

Big Data et mesure d'audience : un mariage de raison ? *

Lorie Dudoignon, Fabienne Le Sager et Aurélie Vanheuverzwyn

Question clé

Apparue en 1995, la convergence numérique a peu à peu modifié l'univers des données, en mettant en évidence le pouvoir des Big Data ou données massives. Les données massives doivent-elles être vues comme un risque de retour potentiel vers le paradigme de l'exhaustif dominant jusqu'à la fin du 19^e siècle, alors que le 20^e siècle a été celui de l'échantillonnage et des sondages ? Telle est la question à laquelle Médiamétrie est confrontée dans le cadre de ses dispositifs de mesures d'audience.

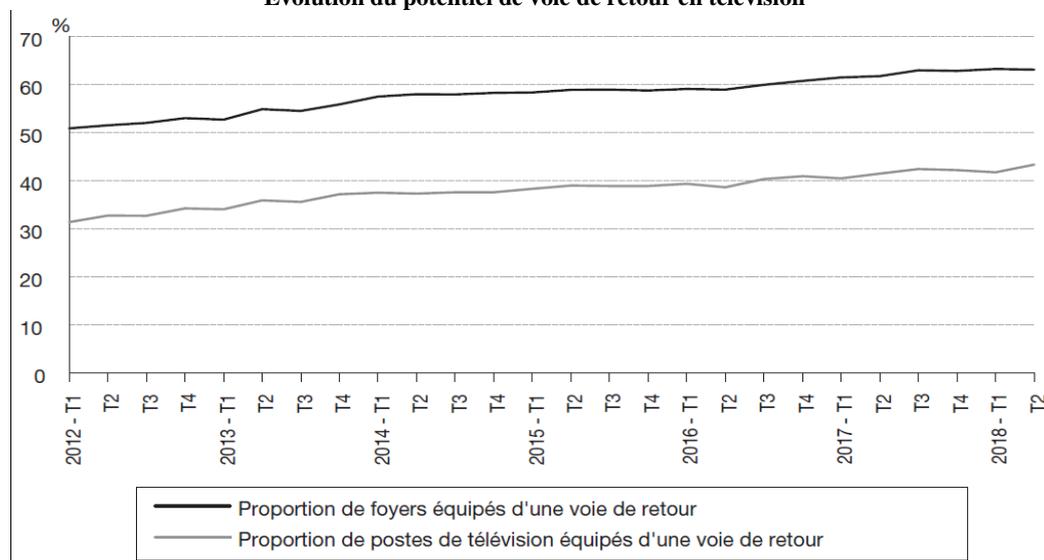
Méthodologie

L'analyse porte tout d'abord sur les avantages et les limites comparées des données d'enquête et des données massives en termes de qualité, ce qui permet d'expliquer le choix d'approches dites hybrides, qui consistent à mélanger deux sources d'informations pour en créer une troisième plus riche ou plus fine.

Principaux résultats

- Les enjeux de qualité concernent tout autant les données massives que les données d'enquêtes.
- La mesure d'audience Internet repose sur deux types de dispositifs : les premiers, dits *user-centric*, centrés sur l'utilisateur, et les seconds, dits *site-centric*, centrés sur le site. L'approche hybride adoptée pour la mesure d'audience Internet consiste à redresser la donnée d'enquête sur les résultats exhaustifs *site-centric* par calage sur marges.
- En télévision, on dispose de données issues de la voie de retour pour certains modes de diffusion comme l'ADSL, le câble ou le satellite. La mesure d'audience par panel ne permet pas toujours de mesurer finement des usages très morcelés. L'utilisation des données voie de retour est donc particulièrement intéressante pour les chaînes thématiques. L'approche hybride adoptée dans ce cas repose sur les chaînes de Markov cachées.

Évolution du potentiel de voie de retour en télévision



Champ : France métropolitaine.
Source : Médiamétrie, Home Devices.

Message

Les données massives ne doivent pas être vues comme une alternative aux mesures d'audience traditionnelles basées sur des enquêtes par sondage mais comme une opportunité de les améliorer.