

## Ces algorithmes dont on ne sait rien alors qu'ils régissent nos vies



Selon des révélations du Wall Street Journal, le moteur de recherche d'Amazon a tendance à privilégier la marque de la plateforme. Les algorithmes sont de plus en plus opaques.

Avec Daniel Le  
Métayer

**Atlantico.fr : D'après une enquête du Wall Street Journal, sous couvert d'impartialité, le moteur de recherche d'Amazon privilégie les produits de la marque de la plateforme, notamment de manière indirecte. Malgré l'opacité de cet algorithme, est-il possible dans ce cas de prouver cette pratique afin d'envisager une pénalité pour abus de position dominante, comme celle infligée par Margrethe Vestager à Alphabet/Google en 2017?**

**Daniel Le Métayer :** De manière générale, à défaut de possibilité d'accéder au code informatique du système, le mieux que l'on puisse faire est de le tester, c'est à dire de lui soumettre un grand nombre de requêtes et de tenter, à partir des résultats, d'inférer les facteurs déterminants (les caractéristiques qui ont influencé ce résultat). En l'occurrence, il faudrait soumettre un grand nombre de requêtes à Amazon (recherches de produits) et tenter d'établir, à partir des classements proposés, si la caractéristique « marque Amazon » est déterminante. C'est ce que la Commission Européenne a fait, entre autres, en analysant 1,7 milliards de requêtes sur le moteur de recherche de Google avant de prononcer sa sanction pour abus de position dominante. Cependant, la tâche est difficile si on ne connaît pas les autres caractéristiques, qui seraient de l'ordre d'une centaine pour Amazon, ni leur valeur. Par exemple, on peut connaître la caractéristique critère « vitesse d'expédition » associé à un vendeur, mais probablement pas la caractéristique « volume des ventes » d'un produit qui devrait être un facteur important. De plus, on n'obtiendra pas de preuve absolue mais un faisceau d'indices, des éléments de preuve.

**Quelles lacunes de gouvernance identifiez-vous en France et dans l'Union Européenne en ce qui concerne la transparence de tels algorithmes, en tenant compte des prérequis du secret industriel?**

Les lacunes sont multiples, aussi bien pour ce qui est de la loi que de son effectivité. Par exemple, la loi pour une République numérique prévoit en France une obligation, pour les opérateurs de plateformes en ligne, de fournir une information « loyale, claire et transparente » sur les modalités de référencement, de classement et de déréférencement des contenus. Le décret d'application précise que ces informations doivent comprendre « les critères de classement par défaut des contenus et des offres de biens et services, ainsi que leur principaux paramètres ». La loi impose aussi d'informer le public sur « l'existence d'une relation contractuelle,

---

d'un lien capitalistique ou d'une rémunération à son profit, dès lors qu'ils influencent le classement ou le référencement des contenus, des biens ou des services proposés ou mis en ligne ». Si la société Amazon a adapté son algorithme pour favoriser ses propres produits, comme l'affirme le Wall Street Journal, elle est en infraction avec cette loi dès lors que les internautes n'en sont pas informés.

Idéalement, il faudrait renforcer les obligations des opérateurs ou des fournisseurs de systèmes algorithmiques qui ont une incidence importante sur la vie des citoyens, soit sur leur vie quotidienne, soit dans des secteurs critiques comme la justice, la police, ou encore la santé. Avant de déployer ce type de système, il devrait être nécessaire de conduire une « étude d'impact algorithmique », comme cela a été introduit récemment pour le secteur public par une directive canadienne. Par ailleurs ces systèmes devraient pouvoir être audités par une autorité indépendante. Une telle autorité étant tenue à des obligations de confidentialité, l'argument de protection des secrets industriels ou de la propriété intellectuelle ne me paraît pas recevable.

**On entend souvent que les algorithmes d'apprentissage (communément appelés intelligences artificielles) complexifient l'exercice de la transparence car leur créateur, par définition, ne les maîtrise pas de bout en bout. Est-ce le cas pour tous les algorithmes d'apprentissage? Comment peut-on identifier et prévenir alors les abus?**

Tout d'abord, il faut être clair sur le fait que ce n'est pas la complexité liée à l'apprentissage qui est en cause dans les cas évoqués à propos d'Amazon et de Google : il s'agit d'une opacité ou plutôt d'un secret délibérément entretenu et non pas d'une conséquence de choix techniques. Pour ce qui est de la complexité des modèles produits par apprentissage, elle est réelle pour certaines techniques, notamment ce que l'on appelle l'« apprentissage profond » qui est particulièrement performant pour des applications comme l'analyse d'images.

De manière générale, je préfère parler d'explicabilité plutôt que de transparence, car montrer ne signifie pas forcément rendre compréhensible. Il s'agit d'un domaine de recherche actuellement très actif. Selon les domaines d'application, on peut adopter différentes stratégies pour assurer l'explicabilité. Une première option est d'entraîner un algorithme à produire lui-même des explications de ses résultats ; quand le concepteur n'est pas collaboratif, comme pour Amazon et Google, on peut essayer de produire des explications d'algorithmes a posteriori par des formes de rétro-ingénierie. Dans certains cas, il est possible de mettre au point des techniques d'apprentissage offrant à la fois une précision acceptable et une bonne intelligibilité. Tout dépend des secteurs d'application : par exemple, pour un algorithme utilisé par un juge ou un médecin, une absence totale d'intelligibilité paraît difficile à accepter ; pour un algorithme de jeu de Go, il ne s'agit pas d'une exigence essentielle.

**Propos recueillis par Aurélie Buytaert**