH Noah A., ETZIONI Oren. Allen Institute for AI, Carnegie Mellon University, University of Washington. *Green AI*. Juillet 2019.
- STRUBELL Emma, GANESH Ananya, MCCALLUM Andrew. College of Information and Computer Sciences, University of Massachusetts Amherst. *Energy and Policy Considerations for Deep Learning in NLP*. Juin 2019.

LIVRES

- AMARSY Stéphane, *Mon directeur marketing sera un algorithme - L'intelligence artificielle remplacera-t-elle les marketeurs ?* Editions Kawa, 2017. 160 p.
- CAZALS François et Chantal, *Intelligence Artificielle - L'intelligence amplifiée par la technologie*, Editions De Boeck, 320 p.
- COURBOULAY Vincent. *Vers un numérique responsable*. Actes Sud, 2021. 214 p.
- JEAN Aurélie. *De l'autre côté de la machine - Voyage d'un scientifique au pays des algorithmes*. Éditions de l'Observatoire, 2019. 205 p.
- JULLIA Luc. *L'intelligence artificielle n'existe pas*. First, 2019. 287 p.
- LEVY Mick. *Sortez vos données du frigo - Une entreprise performante avec la Data et l'IA*. Dunod, 2021. 256 p.

PRINCIPAUX ARTICLES WEB

- *AI in the 2020s Must Get Greener—and Here's How*. IEEE Spectrum, fév. 2020.
- *AI for Green : la branche verte de l'intelligence artificielle*. ImpactAI, mai 2021.
- *Bienvenue au Chief Personalization Officer*. MC Factory, mars. 2021.
- *Big Tech and Climate Policy*. InfluenceMap, sept. 2021.
- *Chatbots et IA: le machine learning n'est pas toujours la solution !* FrenchWeb, 2017.

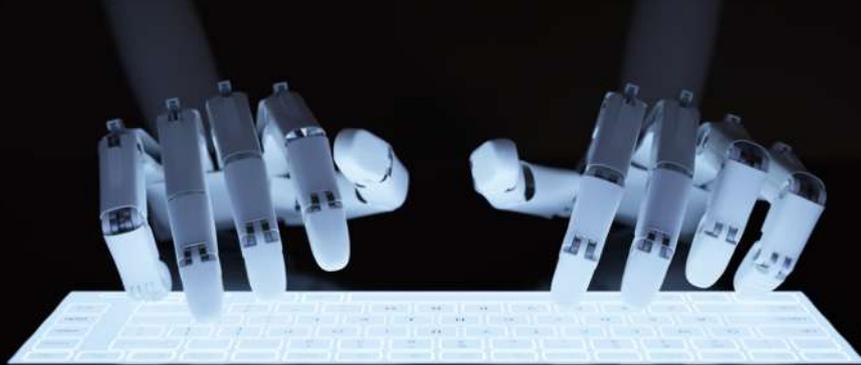
- Comment concilier Marketing Digital et Impact environnemental ? AudreyTips, 2020.
- Créer rapidement une IA business, c'est possible. JDN, juin 2021.
- En finir avec les dirty data : l'heure du grand nettoyage ! Informatique News, 2020.
- En intelligence artificielle, ce n'est pas la taille qui compte ! Le Point, fév. 2021
- Environmental Sustainability And AI. Forbes, août 2020.
- Facebook et Google Ads : le ciblage publicitaire manuel est mort. Vive le machine learning dans vos pubs et stratégies digitales. Com&Web, jan. 2021.
- Google, Microsoft, IBM are using AI to predict climate outcomes – with promising results. The Rising, mai 2020
- Green IT : quelles économies espérer ? GreenIT.fr, oct. 2021.
- How AI Helps Amazon Improve Sustainable Packaging. Triple Pundit, oct. 2020.
- How AI can accelerate a brand's eco-responsibility. Capgemini, mars 2021.
- How chatbots help businesses to improve customer experience. emplifi, juin 2021.
- How Marketers Can Use AI to Tune into Planet-Loving Prospects. MTA, juin 2017.
- IBM présente une nouvelle puce à faible consommation d'énergie. ZDNet, fév. 2021
- Impact de l'IA sur l'environnement : démêler le vrai du faux ! JDN, jan. 2021.
- L'IA s'interroge sur sa voracité énergétique, Data Analytics Post, nov. 2020.
- La recherche en IA est très énergivore. Voici pourquoi. The Conversation, jan. 2021.
- La révolution de l'IA dans le Marketing. AFRC.
- La vérité sur l'intelligence artificielle en entreprise. FredCavazza.net, juillet 2020.
- Le marketing responsable, oxymore ou nouveau défi ? Le Hub La Poste, juin 2021.
- L'impact de l'intelligence artificielle de Google sur le SEO et Entretien "Deep Learning" avec Olivier Andrieu. ia4marketing.fr, mars 2021.
- Livraison le jour-même, est-ce réellement ce que veulent vos clients ? JDN, avril 2021.
- Machine learning et Marketing Digital: le duo incontournable. Eminence, déc. 2020.
- Mesurer l'impact environnemental des algorithmes ? Avec CodeCarbon, c'est désormais possible. Les Echos, déc. 2020.
- Personnalisation prédictive, l'IA au service du marketing. Blog MBAMCI, fév. 2020.
- Pourquoi l'IA est-elle devenue aussi énergivore ? L'actualité, jan. 2021.
- Quand l'intelligence artificielle vient au secours de la nature. Actua IA, nov. 2020.
- Reducing the carbon footprint of artificial intelligence. MIT News, avril 2020.
- Refroidir nos data centers, un vrai casse-tête. L'Express, août 2018.
- RSE et Green IT : vers un numérique (un peu plus) responsable. L'Express, mai 2021.
- Vers une IA moins gourmande en données et en énergie. Hello Future, fév. 2021.
- Why We Should Care About The Environmental Impact Of AI. Forbes, août 2020.

Vidéo de teasing de thèse



Mes articles et podcast #MBAMCI

MBA
MCI
LET'S GO DIGITAL



Dopez votre copywriting à l'Intelligence Artificielle !



Denis Lefebvre

 @lefden  in/denislefebvre



MBA
MCI
LET'S GO DIGITAL

Analyse prédictive

l'IA au service
du ciblage marketing



Denis Lefebvre

 @lefden  in/denislefebvre

IA et ciblage publicitaire

Visez juste et au meilleur prix !

MBA
MCI
LET'S GO DIGITAL



Denis Lefebvre

@lefden in/denislefebvre



Let's Go Digital #MBAMCI

Comment prédire vos résultats marketing ?



Jiang Shen



Denis Lefebvre



Mes articles sur ia4marketing

ia4marketing.fr



Archives de l'auteur : Denis Lefebvre

Accueil / Auteur de l'article : Denis Lefebvre

Entrepreneur depuis plus de 10 ans dans le développement de logiciels innovants pour le livre, je complète mon expertise en marketing digital au sein du MBA MCI (Institut Léonard de Vinci). Ma thèse en cours porte sur l'IA marketing et sur l'enjeu de la réduction de son empreinte écologique.



Botmind : le chatbot hybride qui convertit vos visiteurs en clients

Relation clients • Par Denis Lefebvre •
15/06/2021 • 1 Commentaire

Botmind : le chatbot hybride qui convertit vos visiteurs en clients Vous les voyez fleurir sur les sites sur lesquels vous naviguez : les « assistants virtuels ». Le terme est privilégié à celui de « chatbot » pour s'éloigner de l'image décevante qu'il a longtemps pu renvoyer. Or le chatbot nouvelle génération, boosté à l'intelligence artificielle, n'a plus rien...



GetQuanty : boostez vos ventes B2B grâce au scoring prédictif !

Marketing Prédictif • Par Denis Lefebvre •
13/04/2021 • Laisser un commentaire

GetQuanty a recours à l'Intelligence Artificielle, et plus particulièrement au scoring prédictif, pour permettre aux acteurs B2B d'identifier et de qualifier les visiteurs de leurs sites restés anonymes. Finie, la prospection à froid. Vous allez pouvoir maximiser votre potentiel de conversion en vous appuyant sur des données chaudes, collectées en temps réel sur votre site,...



Chatbot : jusqu'où humaniser la relation client ?

Relation clients • Par Denis Lefebvre •
04/03/2021 • Laisser un commentaire

Vous vous êtes sans doute déjà amusés à jouer avec un chatbot comme on joue avec un chat : en essayant de le titiller. Si vous l'aviez alors jugé inefficace et impersonnel, vous serez surpris de voir à quel point les progrès de l'intelligence artificielle sont en train de le métamorphoser, à l'oral comme à...



Retrouvez également sur ia4marketing
la page dédiée de ma thèse professionnelle