

# Entretien avec Yves GASSOT

*Propos recueillis par Edmond BARANES.*

Yves Gassot a été pendant plus de vingt ans directeur général de l'IDATE DigiWorld. À ce titre, il a participé à de très nombreux travaux sur l'évolution des marchés des télécommunications, et plus largement du numérique, a assuré la direction de la revue *Communications & Stratégies*, du rapport annuel *DigiWorld Yearbook*, et il est l'auteur de nombreux articles. Il a été conseiller spécial de Viviane Reding, Commissaire européenne, pendant la revue du cadre réglementaire des communications électroniques. Il travaille aujourd'hui à la direction générale d'Orange. Yves Gassot est par ailleurs membre associé du Conseil général de l'Économie et du comité stratégique d'Iris Capital. Il a une formation initiale d'architecte (DPLG-Paris), une maîtrise d'urbanisme, et il est diplômé de l'IEP Paris (3<sup>e</sup> cycle d'aménagement).

## **Enjeux numériques [EN] : Qu'appelle-t-on Big Data ?**

**Yves Gassot [Y.G.] :** L'expression Big Data est née à la fin des années 1990 et est donc relativement récente. Elle est l'expression de plusieurs évolutions.

D'abord l'explosion de la production de données, qui combine la croissance continue du nombre d'internautes, amplifiée depuis dix ans par la généralisation des usages de l'Internet mobile, le basculement dans le numérique du monde de la musique et de la vidéo, l'intensification des échanges à travers les réseaux sociaux... Naturellement, l'entreprise est aussi partie prenante de la production de ces données, dans son fonctionnement interne comme dans ses relations avec les fournisseurs ou la clientèle. Enfin, il faut mettre en avant le phénomène émergent de l'Internet des objets (IoT) qui repousse à l'infini les limites des connexions et des flux supportés par l'Internet. Aussi, pour évaluer la production annuelle de données, on ne peut plus parler en gigaoctets (8 milliards de bits), ni même en téraoctets (1 000 gigaoctets) mais en dizaines de zettaoctets (1 000 milliards de gigaoctets<sup>(1)</sup>).

Derrière cette accélération de la production de données, on trouve évidemment des progrès multiples et considérables en matière de capacité de transmission (fibre optique, 4G, 5G et WiFi), de stockage, de calcul et traitement (à travers des composants plus performants, des infrastructures de *data centers* supportant des architectures en *cloud* et l'avènement de nouveaux algorithmes). La société britannique ARM, leader dans le design des composants pour smartphones, considère qu'entre 2009 et 2015, la performance des ordinateurs a été multipliée par 80, celle des cartes graphiques par 158 et celle de la résolution des écrans par 24...

## **[EN] : Peut-on parler de phénomène *data centric* ?**

**[Y.G.] :** On parle aussi de « l'économie de la *data* ». Celle-ci se substituerait à l'économie de l'énergie associée à la révolution industrielle de la même façon que l'on voyait, dans les années 1990, la « société de l'information » succéder à la « société de consommation ». De fait, l'économie de la *data* domine aujourd'hui le *top ten* des plus grosses capitalisations boursières au niveau mondial (Apple, Amazon, Alphabet/Google, Microsoft, Facebook, mais aussi désormais Tencent, Alibaba...), en ayant bousculé les *majors* de l'énergie.

---

(1) On considère ainsi que 90 % des données disponibles aujourd'hui dans le monde ont été produites durant les deux dernières années.

Si, depuis plusieurs décennies, les systèmes d'information (SI) sont considérés comme le système nerveux des entreprises et des organisations, c'est plus récemment que le traitement de données en masse est apparu comme un facteur essentiel qui touche tous les aspects de transformation d'une entreprise. Ainsi, un opérateur de télécommunications cherchera dans les données un signe d'un possible désabonnement d'un client, mais aussi des automatismes pour optimiser en temps réel son réseau ou encore pour déceler une cyberattaque potentielle. Pour certaines entreprises, la donnée est intimement liée au modèle d'affaires. C'est naturellement le cas des grandes plateformes de l'Internet avec Google qui tire encore l'essentiel de ses recettes de la puissance et de la domination de son moteur de recherche tandis que Facebook arrive à un niveau de rentabilité record en commercialisant ses données sur le nouveau marché de la publicité ciblée (avec les risques associés, comme on l'a vu récemment avec les *fake news* et *data breach* de Cambridge Analytica).

Mais on ne doit pas oublier que la *data* trouve aussi sa place hors de l'économie purement marchande, avec le succès des sites d'échange (pas obligatoirement marchands), la gestion de la complexité urbaine (cf. les pratiques d'« open data » pour exploiter les multiples données urbaines afin d'optimiser les transports ou les plans d'urbanisme), et dans les progrès de la météorologie, de la médecine et plus généralement des activités scientifiques.

Dans le même temps, je crois qu'il est honnête de dire que le Big Data est surtout une pratique en devenir. Beaucoup d'entreprises, même importantes, sont encore dans ce domaine en phase d'apprentissage et ne le maîtrisent pas suffisamment pour l'intégrer dans le cœur de leurs activités. La migration vers le *cloud* est loin d'être achevée, l'IoT démarre... L'abondance de *data* peut s'assimiler à un certain chaos si l'on ne dispose pas des compétences (souvent « aspirées » par les leaders de l'Internet) et d'une stratégie sur les priorités de l'analyse et les données pertinentes à privilégier. Il faut dans ces conditions pouvoir compter sur une direction générale impliquée et prête à soutenir les objectifs mais aussi les aléas de la transformation numérique.

### [EN] : Comment caractériser la rupture introduite par le Big Data ?

[Y.G.] : Trois caractéristiques ont été associées au Big Data en mettant en avant la règle des « 3 V » :

- D'abord, et cela apparaît évident, il faut du Volume ;
- Il faut aussi de la Variété ; schématiquement, ce n'est pas en restant avec des données propres à une entreprise ou à un métier que l'on dispose des systèmes d'analyse les plus puissants et intéressants. D'où les nombreuses sociétés qui sont apparues en mettant en œuvre des compétences de traitement de la donnée mais aussi en intervenant comme *brokers de data* ;
- Enfin, c'est le plus souvent en temps réel que le traitement de ces données sera le plus pertinent (*Velocity*).

Mais si l'on parle de rupture, il faut aborder une autre notion, très directement associée au Big Data et qui même de plus en plus s'y substitue dans les commentaires sur la transformation numérique : l'intelligence artificielle ou ses derniers développements sous la forme de *machine learning*.

Ce sont en effet aujourd'hui les progrès fantastiques de l'intelligence artificielle qui peuvent donner de nouvelles ambitions et perspectives au traitement et à la valorisation de la donnée.

L'intelligence artificielle est une expression assez ancienne qui est apparue dès les années 1950 dans les laboratoires de Stanford. On l'a ensuite un peu oubliée, même si l'on a connu, dans les années 1970, l'émergence de systèmes experts. Il s'agissait alors de décortiquer des sujets, pas trop complexes, en leur appliquant des règles de causalité établies sur la base d'un jeu limité de données structurées. Avec le Big Data, on est passé très schématiquement de cette approche causale à une recherche des corrélations à travers l'utilisation d'un très grand nombre de données (souvent non structurées). Les développements sont en particulier focalisés sur les algorithmes basés sur une

méthode d'apprentissage (*machine learning*) qui va permettre à une machine de maîtriser progressivement les multiples combinatoires d'une partie de Go, ou de distinguer l'image d'un chat de celle d'un chien... Mais malgré les performances spectaculaires obtenues et leur importance pour les domaines de perception (reconnaissance des formes et reconnaissance vocale), les experts du domaine insistent sur le fait qu'il s'agit toujours d'une intelligence supervisée. Il a fallu que des milliers d'images soient codées comme étant celles de chiens pour que la machine apprenne à faire la différence avec celles des chats, tandis que la machine de *deep mind* (Alphabet/Google), qui a battu en 2016 le champion du monde de Go, ne saurait pas s'aventurer sans « apprentissage supervisé » dans une partie d'échecs...

On semble donc encore loin d'une intelligence artificielle au sens où elle disposerait de l'intelligence humaine de perception et de raisonnement, sans même parler de conscience.

### [EN] : Quels sont les nouveaux enjeux réglementaires ?

[Y.G.] : À côté des progrès attendus, il y a en effet un ensemble de problèmes réglementaires très complexes. On se contentera d'en citer quelques-uns :

- un risque sur la préservation des données personnelles : la collecte des données qui contribue au financement de l'innovation et des applications offertes aux internautes s'accompagne de risques de cybercriminalité et de craintes sur la maîtrise des données personnelles (*cf.* le règlement européen dit de RGPD qui entre en application en mai 2018) ;
- un problème de concurrence : la domination des grandes plateformes de l'Internet est alimentée par l'accumulation continue de données de par leur modèle et par la puissance financière acquise (qui leur permet d'acheter les talents et les *start-ups* les plus prometteuses dans le domaine de l'intelligence artificielle).

Ainsi, comme souvent, la concentration de la technologie au sein d'un petit nombre d'acteurs s'apparente à un phénomène complexe à réguler car il recouvre :

- des avantages pour le consommateur ainsi que pour les autres acteurs de l'industrie qui trouvent l'occasion de s'inscrire dans des écosystèmes performants ;
- mais aussi des risques sur le contrôle de nos données personnelles et la distorsion du marché par de nouvelles barrières à l'entrée.

Si l'on prend les interfaces vocales, basées sur les derniers progrès de l'intelligence artificielle, qui sont installées dans les smartphones ou les enceintes domestiques par Apple, Google, Amazon, Samsung et Microsoft, le consommateur peut considérer un progrès sans percevoir tout de suite que ce mode de commande peut aussi s'accompagner d'une limitation du choix offert pour accéder à tel ou tel catalogue de musique en *streaming*. On peut ainsi craindre que ces dernières innovations associées à l'intelligence artificielle viennent avantager les applications des principales plateformes mais aussi freiner le changement de plateformes par le consommateur...

De plus, les dispositions réglementaires esquissées au niveau d'un pays ou d'un groupe de pays (Union européenne) doivent, pour être efficaces, être négociées dans un cadre international, alors même que les enjeux se présentent souvent différemment pour les différents acteurs. On peut à cet égard évoquer les négociations à répétition entre l'Europe et les États-Unis à travers le Safe Harbor puis le Privacy Shield, mais aussi les discussions entre les membres de l'Union européenne qui n'ont pas tous la même sensibilité sur ce sujet.

Il me paraît aussi important de mentionner d'autres débats, par exemple :

- celui qui s'intéresse à la symétrie des règles à l'intérieur de la chaîne de valeur ; ainsi les opérateurs de télécommunications soulignent que les fournisseurs de service Internet (OTT) seraient beau-

coup moins encadrés dans la gestion des données personnelles que les opérateurs de télécommunications par leur régulation sectorielle ;

- celui qui oppose de plus en plus souvent les grands acteurs de l'Internet et les autorités publiques ; on a vu Microsoft s'opposer au droit de regard des autorités américaines sur des données inscrites dans un de ses *data centers* en Irlande ; Apple a refusé, dans une affaire criminelle aux États-Unis, d'« ouvrir » le système d'exploitation d'un iPhone ; l'application Whatsapp a été (momentanément) fermée au Brésil sur la requête d'un juge qui souhaitait avoir accès au contenu de cette messagerie cryptée.

D'un autre côté, il faut aussi souligner le poids que peut avoir la régulation « par la *data* », en faisant notamment appel aux consommateurs dans l'esprit des réflexions de Nicolas Colin et Henri Verdier (*L'Âge de la multitude*). Le régulateur français des télécoms (ARCEP) s'est engagé dans cette voie en mettant des cartes géographiques à disposition des usagers pour leur permettre de noter la qualité des services mobiles sur le territoire.

### **[EN] : Faut-il réguler les plateformes et s'assurer de la loyauté des algorithmes ?**

[Y.G.] : Je pense qu'il est assez difficile et peut-être dangereux de prescrire une réglementation *ex ante* à vocation générale mais propre aux plateformes. Il faut peut-être d'abord s'assurer de ce que les différents droits (droits des consommateurs, droits des contrats, droits intellectuels, droit fiscal, droit de la concurrence...) soient respectés, et quand c'est nécessaire construire les rapports de force qui permettent d'imposer les remèdes pertinents. Des progrès sont possibles et cela a commencé. On a vu que la puissance publique pouvait imposer à une grande plateforme de réservation l'abandon du principe d'exclusivité qu'elle imposait aux hôtels. On a également vu que Google a dû amender le fonctionnement de son moteur de recherche. On peut imaginer contrôler la loyauté des algorithmes utilisés par les plateformes dominantes... par exemple, en s'assurant de temps en temps par sondage que les recommandations d'un site d'eCommerce vous informant de ce que vos amis ont aimé .... sont réellement fondées !

Cela étant, je suis très inquiet quand je vois que le premier réflexe devant une innovation de rupture se limite à la réglementation. Il faut bien sûr aussi compter sur la dynamique d'innovation et de concurrence, et pour cela favoriser les investissements dans la connaissance et la création.

Après tout, on a vu apparaître de nouvelles plateformes sectorielles (Uber, AirBnB, Spotify, Netflix...) et les GAFAM<sup>(2)</sup> peinent souvent à réussir leur diversification « horizontale »...

### **[EN] : Quel constat peut-on faire sur le positionnement de l'Europe face au dynamisme des États-Unis et au rattrapage de la Chine ?**

[Y.G.] : L'Europe est absente des dix plus grosses capitalisations boursières qui sont, comme je l'évoquais, dominées par les GAFAM et désormais leurs homologues chinois (Tencent, Alibaba, Baidu). La situation n'est donc pas brillante mais il faut bien voir que nous ne sommes qu'au début de la transformation numérique. Nous conservons des atouts, dans l'IoT, dans l'intelligence artificielle avec des chercheurs reconnus dans le monde entier, dans les télécommunications, dans les industries culturelles (notamment les jeux vidéo). Bien que les investissements européens en capital-risque soient encore en retrait, l'écart avec les États-Unis se réduit et les écosystèmes de *start-ups* à Londres, Paris ou Berlin vont certainement générer de plus en plus de licornes... Enfin, la transformation numérique, en touchant des secteurs de l'industrie dans lesquels l'Europe dispose d'acteurs puissants (aéronautique, automobile, énergie...), offre de nouveaux challenges à notre industrie (Industry 4.0) et des terrains d'émergence de nouveaux leaderships européens dans le numérique.

---

(2) Google, Apple, Facebook, Amazon et Microsoft.

**[EN] : Quelles seraient vos recommandations de politiques à mettre en œuvre au niveau européen ?**

[Y.G.] : 1. L'éducation – par la qualité et la profondeur de la formation initiale, supérieure et « tout au long de la vie » – constitue une priorité connue mais qui n'est pas encore assez prise en compte. Elle doit s'accompagner d'un renforcement de la culture scientifique sans qu'on fasse une mauvaise querelle aux sciences « molles ». C'est la condition pour ne pas subir la transformation numérique, mais aussi pour éviter que l'Europe subisse la polarisation des emplois.

2. Il faut poursuivre les progrès observés dans le domaine des *start-ups* et de la culture d'entrepreneur, adopter les dispositions fiscales permettant de mieux orienter l'épargne pour renouveler notre tissu économique et monter en gamme. C'est en réduisant le *gap* de productivité entre les 10 % des entreprises les plus avancées dans l'intégration numérique et celles qui restent encore très en retrait que l'on évitera la fameuse « stagnation séculaire ».

3. Malgré les vicissitudes du projet européen, les pays membres doivent poursuivre leurs efforts pour disposer d'un cadre réglementaire harmonisé afin de bénéficier d'un grand marché domestique et de favoriser l'émergence de leaders européens. Mais l'Europe du numérique doit aussi s'exprimer comme un projet mobilisateur, alternatif aux dérives qui peuvent ruiner la confiance du plus grand nombre. C'est l'enjeu des discussions sur le régime fiscal des plateformes, sur les abus de position dominante, ainsi que sur la maîtrise des données personnelles et la lutte contre la cybercriminalité.

4. La valorisation de la place éminente du numérique dans les économies d'énergie (pas de systématisation des énergies alternatives sans « réseaux intelligents ») et la diminution des pollutions (*smart cities*) doit enfin être renforcée et devenir un élément de confiance dans l'innovation et l'adhésion à un axe stratégique associant transformation numérique et transition énergétique.

**[EN] : Finalement, quelle est votre vision sur l'évolution du Big Data dans les prochaines années ?**

[Y.G.] : Quand on observe les différentes catégories dans lesquelles se rangent les *start-ups* ainsi que les technologies mises en œuvre, on s'aperçoit de ce que le Big Data, combiné avec les progrès de l'intelligence artificielle, est une source d'innovation qui s'applique à tous les secteurs de l'innovation numérique. L'IoT va permettre de drainer de multiples données et donc de transformer radicalement la maintenance des grands systèmes complexes (« maintenance prédictive »). La cybersécurité, en intégrant de plus en plus d'intelligence artificielle, va pouvoir détecter des « signaux faibles » annonciateurs d'attaques. Le traitement du cancer ne pourra plus se concevoir sans le concours de systèmes capables d'exploiter de gigantesques bases de données pour préciser le diagnostic et évaluer le meilleur traitement. Et l'on pourrait continuer en faisant référence à la voiture autonome, à la virtualisation des réseaux de télécommunications, à la FinTech et l'InsurTech, à la publicité programmatique, etc.

Les perspectives sont donc très impressionnantes mais encore une fois, nous n'en sommes qu'aux balbutiements... et les usages concrets dans les entreprises sont encore parfois très sommaires. Le *cloud*, l'IoT sont encore en phase de déploiement. Les options réglementaires sont encore incertaines. La place que va prendre la cybercriminalité n'est pas bien cernée... Dans le contexte actuel, le rythme auquel se généralisera le Big Data, et surtout les scénarios « industriels » et géopolitiques qui l'accompagneront, ne sont pas encore très clairs.