

Données immobilières : quels enjeux pour les professionnels de l'immobilier ?

Madame JUILLARD, dans son interview ci-dessous, se limite aux prix et loyers des logements du marché privé.

Les données immobilières se multiplient, s'organisent, se structurent. Pouvez-vous nous dresser un état des lieux du marché de la data immobilière ?

Aucune donnée n'est donnée. Une donnée est une construction, le résultat d'opérations destinées à transformer l'information brute en information qualifiée.

Ce préalable posé, il faut à mon sens distinguer deux marchés : celui des données de prix et de loyer fondées sur le recueil d'informations à la source et disponibles à la transaction ; celui des estimations automatiques produites par des agrégateurs de données.

“

La multiplication des informations ajoute à la confusion et augmente le bruit qui entoure les prix et les loyers des logements.

”

Sur le marché des « données », on compte une petite dizaine de producteurs proposant une large couverture du territoire, complétés d'observatoires locaux, membres de réseaux à vocation nationale. Ceux-là forment un écosystème plutôt stable dans ses composantes et ses méthodes : la plupart officie depuis vingt à cinquante ans, ne laissant qu'une place très marginale aux nouveaux entrants ; ils empruntent la voie de la dématérialisation tracée par le numérique, mais restent attachés au travail alla mano, gage de qualité.



Crédit : Xavier Lahache

Les producteurs de « données » couvrent un à deux segments du marché, jamais l'ensemble. Leurs données s'en trouvent très éclatées. Elles reproduisent l'organisation institutionnelle et professionnelle du monde de la production du logement, entre ancien et neuf, collectif et individuel, commercialisation intermédiaire et non intermédiaire. La raison en est précisément que leurs producteurs remontent à la source pour les élaborer. Leurs données sont à la fois concurrentes et interdépendantes. Tandis que la concurrence se joue par segment, la relation d'interdépendance se noue autour du rôle central des données publiques, légitimes car officielles. L'ouverture récente des données de transactions de la DGFIP, une manne pour tous, un socle à présent partagé, promet de l'accentuer.

Les estimations automatiques forment un autre ensemble. Alors que celles qui sont destinées aux particuliers ne se comptent plus, celles qui s'offrent aux professionnels s'avèrent moins nombreuses. Elles ne concurrencent pas vraiment les données existantes, mais les complètent. Elles ne puisent pas non plus à de nouvelles sources de prix et de loyers, mais constituent les résultats de modèles mathématiques qui agrègent les données existantes avec des données de contexte très diverses. Les opérations sur lesquelles elles reposent sont sous-tendues par deux enjeux : produire des données rafraîchies et actualisées à un instant T, celui de la consultation ; offrir des données à l'échelle du logement, celle de prospection et de l'investissement.

“

Il vaut mieux être un bon connaisseur du marché qu'un bon data scientist.

”

Quelles limites voyez-vous à la multiplication des données ?

J'y vois un nouveau rempart contre la transparence des marchés. De fait, la multiplication des informations ajoute à la confusion et augmente le bruit qui entoure les prix et les loyers des logements. Si les « données » et les estimations disponibles reproduisent toutes les mêmes catégories (cela en est même étonnant), elles fournissent des appréciations différentes de la valeur des logements. Surtout, elles sont incomparables, car construites différemment et disponibles à des formats différents. Enfin, la multiplication des canaux d'accès aux données n'aide pas. Elle est quasi exponentielle, puisque tous les producteurs empruntent plusieurs canaux à la fois pour adresser leurs données à des cibles différentes. Cela rend difficile la recherche d'informations.

Comment le marché va-t-il ou doit-il se développer, pour rendre la donnée plus exploitable qu'elle ne l'est aujourd'hui ?

Un premier enjeu se dessine autour de la disponibilité des données brutes. L'idéal serait de disposer d'une plateforme, sorte de dataplace partagée entre les producteurs et avec les utilisateurs. On en est loin. Non seulement les données restent très éclatées, mais elles sont rarement interopérables, donc prêtes à l'emploi. Pour les utiliser, un travail de façonnage s'impose, qui comprend des opérations d'identification, de jointure, de normalisation, d'alignement et de géocodage d'une grande complexité. Un travail de traduction et de transposition s'y ajoute, pas moins complexe. De quoi rappeler que même

les données dites brutes ne sont pas neutres mais rattachées à un périmètre de compétences, un cadre de production et une intention d'usage qui, bien qu'atténués par les ajustements préalables à leur ouverture, rendent leur utilisation particulièrement délicate. Avec les données, il n'y a rien de magique et l'avènement du numérique n'y a rien changé.

La production de statistique renvoie à un second enjeu. Il s'agirait d'harmoniser les statistiques disponibles. On en est loin là encore. Il s'agirait également de libérer les statistiques des cadres trop figés qui les enferment habituellement. L'idéal serait de disposer d'outils tellement agiles qu'ils permettraient à l'utilisateur de délimiter le propre contour de son champ d'observation. On est proche de l'utopie, quoiqu'on assiste aujourd'hui déjà au développement d'outils de consultation interactive de plus en plus agiles et, plus généralement, à un glissement de la vente de données vers la vente de services permettant de les exploiter plus facilement.

Comment les professionnels du patrimoine peuvent-ils s'approprier et utiliser ces données dans leurs préconisations immobilières ?

En s'armant d'une stratégie intégrée et d'outils partagés pour faciliter la collecte, la prise en main et l'utilisation des données utiles disponibles. En utilisant les données en connaissance de cause, c'est-à-dire en étant exigeants sur leur transparence. En les croisant avec leurs propres données, elles-mêmes consolidées, tracées, actualisées et valorisées au moyen d'une véritable politique data. En partageant leurs données dans une démarche de mutualisation et de production ad hoc (les brokers l'ont bien fait pour disposer de données clés sur le marché). Enfin, en s'appuyant et en améliorant toujours leur expertise, la seule qui vaille ! Aujourd'hui encore en effet, les données immobilières sont d'abord affaire d'immobilier, pas de donnée et il vaut mieux être un bon connaisseur du marché qu'un bon data scientist. De quoi se prémunir contre tout leur techniciste. ■

Claire JUILLARD

Docteur en sociologie,
spécialiste de la ville, du logement
et des marchés immobiliers.