



# BIG DATA, ALGORITHMES ET RISQUES DE DISCRIMINATIONS

## L'EXEMPLE DE L'ASSURANCE

### **Algorithmes et risques de discriminations, quelle éthique dans le secteur de l'assurance**

Environ 50 personnes ont assisté le 15 septembre au séminaire « Algorithmes et risques de discriminations, quelle éthique dans le secteur de l'assurance ? », organisé par la Ligue des Droits de l'Homme. Au travers de deux tables-rondes, les discussions ont permis de mieux cerner les enjeux du Big data et des algorithmes dans le domaine de l'assurance, et des verrous à poser pour lutter contre les discriminations que leur utilisation pourrait engendrer. Tel que souligné dès d'introduction par Gérard Aschieri, rédacteur en chef de la revue *Hommes & Libertés* et membre du Cese, il ne faut pas oublier que les algorithmes ne sont pas neutres : ils reflètent des choix techniques mais également des choix politiques, sociaux, éthiques.

### **La première table-ronde a été l'occasion de rappeler le contexte de production et d'utilisation des algorithmes**

Serge Abiteboul, directeur de recherche à l'Inria et à l'ENS de Paris, le rappelle : « *Le Big data n'est pas magique, et les algorithmes ne sont ni justes ni injustes par nature : ils sont une création humaine* ». Il s'agit de techniques permettant l'acquisition des données numériques sur lesquelles sont effectuées des calculs dans le but d'en apprendre de nouvelles informations. En conséquence, les algorithmes utilisés peuvent devenir source de discriminations, comme l'ont montré des cas de discrimination sociale dans les prix de vente en ligne d'agrafeuses du fait du lieu de résidence, ou de discrimination ethnique dans la justice pour la prédiction du risque de récidive. Leur origine provient des types de données utilisés qui peuvent introduire un biais dans les résultats. Il s'agit donc d'un véritable enjeu éthique.

Parmi les données utilisées pour le fonctionnement des algorithmes, on trouve des données à caractère personnel. Un cadre juridique existe avec la Loi Informatique et Libertés, et l'entrée en vigueur de Règlement européen relatif à la protection des données (RGPD) en mai 2018. Or, comme le rappelle Jeanne Bossi-Malafosse, avocate et experte auprès du Conseil de l'Europe, les grands principes de protection des données personnelles établis dans ces deux textes ne s'accordent pas ou mal avec le Big data. A titre d'exemple, le principe de finalité déterminée d'un traitement de données personnelles s'accommode difficilement avec la vocation du Big data de réutilisation des données pour des finalités auxquelles elles n'étaient pas originellement destinées. De même, les techniques de pseudonymisation et d'anonymisation, dont l'usage est recommandé par la réglementation deviennent plus difficiles à mettre en œuvre sur de gros volumes de données. Marie-Hélène Gostiaux, avocate, souligne d'ailleurs que le Big data est constitué de données diverses, personnelles ou non, publiques ou non : leur interconnexion permet de déduire des connaissances sur les individus, voire ré-identifier une personne. Dans ce contexte, comment assurer le consentement et l'information ? Les pratiques doivent être observées par le législateur pour les encadrer au mieux.

Au-delà du cadre juridique de l'utilisation des données personnelles, l'utilisation des algorithmes a un impact social, parce qu'il modifie le rapport des usagers à l'information et au savoir. Il existe une interaction entre les individus et une connaissance forcée par les algorithmes créés par Google, Facebook, Twitter, etc. explique Fabrice Epelboin, enseignant à Sciences-Po Paris. La circulation d'informations déterminée par des algorithmes dont ils ignorent tout, enferme les individus dans une bulle informationnelle réconfortante et dans un univers social qui leur correspond.

**La deuxième table ronde a été l'occasion d'axer les discussions sur les transformations induites par les technologies dans le domaine de l'assurance et de leurs conséquences en termes de risques pour les individus.**

Pour Florence Picard, présidente de la Commission scientifique de l'Institut des actuaires, les données sont la matière première des assureurs, et ont toujours été utilisées pour l'évaluation du risque. Mais le numérique bouleverse les pratiques, et les assureurs se voient fragilisés par de nouveaux acteurs (Google, etc.) qui sont à la source des données. Pour éviter de se voir marginaliser, les assureurs doivent évoluer vers de nouveaux services de prévention individualisés. Ces services vont agir directement sur la source du risque pour tenter d'éviter qu'il survienne explique Fabrice Faivre, directeur R&D au Pôle IARD de la Macif, mais ils nécessitent que les assureurs disposent de données supplémentaires, toujours plus importantes. Cette évolution a pour conséquence de nouvelles tensions, alerte Florence Picard, entre l'offre si ce n'est l'obligation d'un service de prévention et le respect des libertés individuelles. C'est pourquoi il est nécessaire de proposer ces services par rapport à une norme, qui peut être intégrée dans un algorithme par exemple. Paul Olivier Gibert, fondateur de Digital&Ethics précise également que l'utilisation d'algorithmes peut générer une tension entre mutualisation et individualisation des services assurantiels.

Pour Bernard Benhamou, secrétaire général de l'Institut de la souveraineté numérique, la question des algorithmes et de l'assurance est politique au sens du devenir de la société à travers les technologies, et de la question de savoir si notre modèle social doit être remis en cause par ou à cause des technologies. En tant qu'utilisateur chacun doit prendre position et participer au débat sur les formes technologiques de notre société, débat qui est aujourd'hui capté par des experts.

Slimane Laoufi, Défenseur des droits, rappelle quant à lui que l'enjeu central de la thématique est la lutte contre les discriminations. Les risques de discriminations dans le domaine de l'assurance existent, d'autant plus que la construction d'algorithmes à partir de données sert à profiler les individus. Dans le domaine des assurances, les collectes des données peuvent être certes légales, mais l'utilisation qui en est faite peut aller au-delà de ce qui est prévu par la loi, conduisant à des situations de discriminations directes tels que le refus de demande assurance, ou des situations de subordinations, où l'individu se voit octroyer des produits sous certaines conditions. C'est le notamment le travail du Défenseur des droits, en collaboration avec la Cnil, de contrôler l'absence de discrimination dans l'utilisation des algorithmes.

Parmi les solutions proposées, certains réclament un meilleur encadrement juridique pour la création, la maîtrise et la transparence des algorithmes et une éthique dans les assurances. Tous s'accordent sur le fait que la technologie que l'on dit être intelligente ne l'est que si elle est respectueuse des droits. Des actions de sensibilisation et formation sont nécessaires avec les opérationnels sur l'utilisation des algorithmes. Serge Abiteboul précise cependant que ce n'est pas aux informaticiens de définir ce qui est éthique, mais à la société. En effet, des moyens sont à penser, des cadres à définir, des évaluations et des contrôles à réaliser conclue Malik Salemkour, président de la LDH. « *En soi, le progrès technique n'est pas par essence une menace* » rappelle-t-il, « *cela dépend de son usage : il doit être clair et transparent dans ses modalités et dans ses finalités, et dans chacune de ses réalités* ». C'est un véritable débat de société auquel tous les citoyens doivent pouvoir participer, et que des associations comme la Ligue des droits de l'Homme souhaitent faire vivre et animer.

L'événement ayant été diffusé en streaming est disponible sur youtube  
<https://www.youtube.com/watch?v=XA14yd70dQA&t=2563s>  
[https://www.youtube.com/watch?v=B\\_-B6B72CBU&t=926s](https://www.youtube.com/watch?v=B_-B6B72CBU&t=926s)