



## « Big Data : quels changements pour les cadres ? »

*La présidente de l'OdC, Soraya Duboc, présente le séminaire.*



Après avoir salué les présents dans l'auditorium et sur la Toile, et remercié l'AG2R qui nous permet d'être ensemble ce soir dans ces locaux, **Soraya Duboc**, présidente de l'Observatoire des Cadres, présente le séminaire de ce soir :

« Au nom de toute l'équipe de l'Observatoire des Cadres, j'ai le plaisir de vous accueillir pour ce nouveau séminaire, sur la thématique du *Big data*.

Le phénomène *Big data* - mégadonnées devrions-nous dire en cette semaine de la langue française et de la francophonie - participe de la transition numérique en cours dans tous les secteurs

d'activités, sans exception.

Chacun de nous le constate : cette transition se caractérise par une dématérialisation croissante non seulement des services mais aussi des industries, par la connexion progressive de nombreux objets (internet des objets), et par l'apparition de nouveaux modèles d'affaires issus des plateformes.

Elle présente un caractère massif et quasi inéluctable, doublée d'une incertitude forte sur le volume, la qualité et la consistance des emplois à venir. À quoi ressemblera le travail demain et plus spécifiquement celui des cadres et du management ?

Le sujet est crucial et fait l'objet de nombreux et intéressants rapports et études - rapport Mettling "Transformation numérique et vie au travail" de septembre 2015, rapport du Conseil national du numérique (CNnum) "Travail, emploi, numérique, les nouvelles trajectoires" de janvier 2016, le document de travail de France Stratégie "L'avenir du travail : quelles redéfinitions de l'emploi, des statuts et des protections", publié le 10 mars 2016 -.

Mais ce soir, nous allons nous confronter à un aspect de ces transformations, au cœur des organisations et des fonctions occupées par les managers. Nous allons tenter de mieux comprendre les effets de la numérisation et des *data* sur le travail des cadres et sur le management.

Pour éclairer cette question d'avenir, au centre des missions et de la dynamique de l'Observatoire des Cadres et du management, nous faisons appel à Dominique Cardon, sociologue, professeur associé à l'Université de Marne-la-Vallée, auteur de *A quoi rêvent les algorithmes, nos vies à l'heure des big data*, La République des Idées, Seuil, 2015, et à Sylvie Joseph, experte en *data management* et élue CFDT, qui seront interrogés par Jean-Marie- Bergère, membre du Conseil Scientifique de l'OdC. »



### AlphaGo n'est pas venu

**Jean-Marie Bergère** nous dit avec un sourire qu'il aurait bien invité AlphaGo, le programme informatique « qui fait preuve de qualités humaines » et qui vient de remporter quatre parties de jeu de go contre un champion coréen, car le jeu de go nécessite stratégie, vision, astuce, imagination, c'est-à-dire les qualités que l'on prête aux managers et aux cadres. Mais AlphaGo fêtait sa victoire dans un bar de Séoul avec un des créateurs de la société DeepMind, son concepteur, qui est non pas ingénieur en informatique mais docteur en neurosciences. L'animateur n'a pas non plus laissé sa

place à un robot humanoïde qui l'aurait avantageusement remplacé en analysant en direct les attentes du public voire en les anticipant. Il n'y aura donc que des humains dans la soirée.

### Le Big Data est une rupture

Le premier d'entre eux à intervenir est Sylvie Joseph, directrice de la transformation interne à la direction numérique de La Poste.

Le *Big Data* aujourd'hui, explique **Sylvie Joseph**, n'est plus seulement la manipulation d'une grande masse de données pour de moins en moins cher mais un ensemble de processus, qui se raccorde à l'intelligence artificielle et à la machine apprenante (*machine learning*). C'est une rupture par rapport à ce qu'on connaît depuis une trentaine ou une quarantaine d'années.

On malaxe de la donnée en masse depuis longtemps, ainsi la grande distribution analyse les données de vente et de clientèle depuis vingt ans ; ce qui change aujourd'hui, c'est la capacité de le faire à la nanoseconde sur de très gros volumes, et d'intégrer des données non structurées telles que les traces laissées sur les réseaux sociaux par exemple, des données beaucoup plus sémantiques.

C'est vraiment un tout et la rupture se produit en ce moment, l'exemple du jeu de go illustre bien ce qui va se passer, la conjonction de l'intelligence artificielle et de l'intelligence humaine. Avec encore pour un certain temps une légère différence entre elles, il y en a une qui *a priori* pour le moment n'éprouve pas d'émotions et l'autre encore un peu, c'est notre cas, ce qui fait toute la différence.

Il faut cependant demeurer prudent, par rapport à la data et à l'analyse de la data et à la capacité des machines, selon l'adage bien connu GIGO (*garbage in, garbage out, i.e.* foutaises à l'entrée, foutaises à la sortie). Si vous mettez de la mauvaise donnée à l'entrée, vous aurez de mauvaises corrélations et de mauvaises données de sortie.

### Libre arbitre versus pensée unique

L'intelligence humaine, la prise de recul, l'analyse stratégique, l'analyse tout court, sont extrêmement importantes, et surtout le libre arbitre. L'humain, c'est la vigilance et l'analyse.

Il faut aussi prendre garde à ce que l'intelligence artificielle ne conduise pas au clonage des comportements. Les analyses de données produisent des recommandations, ce qui a tendance à emmener les gens dans un tunnel, dans une espèce de pensée unique, quelle qu'elle soit, dont on ne sort pas.

Jean-Marie Bergère approuve, il a entendu récemment « la merveilleuse expression d'endogamie algorithmique ».

### Des ruptures techniques qui ont des conséquences sur les compétences

Néanmoins, reprend Sylvie Joseph, il existe aujourd'hui de grosses ruptures techniques, pas uniquement pour affronter des joueurs de go sud-coréens, mais aussi prosaïquement dans la vie de l'entreprise de tous les jours, par exemple pouvoir faire une requête en langage naturel sur les outils de gestion. Les requêtes SQL (*Structured Query Language*, en français *langage de requête structurée*,

un langage informatique normalisé servant à exploiter des bases de données relationnelles) et autres du même genre consistaient un langage très ésotérique ; aujourd'hui vous posez la question de la manière dont vous parlez et vous avez votre réponse. Cela veut dire qu'il est aujourd'hui possible de faire du contrôle de gestion, de l'analyse, sans être un spécialiste de la requête SQL. C'est un très gros changement y compris en matière de compétences internes dans l'entreprise. Ces nouveautés, qui proviennent d'éditeurs essentiellement américains, de type Microsoft ou IBM, permettent de regarder l'organisation de l'entreprise d'une autre façon.

Les outils dont on dispose aujourd'hui permettent de plus en plus de projeter des données, de malaxer des données de nature extrêmement différente, pour essayer, à travers d'algorithmes, d'en déduire un certain nombre de comportements prédictifs. Et ces comportements prédictifs sont des formes de scénarii construits très rapidement qui permettent de poser un diagnostic et surtout de mener un travail de stratégie, ce qui devrait encore pour un certain temps appartenir à l'humain.

### *Des outils encore classiques mais omniprésents*



La question posée était celle de l'état de l'art : quelle est l'utilisation de ces outils dans l'entreprise ?, sachant que par outils on entend les données et les dispositifs qui permettent de malaxer celles-ci. La réponse est qu'aujourd'hui on en est encore à des outils de pilotage relativement classiques, mais qui en revanche touchent à peu près tous les compartiments de l'entreprise.

Prenons quelques exemples. Dans le domaine de la finance et de la gestion, il existe évidemment les ERP tels que SAP, etc. Ces ERP (*enterprise resource planning*, en français *progiciels de gestion intégrée*) sont aujourd'hui nourris d'énormément de données, ils permettent d'affiner les organisations, les gains de productivité, etc. Il existe aussi dans le domaine de la finance, d'ailleurs certains s'en plaignent beaucoup, des outils qui permettent de miser sur des actions, de sortir et de prendre des plus-values - ou des moins-values, cela dépend - de manière extrêmement rapide, à la nanoseconde, uniquement par l'action des robots, et donc de l'analyse de *data*. Il y a donc tout un marché financier qui fonctionne ainsi, avec ses bons et ses mauvais côtés.

### *Des possibilités immenses*

Ce type d'outil est de plus en plus utilisé dans la gestion de la production. Par exemple, la Poste a un outil appelé TAE qui permet de lire toutes les enveloppes qui passent dans les machines de tri. On l'utilise aujourd'hui exclusivement pour la gestion de production de tri et de redressement d'adresse. Tous les expéditeurs n'écrivant pas lisiblement ni ne remplissant pas correctement les adresses, TAE permet de redresser les erreurs en quelques secondes et ainsi d'avoir une qualité de service très importante. Mais si La Poste le voulait, elle pourrait aussi faire, par exemple, un observatoire de la pression publicitaire à travers le marketing direct papier ou le courrier des facturiers, par département, par localisation, par groupe IRIS, etc., et dresser une cartographie de la France sur ces sujets. Ce n'est pas ce qui intéresse aujourd'hui cette entreprise mais elle est en mesure de le faire.

Grâce à ces outils, on trouve de plus en plus des pilotages de maintenance prédictive, notamment dans l'aéronautique. On n'arrête plus les avions à des dates prévues, on les arrête quand un certain nombre d'indicateurs déclenchent un seuil de maintenance, ce qui permet de faire des économies substantielles tout en gérant la sécurité.

### *L'analyse devient prédictive*

Dans un domaine qui est aujourd'hui quasi banalisé comme celui de la gestion commerciale, du marketing, et de la relation client, on utilise évidemment les *data* depuis très longtemps. Ce qui est en

train de changer, ce sont les sujets autour de l'Intelligence artificielle et surtout de l'analyse sémantique.

On a donc aujourd'hui la capacité à traiter des données structurées et même, ce qui est le plus important, non structurées, et par conséquent la possibilité de produire des profils, des typologies d'individu et de comportement, à la fois avérés et prédictifs. Ce côté prédictif est très intéressant car il permet, quand les données sont bien utilisées, d'améliorer le taux de concrétisation des ventes et la performance commerciale en général, ainsi que l'amélioration de la relation client. On peut même faire traiter partiellement cette relation client, non par des humains mais par des robots. Les machines peuvent vous envoyer des messages, et elles peuvent le faire de manière sémantique et intelligente. D'ailleurs, les Français sont très bons dans le domaine de l'analyse sémantique.

Comme illustration de ceci, on peut citer la toute récente intervention du président d'Orange au *Hello show* - la présentation des innovations de l'année - . Il a beaucoup insisté sur ce qu'apporte l'intelligence artificielle connectée au réseau et au contenu : l'analyse du contenu télévisuel permettra d'enrichir l'expérience utilisateur de chaque client à travers la *box* et le téléviseur de celui-ci, expérience approfondie en temps réel de façon extrêmement simple.

### *Enchères et diagnostics*

Pour rester dans le domaine de la communication, les achats média sur les sites Internet utilisent énormément, comme le fait aussi la finance, le *real time bidding* (RTB, enchères en temps réel en français). Il s'agit de ventes aux enchères qui sont connectées directement aux audiences web. Le prix payé par l'annonceur, à la seconde ou la nanoseconde, est fonction de l'audience au moment où passe sa publicité. Tout est robotisé et se passe à l'échelle mondiale. Quasiment toute la partie numérique de la communication, de l'achat d'espace, fonctionne ainsi.

Dominique Cardon précise que le marché digital est divisé en deux : d'une part le marché du *display* sur les *homes* des pages (affichage sur la page d'accueil du site) qui est toujours CPM (*cost per mille* ou coût par mille pages vues), c'est celui qui rapporte de l'argent ; d'autre part, un autre, représentant 30% du marché total, qui est *back* en RTB (*real-time bidding* - enchère en temps réel), et plutôt sur les pages profondes des sites web. Et lui il n'a pas encore colonisé la *home*. Pas encore ! souligne Sylvie Joseph.

Le domaine de la santé utilise de plus en plus l'analyse de la *data* et les systèmes prédictifs. Aujourd'hui, dans le diagnostic de certains types de cancer et surtout dans la capacité à cibler de manière extrêmement fine les traitements adaptés au cas précis de l'individu, à sa morphologie, à sa pathologie, les machines sont plus performantes que le meilleur des médecins.

Du mauvais côté des choses, la guerre est elle aussi devenue en grande partie électronique, que ce soit en matière de renseignement qui malaxe beaucoup de *data*, en particulier la guerre contre le terrorisme, ou le tir à distance avec le pilotage de drone qui permet à un homme de bombarder un pays en manœuvrant une manette de son bureau aux États-Unis.

### *Prévoir les comportements, adapter les processus*

Quels sont les impacts du *big data* sur le travail, des cadres et des salariés en général ?

On se trouve devant une émergence d'expérimentations. Par exemple, les data R.H. permettent de regarder ce qui s'est passé dans la vie de l'entreprise, grâce à des indicateurs et d'essayer de réaliser des projections, de prévoir des comportements.

Si on peut prévoir l'absentéisme, on pourra, quand arrive une période d'absentéisme massif, réorganiser les équipes par anticipation. Cela soulagera les cadres et les encadrants, lesquels sont plus à l'aise quand ils peuvent anticiper que lorsqu'ils doivent réagir à chaud et risquent de se trouver dans l'incapacité de mettre en place des équipes de substitution. Cela diminue les risques psychosociaux des encadrants et a donc des vertus positives.

Qui dit *data* dit traçabilité, qu'elle soit géographique ou de connexion. Et la traçabilité est un vrai sujet. Cela peut être pour le meilleur, lorsqu'un calcul de temps de travail alerte le cadre qui dépasse trop ses horaires et génère un droit à la déconnexion ou quand il permet au manager d'accompagner et de conseiller ses collaborateurs sur la façon dont ils répartissent leur temps, afin de les rendre plus efficaces tout en ayant un meilleur partage entre temps professionnel et temps personnel. Il est bien de factualiser, d'identifier, de quantifier l'utilisation du temps. Ça peut être aussi pour le pire, avec cette notion souvent redoutée de « flicage ».

On a aujourd'hui de plus en plus d'outils de mesure qui permettent d'organiser ou de réorganiser l'allocation des ressources ; ils permettent aussi de travailler sur une adaptation des processus, notamment en regardant comment se comportent les clients dans leur parcours ou les collaborateurs qui utilisent ces outils et processus.

En analysant de manière extrêmement précise les processus, avec les *data* et les machines, on peut trouver des solutions qui soient pertinentes pour les deux parties.

### **Baromètre et organisateur**

En matière de relations sociales, par exemple, on commence à voir apparaître ces analyses, à travers les réseaux sociaux internes d'entreprise ou des applications spécialisées. Une application appelée Teamood est en expérimentation assez avancée, c'est un baromètre d'humeur sociale. Les salariés témoignent par quelques clics de la manière dont s'est passée leur journée, de manière anonyme d'ailleurs, cela peut donner lieu à analyser des sociogrammes dans l'entreprise : qui est relation avec qui, qui parle avec qui, qui parle de quoi. Les réseaux sociaux internes permettent de faire émerger des signaux faibles. On est aujourd'hui dans des phases d'expérimentation mais un champ nouveau d'analyse R.H. et d'études sociales apparaît.

Pour ce qui est de l'organisation du temps, notons que des éditeurs américains sortent des plateformes bureautiques et collaboratives contenant de l'intelligence artificielle et donc du *machine learning*. Ces plateformes commencent à sortir sur le marché, elles devraient permettre de soulager le travail du collaborateur, en particulier du cadre, dans l'organisation de son temps, de ses tâches, de ses contacts, de ses priorisations, etc. En sachant qu'il est toujours possible de débrancher la machine, de ne pas suivre ces recommandations, de garder son libre arbitre, ce qui est le cœur du sujet. La sortie de ces outils sur le marché induit une vraie transformation, un vrai changement voire même une vraie rupture par rapport à ce que nous connaissions jusqu'à présent en matière de messagerie, d'agenda, etc.

### **Automatiser le reporting**

Et puis on voit émerger en matière de contrôle de gestion cette intelligence artificielle qui permet donc de faire des requêtes en langage naturel, ce qui ouvre des perspectives en matière d'allocation de ressources et de formation de collaborateurs. Certains, qui n'étaient pas profilés pour être, par exemple, contrôleur de gestion, pourraient petit à petit le devenir avec des formations beaucoup plus légères que ce qu'on aurait pu imaginer. Cela permet aussi, et peut-être même surtout, d'automatiser une grande partie du *reporting* qui aujourd'hui prend beaucoup de temps, que ce soit pour les cadres ou pour certaines catégories de collaborateurs. Souvent d'ailleurs il faut répéter des tâches qui demandées par plusieurs services ou différents niveaux dans l'entreprise, le fait de pouvoir automatiser toutes ces tâches permet de libérer du temps et de réinvestir ce temps plus intelligemment sur des tâches à plus forte valeur ajoutée, de créativité, d'analyse stratégique, de plan d'action, de mise en œuvre opérationnelle, de pilotage de la qualité, etc.

### **Dans la carrière et le recrutement**

En matière de formation et de mobilité professionnelle, on note l'apparition sur le marché d'entreprises qui sont en capacité de cartographier ou de faire cartographier à l'intéressé lui-même

ses compétences, ses appétences, ses potentiels, etc. Ce ne sont plus des *start-up* mais des entreprises qui commencent à remporter des succès commerciaux patents

En captant les données laissées par vous sur les réseaux sociaux professionnels de type LinkedIn ou Viadeo, en analysant ce que vous dites de vous-même, on peut établir des comparaisons, du *benchmark*. Ces masses de données extrêmement importantes – en France il y a plus de dix millions d'inscrits à LinkedIn – permettent d'établir des bases de données très solides et très riches, que les entreprises ne possèdent pas, le plus souvent. À travers ces éléments, il devient de plus en plus possible de cartographier vos compétences et de recommander de manière automatique ou semi-automatique des formations adaptées pour à la fois acquérir les compétences qui vous manquent et répondre aux attentes que vous avez pu évoquer, par exemple dans le cadre de votre évaluation annuelle.

Pôle Emploi s'intéresse de très près au sujet, il mène quelques expérimentations dont l'une près de Cholet. On utilise l'analyse, les *big data* et la *data visualisation* pour essayer de rendre beaucoup plus rapide, concrète et performante la jonction -ou la corrélation- entre un demandeur d'emploi qui a un certain nombre de compétences à promouvoir et des postes qui pourraient lui être ouverts mais pas forcément dans ce qu'il imaginait être possible pour lui. On va chercher des écarts et donc des *scorings* qui permettent de dire « vous n'allez pas penser à ce type de poste mais il se trouve que par rapport aux compétences que vous avez, vous correspondez à 60, 70%. Et voilà ce qu'il vous faut faire pour arriver aux 100%. » Ce n'est pas très compliqué et la personne peut alors prétendre plus pertinemment au poste, on maximalise les chances du candidat de retrouver un emploi plus rapidement. Pôle Emploi est évidemment intéressé par ces logiques.

### Un équilibre pas simple à trouver

Les cadres sont-ils dépossédés de leur rôle ? Ce n'est pas le cas encore actuellement mais on est en pleine transformation, et effectivement le risque est grand de les voir dépossédés d'un certain nombre de leurs attributions d'aujourd'hui, voire même d'être complètement happés par ces logiques de machines, de corrélations et de prédictions, et d'y perdre finalement la substantifique moelle de leur rôle. Néanmoins, le rôle d'un cadre est toujours le même et devrait le rester, c'est un rôle de prise de recul, d'analyse, d'encadrement, de développement des collaborateurs, et à cela aujourd'hui la machine ne peut y répondre.

En tout cas, il ne faut pas être pessimiste, c'est potentiellement une grande chance pour les cadres de pouvoir disposer d'outils de ce type, à condition qu'ils sachent ce qu'on peut en tirer et surtout comment structurer les choses en amont pour avoir de la qualité d'analyse en aval. Ils doivent aussi réinvestir du temps sur ce qui est le cœur du métier du cadre, le sens, la vision, la stratégie, la créativité, l'accompagnement et l'animation des équipes. Néanmoins, il faut prendre garde à ne pas tomber dans l'abêtissement, dans l'asservissement à la machine ; il faut doser le délicat équilibre entre les bénéfices d'usage et le renoncement à sa liberté personnelle (*privacy*), qui ne sera plus jamais telle qu'on l'a connue.

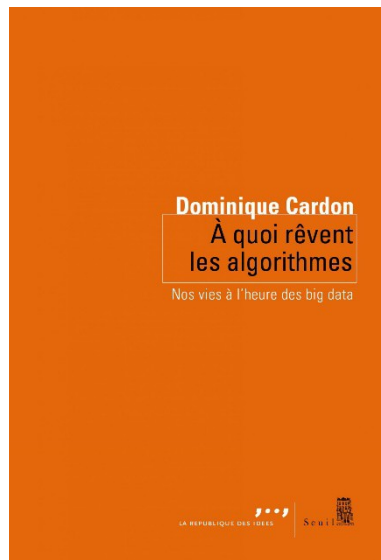
Le cadre peut réaffirmer sa fonction mais, dans un environnement numérique et de machines intelligentes, cela passe probablement aussi par une révision de sa formation de base. Un sujet sur lequel on devrait travailler, particulièrement en France, car ce n'est pas particulièrement développé dans les élites françaises qui se forment dans les grandes écoles d'ingénieur ou de commerce, c'est la connaissance des sciences humaines, de la philosophie, de la sociologie. Un point extrêmement important pour vivre et survivre intelligemment dans le monde numérique, c'est l'éthique.

### Échange avec les organisations syndicales

Mais quel est l'accueil réservé à de telles analyses de la part de La Poste, demande Jean-Marie Bergère, avez-vous un mandat fort de la direction générale, y a-t-il une forte demande de la part des cadres et des collaborateurs ?

Le but du programme est d'utiliser les outils informatiques, expose Sylvie Joseph, qui a suivi un parcours marketing, ce qui lui permet de prendre les questions par l'usage et non par les outils. Elle travaille particulièrement à mettre en symétrie les intentions du collaborateur et celles du client, ce qui est la volonté de la Direction générale. Nous échangeons régulièrement avec les organisations syndicales sur nos expérimentations. L'utilisation de la donnée rentre évidemment dans le cadre de ces discussions.

### Quel rôle pour les cadres dans une société collaborative ?



**Dominique Cardon**, sociologue à Orange Labs et professeur associé à Marne-la-Vallée, a travaillé autrefois en entreprise ; il vient de publier *A quoi rêvent les algorithmes, nos vies à l'heure du big data*.

En tant que sociologue, Dominique Cardon veut dire au préalable qu'il existe des effets de mode, dus au fait que la Silicon Valley vend un concept nouveau tous les cinq ans. Il existe une réalité derrière mais aussi des promesses loin d'être réalisées, AlphaGo ou pas.

#### *Les données et les artefacts*

Il faut distinguer les données et l'enjeu qu'il y a derrière elles. Les données existent de tout temps, dans la science comme dans l'entreprise ; ce qui est nouveau et original avec le *Big Data*, c'est qu'elles sont beaucoup plus nombreuses, les capteurs qui enregistrent tout donnent une hétérogénéité et un volume de

données considérables. À côté des données, il y a les calculateurs, les algorithmes.

Un débat très clivant présente les hommes et les machines, et une sorte de combat entre eux, c'est une représentation très caricaturale des choses.

Quand on se représente la machine, soit on lui prête une intelligence incroyable, et c'est le cas avec le retour de l'intelligence artificielle, soit on lui prête une bêtise crasse parce qu'elle calcule bêtement, elle rationalise, elle est idiote. Il faudrait savoir si la machine est subtile ou pas subtile ! Mais surtout, ce qu'il est important de souligner, c'est l'idée que les calculateurs sont fabriqués par des équipes d'ingénieurs qui appartiennent à des entreprises qui ont des intérêts, qui ont des logiques, et qui prêtent aux calculateurs des intentions pour produire un monde.

Les artefacts existent depuis les débuts de l'humanité, avec l'écriture et le livre. Le livre a été au cœur de ce transfert de la mémoire aux objets, à l'environnement. Aujourd'hui on en transfère de plus en plus vers des calculateurs numériques et évidemment dans ce processus on produit un environnement dans lequel nous nous plions pour vivre, aimer, rencontrer, etc.

#### *Le système électoral est un algorithme*

La meilleure illustration de ce qu'est un algorithme, c'est l'algorithme électoral. Un système électoral est un algorithme, une suite d'instructions pour transformer une masse de bulletins en une représentation dans une assemblée. Si on choisit le système uninominal à un tour on a une assemblée bipolaire, comme au Royaume-Uni ; en conséquence, les Anglais vivent dans un espace bipartisan qui va structurer leur manière de se positionner, cela forme leur monde. Un Israélien qui vit dans un monde proportionnel est dans un espace fondamentalement multicolore, dans lequel il y a des alliances entre les partis. Ce n'est pas le même espace politique, ce n'est pas la même socialisation ; le système de représentation va structurer différemment l'espace dans lequel la population va vivre.

De la même façon, les choix fait par les concepteurs d'un algorithme vont conduire à structurer différemment le monde dans lequel on vivra.

### *Popularité, autorité, réputation, prédiction*

Dans l'histoire du Web, on a utilisé différentes manière de ranger, de classer l'information, avec des techniques fondamentalement différentes, qui correspondent à des représentations différentes.

L'audience, la mesure du clic du visiteur unique construit un monde avec un centre où chacun dispose d'une voix, c'est le concept de **popularité**. A l'inverse, Google est basé sur le principe d'**autorité**, en se référant à ce qui est cité par des sites qui sont eux-mêmes cités. C'est semblable à la manière dont les chercheurs sont évalués : les articles cités, une reconnaissance par les pairs. Google cherche à mesurer l'autorité comme meilleure approximation de la qualité de l'information.

Un monde organisé par la popularité ne ressemble pas à un monde organisé par l'autorité.

Les métriques du web social, les compteurs d'amis, de viralité, ont décidé qu'il n'y avait de centre, l'individu quand il choisit des amis dessine un espace affinitaire dans lequel chacun va produire sa **réputation** ; la manière de calculer la réputation par les techniques du web, c'est de mesurer les effets que les messages ont sur les autres parce ceux-ci les reprennent (*like, retweets, etc.*). C'est un espace dans lequel la réputation organise la hiérarchie.

La quatrième famille de calcul algorithmique, celle qui nous préoccupe, n'est pas encore très présente dans le monde de l'entreprise. Elle est encore à l'état de promesse mais elle est en train de bouleverser les trois autres familles, de s'intégrer dans les trois autres familles, popularité, autorité, réputation. Il s'agit de la **prédiction**.



### *Se méfier de la prédiction*

La prédiction ne s'occupe pas de ce que les gens disent mais seulement de leur comportement. Certes, on peut utiliser l'analyse sémantique -ce que disent les gens- pour en tirer du signal mais on n'arrive pas bien à prévoir. Pour le moment, on ne sait pas encore prévoir le résultat des élections, le cours de Bourse, la réussite d'un film, par exemple. Si on écoutait Twitter, l'Ecosse serait indépendante, parce que ceux qui étaient pour l'indépendance le disaient et que ceux qui étaient contre ne le disaient pas, c'est ce qu'on appelle l'auto-sélection. Autre exemple, vingt pour cent des

personnes interrogées disent regarder Arte, dans les statistiques, c'est trois pour cent.

Ce décalage-là doit être rattrapé par l'étude des comportements, des traces de navigation. Avec le RTB, on s'occupe du comportement, on se place « sous » les données, on essaie de ne pas être identifié par les utilisateurs pour éviter la mise en place d'une boucle réflexible. En effet, dès que les individus savent qu'ils sont enregistrés par une métrique, pour un palmarès, ils changent de comportement.

### *La troisième vie de l'intelligence artificielle*

Une autre chose importante est la « nouvelle » technique de l'apprentissage des machines. En réalité ce n'est pas nouveau du tout, on peut faire remonter son origine à Thomas Bayes, un révérend anglais du XVIIe siècle. Mais les nouvelles capacités de calcul permettent de réaliser un modèle de prévisions subjectives qui n'était pas possible auparavant ; il y a là un véritable effet technologique.

Ce qu'on appelle l'intelligence artificielle en est à sa troisième vie : dans les années cinquante, avec la cybernétique, tout le monde parlait d'intelligence artificielle, c'est mort dans les années soixante-dix, cela revit dans les années quatre-vingt, avant de re-disparaître provisoirement dans les quatre-vingt-dix. Qu'est ce qui s'est joué avec ces naissances et morts de l'intelligence artificielle ?



L'idée anthropomorphe est de confier à la machine le raisonnement humain. Dans les années quatre-vingt, on apprend les règles symboliques à la machine, et par symboliques on entendait que la machine peut comprendre les règles, on va lui apprendre toutes les règles et puis elle va prendre les bonnes décisions parce qu'elle aura un système expert. Ce programme a échoué car les variations de contexte sont trop complexes. La machine ne peut pas tout prévoir, en appliquant la règle elle devient idiote. Cela se passe même dans des situations sociales qui sont très réglées comme le diagnostic médical.

Face à cet échec, une autre stratégie s'est développée pendant les années quatre-vingt-dix et deux mille : on n'apprend pas à la machine les règles symboliques mais on se base sur les statistiques. Les machines apprennent des micro-règles, de petits *patterns* (modèles) statistiques qui se rétro-propagent. La machine apprend un modèle que nous ne pouvons plus comprendre, elle a seulement besoin que l'on lui donne des données en entrée et un objectif. Par exemple, on entre dix mille photos de chats, la machine va faire tourner et après elle trouvera les chats dans les autres photos. Et ça, ça marche.

Jouer au go est extrêmement simple parce qu'il n'a pas de contexte, si la machine jouer des millions de parties contre elle-même, elle peut apprendre, parce que l'espace de déplacement des pièces est un espace limité, immense mais limité.

Dans l'entreprise, il va falloir donner au calculateur des signaux très nombreux afin de produire des prédictions justes. La tâche du cadre, la stratégie, la complexité, font partie du contexte. Il est exact que le numérique est partout et capte beaucoup de signaux, mais toute notre vie n'est pas dans les signaux numériques.

### **Boucler en temps réel**

Ce qui est nouveau dans le *big data*, c'est du temps réel, une base de données et un signal en boucle qui dit si on a été efficace, ce qui permet de réviser le modèle. Le système boucle constamment. Google sait si pour tel mot-clé les gens ont cliqué sur le premier, le deuxième, le troisième, le quatrième lien, c'est une information absolument décisive parce que cela lui donne la réponse à l'efficacité de sa prédiction. Vous avez cliqué sur le troisième lien et n'êtes pas revenu sur la page, cela veut dire que le troisième lien est sans doute le meilleur ; Google va alors réviser le modèle pour les suivants ou pour le personnaliser. Cette boucle constante dans lequel le modèle peut utiliser une information pour se réviser marche pour la recommandation musicale, mais c'est beaucoup plus difficile quand on a des données à froid où on n'a pas la mesure de l'efficacité de la prédiction à réintégrer constamment dans le modèle.

On va vers une accélération, une intensification des processus de rationalisation, on se met à tout compter pour mettre tout en indicateurs mais on continue à produire des indicateurs qui sont à moitié vrais à moitié faux, toujours incertains, souvent contestés. Si on a pas le moyen de les valider dans une boucle, on ne fait pas vraiment de la *big data* prédictive, on est simplement en train de poursuivre un travail de rationalisation du travail salarié et du coup peut-être d'augmenter pour les cadres des charges de *monitoring*, de *reporting*, plutôt que des tâches stratégiques. Sans vouloir être trop pessimiste, cela peut-être un risque associé au déploiement d'une logique de calcul qui est assez proche des ERP++, finalement.

### **Débat**

Observe-t-on des réseaux professionnels sauvages ? demande Jean-Marie Bergère. Bien sûr, répond Sylvie Joseph, c'est le cas s'il existe un besoin et que l'entreprise n'offre pas les bons outils. À La Poste, il existe des innovateurs qui utilisent aussi un outil ouvert sur l'extérieur. Il faut se méfier de l'usage brutal de la déconnexion, en empêchant les cadres de se connecter la nuit, on a créé des embouteillages le matin. Il faut éduquer les gens à se déconnecter quand ils le veulent.

Lors du débat avec la salle, Dominique Cardon explique que la prédiction équivaut à un calcul probabiliste ; c'est uniquement dans la tête des analystes qu'il y a causalité ! Le modèle comportementaliste montre que nous, consommateurs ou salariés, sommes très réguliers dans nos comportements ; la causalité est la croyance que nous mettons sur la corrélation.

Comme le fait remarquer un participant, les traducteurs automatiques ont commencé à fonctionner à peu près correctement quand ils se sont basés sur les statistiques.

La Poste peut aujourd'hui proposer de nouveaux services ou une nouvelle façon de rendre des services, explique Sylvie Joseph, ce qui crée de la valeur. L'entreprise effectue des gains de productivité en s'organisant différemment, par exemple l'automatisation des tâches simples permet de vendre de la poste mobile. Dominique Cardon précise que cela crée un peu de valeur financière mais surtout de la valeur d'usage.

Des écoles d'ingénieur mettent en place des formations unissant *Big data* et sciences humaines, ce n'est que rarement le cas dans les entreprises. Quand on parle de *Big data*, parle-t-on de données brutes, de données traitées ? Les bases de données sont-elles du *Big data* ? Comme disait Bachelard, les données ne sont pas données, elles sont construites. Et il n'y a pas tant des données, structurées ou non, que les algorithmes. La Poste met en œuvre des codes éthiques, avec une vraie volonté de donner des garanties, mais la formation épistémologique est peut-être plus importante que la formation éthique.

### Éthique, gouvernance et déontologie

Aujourd'hui, la base de données unique et le *Big data* sont des sujets parallèles, dans un cas comme dans l'autre, ils ne sont pas destinés à définir la stratégie de l'entreprise mais à éclairer le conseil d'administration. Disposant d'un ensemble de faits partagé, le Conseil se rapproche de la vie opérationnelle de l'entreprise. La disponibilité des informations est un enjeu de démocratie dans l'entreprise, affirme Jean-Marie Bergère, mais l'entreprise est-elle loyale ? Sylvie Joseph n'a pas constaté d'approche malveillante mais pense qu'il peut y avoir incompétence. De toute façon, la question de la bonne gouvernance sur ces enjeux est posée, gouvernance qui devrait intégrer des parties prenantes internes et externes, dont les partenaires sociaux, mais est très compliquée à mettre en œuvre. La question de la déontologie se pose, l'audit interne (ou l'inspection générale) peut en être le garant. Il est important que les partenaires sociaux soient associés à la gouvernance et aient le droit de partager les données pour alimenter leur réflexion et de réaliser leurs propres analyses, dans le cadre des lois existantes.

Néanmoins, dans la plupart des entreprises, les données ne sont pas accessibles aux syndicalistes et aux élus du personnel, il arrive aussi que la direction qui ne veut pas donner l'information la cache par obfuscation, par exemple en noyant les données intéressantes dans un flot d'indications non pertinentes. Dans un certain nombre de cas, le retro-calcul est très utile.

### Cet algorithme est-il loyal ?

Un thème qui monte est celui de la nécessité que la Société soit active sur ces questions, pour vérifier, non pas la neutralité mais la loyauté de l'algorithme. Car il y a des soupçons : Amazon propose-t-il les livres lus par ceux qui ont lu les mêmes livres que vous ou bien les livres dont ce distributeur souhaite écouler les stocks ? Mais face à la perversité des géants du Web, les internautes développent des tactiques de défense, ainsi quarante pour cent des Français utiliseraient des logiciels qui bloquent ou brouillent les bannières publicitaires intrusives.

La question de la régulation est posée. Pour Sylvie Joseph, la CNIL est une forme de dé-régulation et un véritable régulateur serait très utile. En matière de données médicales, la connaissance des

données permet de personnaliser le diagnostic et donc le traitement du cancer mais peut aussi conduire un assureur à ajuster ses tarifs en fonction de l'état de santé de l'assuré. On se trouve devant un sujet de philosophie sociale, de l'ordre du politique. Il n'y a pas de gouvernance structurée ni de comité d'éthique dans les entreprises sur ces sujets, alors même qu'il peut être dangereux de laisser ces sujets aux seuls experts. Quand la directive sur les données personnelles arrivera, elle renforcera la vérification. D'une manière générale, les cadres doivent apprendre à trouver du signal dans le « bruit », apprendre à se méfier des apparentes évidences, à savoir ce qu'il fait réellement. Mais qu'en sera-t-il de la nécessaire prise de risque si l'examen d'une grande quantité de données devient un préalable à la décision ? Le rôle du cadre est justement de prendre des risques en appréciant le contexte, et on peut à ce titre garder de l'espoir.

### *Syndicalisme et contre-calculs*

Des syndicalistes d'IBM ont monté une formation sur le *Big data* afin de s'en servir comme arme syndicale. Cette entreprise comporte 80% de cadres, dont 70% sont des travailleurs du savoir aux ordres, dans une entreprise très dure qui essaie d'éliminer les syndicats. Une expertise par le CHSCT a montré grâce au test du  $\chi^2$  qu'il existait des quotas et une discrimination envers les plus de cinquante ans.

Dominique Cardon souligne que le calcul dépend de l'usage. Le calcul peut combattre d'autres calculs, comme le montre le débat sur le PIB et les indicateurs. Pour sa part, il encouragerait les organisations syndicales à opposer des contre-calculs aux calculs de la direction.



### **Chronique de synthèse, par Anne-Florence Quintin**

Le but de ce premier débat est de « faire atterrir » l'avion Big data, pour éviter l'atterrissage d'urgence. C'est l'objectif de l'OdC que d'éviter les atterrissages d'urgence de théories managériales trop ficelées « là haut », dans le *cloud*.

Le joueur de go a perdu ; la machine a appris et elle a gagné. La « trace » qui apprend si bien a fait perdre le joueur de go. À faire perdre le joueur de go, et demain le salarié, le cadre ? Cet événement est présenté comme historique pour l'humanité.

Tentation de la démission de l'homme, du cadre ? Non, Cet événement nous rappelle juste que c'est pour la machine « que tout est toujours possible », pas pour l'homme.

L'homme a perdu au jeu de go. Et si, comme Dominique Cardon nous y a invités, l'algorithme demandait, sur Twitter, « *Que faut-il donc faire de ce cadre, de ce manager, et comment le faire évoluer ?* »

Il dirait peut-être : « *la machine, la trace d'algorithmes t'égale en imagination et en astuce, Jean Marie.* »

Il dirait aussi : « *moi, algorithme, mes capacités déductives sont infinies mais ma responsabilité managériale inexistante, ma capacité de choix, de trancher, d'ajuster aux situations humaines qui font le « vrai » travail... aucune trace, je ne vois pas.* » Décidément pas de « boucle », de *deep learning*, là. Il dirait peut-être (c'est prédictif) : « *il me manque la parole.* » Avez-vous essayé de parler avec SIRI, le logiciel d'Apple dans des moments de solitude ? Il vous répondra inmanquablement au bout de quelques échanges « *je respecte vos idées* ». Il dira qu'il laisse des traces de signaux infinis, mais des traces implicites (c'est Dominique qui l'a dit). *And this is the question !* Ces traces ne sont que des reflets.

Il dira : « *apprends-moi, cadre, à expliciter l'activité, à la faire exister, en vrai ! À approcher le boulot, celui auquel toi, cadre, toi manager, tu tiens ! Et puis, tant que l'on y est, des pistes sur le collectif, la controverse, celle qui invente l'éthique dont parlait Sylvie, ça m'aiderait à tenir la route. »*

Il dira (notez le glissement - du prédictif il devient prescriptif) : « *Ne me laisse pas seul, cadre. Ne me laisse pas tout seul avec les exxperts !, les experts entre eux, comme dans la finance, disait Sylvie. J'ai bien des amis, de rares amis, très rares, ils sont data scientist. Avec eux, c'est facile, on est en réseau. Ils me créent pour que je vive ma vie. Comme ce n'est pas toujours possible, je pars en révision, ils changent les réglages, ils changent le pilotage dans leur coin, là tous seuls- ils sont seuls aussi eux. Ce serait pas mal, s'ils me présentaient au Conseil d'administration, aux administrateurs salariés, j'aurais de nouveaux amis, le DAF, le DirCom, mais pour l'instant personne ne nous demande rien, alors... »*

Il dira aussi : « *Travaille la manière dont tu peux articuler mes données avec les données sociales. Ne soit pas tétanisé par ma puissance, mes données peuvent être au service de ton expertise syndicale. De là à opposer des contre-calculs...Dominique ? »*

Il dira « *Ne me laisse pas seul, mais ne cherche pas à me contrôler, tu n'y arriveras pas- c'est comme pour ton équipe. L'OdC connaît bien ces phénomènes depuis sa création en 1996. Cherche à m'orienter, à me comprendre, comprendre que je suis un nouvel outil de gestion dans l'entreprise, et qu'il s'agira de l'analyser, voilà une première piste de travail pour l'OdC. »*

Il dira, « *S'il te plaît, cadre, encadre-moi, oriente-moi. On peut s'arranger. Moi, je suis prêt à t'aider à prendre des risques, n'est-ce pas ce que l'on attend de toi en ce moment ? »*

=====