

# CAROLINE GOULARD

Co-fondatrice et PDG, Dataveyes

Bonjour,

Merci de m'avoir donné la chance de m'exprimer devant vous. Je m'appelle Caroline, je suis cofondatrice de la société Dataveyes spécialisée dans les interactions entre les humains et les données.

Notre métier chez Dataveyes est de designer et de construire des dispositifs qui permettent aux humains de mieux comprendre les données, de mieux travailler avec, de mieux s'informer avec, et tout simplement de mieux vivre avec au quotidien.

Le fait même qu'il existe aujourd'hui des personnes, comme mon équipe et moi, qui sont spécialisées dans cette question de la relation entre humain et données, est un signe que cette relation ne va pas de soi.

Je vais, dans les minutes qui viennent, faire un état des lieux de ces relations entre humains et données, en lien avec la problématique de cet après-midi.

Les données sont un sujet dont on parle de plus en plus, parce qu'elles se sont beaucoup massifiées ces dernières années. Elles sont devenues le format privilégié pour encapsuler les grandes quantités d'informations qui circulent autour de nous.

Cette massification des données a été vue comme une opportunité par beaucoup :

- Une opportunité pour les entreprises tout d'abord. L'opportunité d'améliorer leur capacité à prendre de bonnes décisions, une opportunité d'enrichir leur business model, de mieux répondre aux besoins de leurs clients, mais aussi une opportunité de moderniser leur processus, de gagner en productivité.
  - *ex. On peut citer ici l'exemple de Procter & Gamble qui a équipé 50 000 de ses collaborateurs avec un système de visualisation de données de vente pour leur permettre de prendre des décisions mieux informées.*
- C'est aussi une opportunité pour les organisations publiques. L'opportunité de construire de meilleures relations entre citoyens et gouvernants, grâce à l'ouverture des données publiques. Le mouvement open data permet de fertiliser l'innovation du quotidien, d'incarner un élan vers plus de transparence, et de favoriser une co-construction des politiques publiques avec les citoyens.
  - *Ex. La ville de New York a non seulement ouvert des jeux de données significatifs mais a aussi créé les conditions pour que des initiatives innovantes en sortent.*
- C'est aussi une opportunité pour les villes. L'opportunité de devenir une *smart city*, une ville qui répond aux défis du développement humain, qui répond aux tensions sur les ressources naturelles, par une gestion optimisée de ses réseaux grâce à l'analyse des données.
  - *ex. La ville de Barcelone a été classée parmi les villes les plus intelligentes du monde grâce à la mise en place de technologies de pointe qui permettent grâce à des capteurs de mieux gérer le trafic routier et l'éclairage urbain.*
- C'est une opportunité pour les citoyens enfin, pour des habitants qui auront peut-être un plus grand confort

de vie grâce à une série d'appareils intelligents, de robots assistants, des outils qui vont simplifier le quotidien.

- *ex. On peut regarder du côté des thermostats et des compteurs intelligents qui aident les habitants d'une maison à optimiser leur consommation d'électricité et faire des économies d'énergie, grâce aux données collectées sur un foyer.*
- C'est aussi, de façon très générale, l'opportunité d'une meilleure connaissance du monde, et de soi, parce que les données peuvent nous permettre de renseigner des choses qui n'étaient pas observables jusque là. De la même manière que nos oreilles ne perçoivent pas tous les sons, il y a des choses, des événements dans notre monde, que nous ne pouvons pas qualifier, ou quantifier, sans l'analyse des données.

Un monde d'opportunités donc, qui devrait s'offrir à nous.

Pourtant, ce qu'on ressent aujourd'hui, ce qu'on lit dans les médias, et de la bouche de nos concitoyens, ce n'est pas ce nouveau monde des données, plein de promesses, mais des constats beaucoup plus pessimistes sur la façon dont sont aujourd'hui utilisées les données :

Les récents événements américains nous en donnent plusieurs signes :

- Premier symptôme : le résultat inattendu de l'élection présidentielle américaine avec l'élection de Donald Trump, qui a jeté un énorme discrédit sur les sondages, plus généralement sur toutes les méthodes d'analyses prédictives, et notre capacité à anticiper grâce aux données.
  - Effectivement, c'est problématique si au lieu de nous aider à mieux comprendre le monde, les données nous éloignent de la réalité de terrain.
- Deuxième symptôme : à la suite de cette élection, l'algorithme de Facebook a été fortement critiqué. Il a été accusé de favoriser la propagation de fausses informations, et d'encourager par là le populisme.
  - Effectivement, c'est un phénomène qui a été appelé les "bulles de filtres" : le fait que les contenus proposés aux internautes sont de plus en plus finement personnalisés grâce aux données, à partir des données personnelles collectées sur ces internautes lors de leur consultation d'information. Ces bulles de filtres sont accusées d'empêcher les internautes d'être confrontés à d'autres idées que les leurs, en proposant prioritairement du contenu censé plaire à l'internaute.
  - Là aussi, il semblerait que l'utilisation faite des données ne serve pas une meilleure connaissance du monde, mais au contraire son rétrécissement.
- Troisième symptôme : l'élection américaine a aussi mis en lumière notre difficulté à faire société. C'est un constat qu'on retrouve aussi en Europe, et en France. Je ne rentrerai pas dans les détails des mécanismes sociologiques ou philosophiques à l'œuvre, mais du point de vue des données. Nous avons aujourd'hui la capacité de regarder les données à un niveau beaucoup plus fin qu'auparavant, si bien que l'on peut faire des statistiques à une échelle individuelle. Dès lors, ça n'a plus beaucoup de sens de parler de grande catégorie de population, de classe sociale ou de nation. Quand on regarde les données à un niveau aussi fin, on trouve une infinité de *patterns*, de motifs individuels.
- Au delà de l'élection américaine, il y a beaucoup d'autres signaux révélateurs : comme par exemples les vagues de protestations des utilisateurs de Facebook lorsque ce dernier met à jour l'algorithme de présentation des news dans le *feed* des utilisateurs, ou lorsque Google modifie son algorithme de *ranking*.

- Ces grands géants numériques utilisent et monétisent des données personnelles, tirées du comportements des internautes, sans forcément apporter la transparence, la déontologie et la pédagogie nécessaires à ce que ça ne soit pas vécu par les utilisateurs finaux comme des démarches intrusives, et problématiques pour leur vie privée et leur droit à l'oubli.
- Enfin, des questions éthiques se posent aussi aujourd'hui, et de plus en plus fortement.
  - Avec par exemple l'avènement des voitures autonomes : lorsqu'un algorithme devra juger de la conduite à tenir en cas d'accident inévitable, et de l'individu à sauver en priorité. Qu'advient-il de nos croyances morales, de notre libre-arbitre dans ces cas-là ?

Voilà pourquoi nous avons un problème aujourd'hui : nous sommes en train de construire un monde que nous comprenons de moins en moins. Il est en train de se dessiner un trou noir entre les données et le savoir : notre capacité à mettre en place des technologies intelligentes dépasse de très loin notre capacité à comprendre comment fonctionnent les algorithmes qui les alimentent.

Et personne ne veut vivre entouré de robots et d'algorithmes qui se comportent comme des boîtes noires. Cette perspective d'un monde dans lequel nous ne comprenons plus les informations qui s'y échangent est inquiétante.

C'est bien pour cela qu'il est grand temps de réagir, et de changer notre approche des données et des systèmes qui utilisent de grandes masses de données :

- il faut changer d'approche sur la question du sens tout d'abord, se donner les moyens de donner du sens aux données. Le langage des données est souvent le langage du code et des mathématiques. Il faut trouver quelque chose de plus heuristique. Un langage qui parle aux humains. Il faut mettre en place une meilleure *data-literacy*, c'est à dire une meilleure capacité à parler et à comprendre le langage des données.
  - Par exemple, la visualisation de données qui permet de rendre compréhensibles les informations contenues dans les données et d'appréhender des phénomènes compliqués en s'appuyant sur d'autres chose que la pure rationalité logique ou verbale.
  - ex. Lorsque nous créons une plateforme de visualisation de données qui permet au grand public en France de comprendre les données d'évaluation de la qualité des soins dans les hôpitaux, qui étaient jusque là des données très technocratiques et compliquées.*
- il faut changer d'approche sur la question des usages : faire en sorte que les nouveaux usages des données soient pensés pour augmenter l'humain, pour améliorer sa capacité à connaître le monde, et à prendre des décisions éclairées, plutôt que de construire des dispositifs qui prennent des décisions à notre place.
  - *C'est le cas par exemple lorsque nous produisons des cartes qui reflètent en temps réel l'état d'un réseau de métro, et les options qu'offre la ville autour, pour aider les voyageurs dans leurs déplacements*
  - C'est vrai à l'échelle de tout un chacun mais c'est aussi vrai à l'échelle des entreprises. La question de l'usage aujourd'hui est ce qui explique que bien des entreprises se soient équipées en technologies pour valoriser leurs données sans pour autant en constater les effets sur leur productivité. C'est un paradoxe qui a déjà été observé dans les entreprises à la fin des années 80 par Robert Solow sous le nom de paradoxe de la productivité et qui s'appuyait sur le constat que l'informatique était partout dans les entreprises sauf dans les statistiques de productivité. Car tant que ça ne passe pas dans les usages, les investissements dans les « Big data » n'auront pas d'effet.



Ce qui compte pour nous, chez Dataveyes, lorsque nous travaillons pour de grands groupes, c'est par exemple de créer des outils qui permettent aux salariés de prendre véritablement en main cette nouvelle source d'information.

- Enfin, il faut changer d'approche sur la gouvernance et l'éthique des données : cela signifie sans doute encadrer, produire des normes, ce qui ne pourra se faire qu'à l'échelle internationale. Et c'est là que je me tourne vers vous, car vous, dans cette salle, êtes bien mieux placés que moi pour que cela se réalise.