
TABLE RONDE DE CLÔTURE

Philippe TASSI

DGA, Médiamétrie

Le contexte des instituts nationaux de statistique est dominé, en Europe, par l'apparition de nouveaux phénomènes à mesurer, des besoins changeant rapidement, un public donnant de plus en plus difficilement sa confiance (confidentialité, réponses aux enquêtes), des données massives appartenant à des acteurs privés, des réductions budgétaires, et des difficultés pour attirer les bonnes compétences.

Les organismes privés connaissent un contexte proche, *mutatis mutandis*.

Pour l'illustrer rapidement : en audience de la télévision, toujours plus de comportements à mesurer, avec l'ajout, à l'initial téléviseur, de nouveaux écrans (smartphone, tablette, PC), trois modes de vision (« live », différé, *replay*), le domicile et le hors domicile, de nouvelles offres concurrentielles (VOD, SVOD), le tout dans un environnement publicitaire peu favorable, sans même parler de la réflexion en cours sur l'audiovisuel public et de ses probables conséquences budgétaires ; des Big Data propriété des opérateurs ; une flambée des salaires offerts aux Data Scientists, et donc une tension sur le marché des diplômés de bon niveau, avec pour effet une baisse de la fidélisation.

Thème 1 : le défi des données massives et de l'open data

Les données massives, dans le domaine des médias, sont nées avec l'émergence du digital et la convergence numérique, au milieu des années 90 ; plus exactement, elles ne sont devenues massives qu'avec l'équipement des foyers et des individus en objets reliés au réseau internet. Smartphone, tablettes, ordinateurs, téléviseurs connectés, box ADSL enregistrent des logs élémentaires. Ces Big Data sont caractérisées par deux paramètres : la quantité, pouvant aller jusqu'à l'exhaustivité, et la fréquence pouvant aller jusqu'au temps réel. Le croisement de ces deux dimensions permet d'atteindre des volumétries difficilement imaginables il y a encore quelques années.

En médias, ces données massives sont souvent de nature différente des données d'échantillons ou de panels : ces dernières proviennent d'individus, et de leur plein gré, alors les Big Data sont souvent au niveau des objets, et donc de granularité plus grossière. Ainsi, l'état d'une box ADSL donne une information sur la chaîne diffusée à l'écran mais aucune sur le nombre des téléspectateurs et qui ils sont. Plus encore, big ne veut pas dire parfait : ainsi, si quand un téléviseur est éteint et si la box reste active, elle enregistre comme chaîne diffusée la dernière chaîne regardée avant l'extinction. Exhaustif, oui, mais exhaustivement faux !

Il est intéressant de mélanger alors les deux sources, panels et données massives, après avoir apuré les données massives de ce genre d'aberrations. Hybrider les deux sources permet d'en créer une troisième, plus riche, plus précise, et d'enrichir les services rendus aux groupes médias.

Par exemple, la mesure de référence de l'audience d'internet est issue, depuis 2012, de l'hybridation des données d'un panel d'individus et de Big Data fournissant l'exhaustivité des visites, une fois celles-cicorrigées des fausses visites venant de robots ou de « fermes à clics ».

Les Big Data sont une innovation, autant s'en servir. Comme toute innovation, elle arrive avec de nouvelles questions théoriques ou opérationnelles à résoudre, souvent passionnantes.

Thème 2 : l'intégration de la géographie dans la statistique

C'est probablement une évidence pour tous.

Dans les médias, cela signifie dépasser le niveau national avec la volonté d'agréger les résultats d'un média, d'un programme, sur plusieurs pays ou zones géographiques : la diffusion par satellite ou par internet abolit les frontières.

Cela veut dire aussi passer du national en local. Un exemple est donné par le média radio. Connaître la hiérarchie des stations au niveau national est intéressant, certes, mais il est tout aussi fondamental pour les annonceurs publicitaires locaux d'avoir des éléments sur les préférences des personnes vivant dans leur zone de chalandise, compte tenu de l'importance parfois non négligeable des stations locales. Nous sortons donc des résultats pour chaque agglomération de plus de 50000 habitants.

Autre effet de la géographie : elle permet d'aller au-delà de la notion de moyenne et d'introduire de la variance. Pour illustrer cette variabilité, si 48 % des foyers ont la télévision par réception hertzienne, cette proportion varie de 38 % en région parisienne à 56 % dans le nord.

Enfin, il y a un lien entre la datavisualisation et la dimension spatiale.

Thème 3 : la méthodologie classique à l'épreuve des nouveaux comportements de réponse

L'un des mots-clés est la confiance. Confiance des clients d'une part, du grand public dont nous recueillons les comportements d'autre part.

Côté clients, la certification ISO 9000 est utile, et il existe en outre des organismes de contrôle et d'audit, qui valident les protocoles employés et le traitement des données, et donc la conformité au cahier des charges initial.

Côté grand public, il est de fait que dans la plupart des pays ayant un système statistique développé, on observe des difficultés de participation de plus en plus grandes, en particulier auprès des individus les plus jeunes, et pas seulement pour des panels de longue durée.

Je ne suis pas persuadé que les sujets de confidentialité soient la seule cause de cette désaffection, même si l'actualité récente liée à Facebook et Cambridge Analytica ou à toute la communication faite autour de la mise en oeuvre du RGPD n'a pas joué un rôle positif. L'existence de sources de données privées, propriétaires, souvent quasi-instantanées, mais qui ne sont représentatives que d'elles-mêmes, amène certains à se poser des questions sur l'intérêt de mesures globales.

La méthodologie doit donc continuer à s'enrichir, comme elle l'a toujours fait, d'ailleurs : j'ai évoqué l'hybridation de sources, et il faut aussi citer les multisources de recueil et leur agrégation – mettre la personne sollicitée dans sa zone de confort est une condition nécessaire à une bonne participation –, l'utilisation des parodonnées, le traitement des non-réponses.

Il est important également d'activer une veille de moyen terme sur un existant connexe, pour voir s'il est intéressant et transposable dans notre milieu; trois sujets me viennent à l'esprit : la blockchain, l'intelligence artificielle (trois rapports à lire absolument : « Anticiper les impacts économiques et sociaux de l'Intelligence Artificielle », publié le 22 mars 2017 par France Stratégie et le CNN ; « Pour une intelligence artificielle maîtrisée, utile et démystifiée » de l'OPECST, Office parlementaire d'évaluation des choix scientifiques et technologiques, publié le 29 mars 2017 ; « Donner un sens à l'intelligence artificielle », Cédric Villani, sorti en mars 2018), et enfin les techniques quantiques.

Thème 4 : le management, les compétences et les méthodes de travail

Vaste sujet, de pleine actualité, dans le public comme dans le privé, même si les environnements sont différents. Une agence de RH bien connue a publié il y a quelques jours un communiqué dont le

titre est « Au diable la stabilité ! Changer de boulot fréquemment, c'est s'assurer plus de curiosité, de savoir-faire ... et d'argent ».

Les compétences nouvelles, par exemple autour des Big data, sont très recherchées. Compétences en modélisation, en mathématiques de très grande dimension, en algorithmes et en learning, en codage, en systèmes d'information, en cloud, en datalake, etc, etc ...

Il faut donc attirer les personnes ayant ces compétences et les fidéliser.

Les thèmes actuels du management tournent en conséquence autour des nouvelles formes de travail (télétravail, co-working, réseaux personnels, mobilité), sur la mobilisation, l'engagement, sur l'acquisition permanente des connaissances nouvelles

– la formation continue est plus que jamais indispensable –, sur les aspirations des jeunes générations – qu'on les appelle Y, Z, millenials –, et sur un point vraiment à travailler : la reconnaissance non financière.

Beaucoup de travail, donc, et c'est heureux d'être dans un domaine statistique en constante évolution. La monotonie est bien loin.