



La transparence des algorithmes

Céline Teyssier, Philippe Roose, Diane Adrianirina

► **To cite this version:**

Céline Teyssier, Philippe Roose, Diane Adrianirina. La transparence des algorithmes. Les convergences du droit et du numérique, 2018, Bordeaux, France. hal-02437217

HAL Id: hal-02437217

<https://hal-univ-pau.archives-ouvertes.fr/hal-02437217>

Submitted on 13 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La transparence des algorithmes

Céline Teyssier¹, Philippe Roose², Diane Adrianirina³

¹ CDRE, /Université de Pau et des Pays de l'Adour

² LIUPPA/Université de Pau et des Pays de l'Adour

³ Université de Bordeaux

Le terme l'algorithme nous est devenu familier. Jusqu'à lors réservé aux mathématiciens, statisticiens et informaticiens, la société de l'information et de la communication a contribué à la diffusion de ce mot sur la place publique.

Déjà utilisé sous l'antiquité, l'algorithme est un programme de calcul qui permet à partir des données d'entrée de trouver un résultat. Aujourd'hui les algorithmes servent des usages très variés. Ils permettent d'identifier le meilleur candidat lors d'un recrutement ; ils aident les entraîneurs sportifs dans le choix des tactiques de jeux ; sur les marchés financiers, ils passent des ordres d'achat et de vente ; ils vous proposent la musique que vous aimez ; ils calculent l'itinéraire le plus rapide et vous trouvent le ou la partenaire sexuel(le) idéal(e).

L'algorithme a pris une part considérable dans la société grâce à deux phénomènes : d'une part, la miniaturisation de l'électronique permettant un développement exponentiel des puissances de calcul des ordinateurs et des capacités de stockage des serveurs et d'autre part, la circulation instantanée et la collecte à grande échelle des données, démultipliées par l'utilisation des Smartphones.

Les recherches sur l'ordinateur quantique permettant de dépasser les limites de la physique laissent penser que les vitesses de calcul vont encore augmenter et ce de manière très importante. Il faut conjuguer à ceci l'explosion des données personnelles qualifiées « d'or noir du XXIème siècle ».¹ Nouvelle manne financière dont les experts s'accordent cependant, à dire qu'elle n'est rien sans une méthode de tri efficace et pertinente, c'est-à-dire rien sans un bon algorithme. Ainsi, selon la société de conseil Gartner « les données sont intrinsèquement passives. Elles ne font et ne représentent rien sauf si vous savez comment les utiliser, comment agir sur ces données, car la véritable valeur réside dans les algorithmes »² qui réalisent des corrélations entre elles.

¹ <http://www.challenges.fr/high-tech/20140926.CHA8245/vos-donnees-personnelles-sur-internet-peuvent-valoir-de-l-or.html>

² <http://www.zdnet.fr/actualites/sans-algorithmes-le-big-data-ne-sert-a-rien-explique-le-gartner-39832842.htm>

Vitesse de calcul des ordinateurs et explosion des données personnelles ont permis l'arrivée d'une nouvelle catégorie d'algorithme : les algorithmes prédictifs.

Un algorithme prédictif est un programme mathématique permettant de calculer des scores prédictifs en fonction des différents types de données disponibles sur les individus étudiés et sur leur comportement de consommation ou d'usage³.

Les algorithmes prédictifs sont présentés par les experts comme des outils révolutionnaires et incontournables. Tous les secteurs d'activité sont concernés. Donnons quelques exemples. Les entreprises commerciales s'en servent pour ajuster au plus près l'offre, pour définir des profils de consommateurs en fonction de leurs goûts, de leurs centres d'intérêt, de leur appartenance à une communauté. Les banques et assurances les utilisent pour prédire les risques de fraude ou d'impayés. Les réseaux sociaux en font un grand usage pour mettre en relation les amis et « les amis de vos amis ». Les objets connectés dans le secteur de la santé fournissent une manne d'informations permettant tris, recoupements et ciblage en tout genre. Les services de police en font également usage pour prévoir notamment les zones de rassemblement lors de manifestations.

De cette énumération pourtant non exhaustive, il est possible de tirer l'enseignement suivant : le chiffre a pris une place immense dans notre société. Il est également possible d'identifier deux risques : une croyance inconditionnelle de la société dans le résultat mathématique ; l'opacité du fonctionnement des algorithmes. Or leurs usages peuvent conduire à influencer des comportements sans même que l'individu concerné en ait conscience ; ou peuvent conduire à des décisions discriminatoires.

Le traitement des données par les ordinateurs grâce aux algorithmes s'apparente pour un néophyte à de la magie. Or, la magie fait référence à des puissances cachées, à des croyances. Dès lors, comment accepter l'utilisation des algorithmes si on n'en comprend le fonctionnement ? Dans ce contexte, méfiance et suspicion sont de mise ; pourtant, il est certain que leur utilisation à grande échelle requiert un capital confiance. La transparence peut-elle être une réponse face aux difficultés de compréhension d'un algorithme ?

³ <https://www.definitions-marketing.com/definition/algorithme-predictif/>

Aujourd'hui, **la transparence dans la société de l'information** a pris une place majeure. La quête de transparence a pris de multiples facettes et ce besoin s'est répandu dans toutes les strates de la société.

L'État et l'administration s'y soumettent. Inhérente à la démocratie, la transparence contribuerait à la confiance des citoyens dans les institutions faisant échec aux privilèges liés à l'asymétrie de l'information. Ainsi Jeremy Bentham estime que « l'œil du public rend l'homme d'État vertueux » (J. Bentham, *The works of Jeremy Bentham*, J. Bowring, Edimbourg, W. Taitn, 1838-1843, vol. X, p145)⁵

Dans l'entreprise, la transparence n'est pas une donnée naturelle. Opacité, secret d'affaires, secret de fabrication ont longtemps dominé le monde de l'entreprise. Le capitalisme financier et la responsabilité sociale des entreprises sont à l'origine d'une plus grande diffusion des informations financières et comptables.

Pour les citoyens, le règne de la transparence trouve son point d'orgue avec Internet. Les réseaux sociaux permettent le développement du culte de la transparence. Avec la société de l'information on assiste à un renversement de principe. Alors que traditionnellement le secret prime et la publicité nécessite une action délibérée, les technologies de l'information et de la communication inversent le principe, (comme le montre l'accroissement des logiciels open-source ainsi que l'intérêt des utilisateurs envers eux). La transparence est devenue la règle (tracer, suivre, localiser, collecter, trier, traiter les données) au point que le secret nécessite une action. L'individu lui-même est devenu transparent.

Alors que l'absence de transparence est source de suspicion, de collusion, de risque de détournement de l'intérêt général au profit d'intérêts particuliers, la recherche de transparence absolue porte a contrario les germes du totalitarisme permettant l'avènement d'une société du contrôle où tout est su sur tout le monde. C'est tout le paradoxe de la transparence. Le bon équilibre est donc la conciliation des informations qui peuvent être partagées et celles qui doivent rester secrètes.

Un algorithme doit-il être transparent ? Peut-il être transparent ? Quels en sont les enjeux ?

Pour comprendre les enjeux, il faut d'abord définir la transparence et les algorithmes.

⁵ J.-F. Kerléo, « la transparence en droit », recherche sur la formation juridique d'une culture juridique, mare&martin, bibliothèque des thèses, 2015, 995p

I - Transparence et algorithmes : définir pour mieux comprendre les enjeux

A - Définition de la transparence

La transparence résulte d'un souci de moraliser les relations et de protéger les intérêts individuels .
Moraliser implique la diffusion d'informations pertinentes. Or, la publication d'informations peut aller à l'encontre du secret professionnel qui entoure les algorithmes ; elle est cependant indispensable car elle permet de trouver un équilibre entre les intérêts des usagers et le propriétaire de l'algorithme. Elle permet de lutter contre l'asymétrie informationnelle, source de suspicion.

Protéger ? Dans cette société gouvernée par le chiffre, que devient le citoyen ? Quelle liberté de pensée pour le citoyen dont l'information lui parvient en fonction de ses centres d'intérêt eux-mêmes identifiés grâce aux calculs algorithmiques ? Quel libre arbitre pour le consommateur ? Quand l'internet est construit initialement comme un espace de liberté basé sur l'échange collaboratif ; l'algorithme prédictif n'est-il pas susceptible de conduire à l'enfermement de la personne ? De manière plus large ce sont les questions relatives à la protection de la vie privée et à la (nécessaire ?) anonymisation des données qui se posent.

Deux exemples illustrent parfaitement le besoin de transparence qui donne aux usagers un pouvoir de contrôle.

Le premier exemple porte sur les conclusions d'une étude Suédoise⁷ ayant révélé que sur 6000 candidatures étudiées, les individus âgés de plus de 40 ans étaient exclus des procédures d'embauche ; ce qui aboutit à un processus de recrutement discriminatoire. Bien évidemment l'algorithme utilisé n'était pas public, et donc non transparent

Le second exemple concerne l'algorithme d'Admission Post-Bac (APB). En 2016, la communauté lycéenne a réclamé plus de transparence en ce qui concerne son code pour savoir sur quels critères les bacheliers étaient départagés, notamment pour les filières universitaires ayant recours au tirage au sort ou ayant une capacité d'accueil incompatible avec le nombre de demandes, et pour vérifier *in fine* si cette sélection était loyale. En effet, l'algorithme n'était, au départ, pas du tout transparent pour les utilisateurs, chose assez intrigante pour eux, étant donné l'impact que peuvent avoir ces résultats sur leurs vies. La transparence demandée passait évidemment par la publication du code de

l'algorithme mais aussi par sa spécification (c'est-à-dire la traduction et le détail du cahier des charges, des jeux et des résultats des tests utilisés) et même par une possibilité d'analyse, donc de tests de l'algorithme avec n'importe quelles données. Ces informations ont permis de mettre en lumière que certaines modalités de sélection ne respectaient pas les dispositions du code de l'éducation.

Le centre National des Ressources Textuelles et Lexicales définit la transparence comme la « propriété qu'a un corps, un milieu, de laisser passer les rayons lumineux, de laisser voir ce qui se trouve derrière ».

Au sens figuré c'est « la qualité d'une personne dont les pensées et les sentiments sont faciles à comprendre, à deviner » ; c'est aussi « la qualité d'une institution qui informe complètement sur son fonctionnement, ses pratiques » ; c'est encore « la qualité de ce qui est facilement compréhensible, intelligible ».

En droit, la transparence est associée à la notion de loyauté. Agir loyalement, c'est agir sans tricher. Et ne pas tricher consiste à ne pas dissimuler ; ce qui revient à être transparent.

La transparence est source de confiance ; l'absence de transparence source de suspicion. Selon Desmoulin, « la publicité est la sauvegarde du peuple » (maxime de Bailly reprise par Desmoulin, « le vieux cordeliers, n°7 p108) tandis que Bentham estime que « là où il n'y a pas de publicité, il n'y a pas de justice » (J. Bentham, Constitutional code, The works of Jeremy Bentham, 1843. J. Bowring, New-York, Russell & Russell, 1962, p.463, vol. IX) »¹⁰.

Un certain nombre de textes de loi fait référence à la notion de transparence sans que l'on dispose pour autant d'une définition uniforme.

Ainsi l'article 1er de la loi n° 2006-686 du 13 juin 2006 relative à la transparence et à la sécurité nucléaire définit la transparence comme « l'ensemble des dispositions prises pour garantir le droit du public à une information fiable et accessible en matière de sécurité nucléaire ».

Ce droit à l'information est également présent dans les relations des administrations avec le public. Les administrations garantissent le droit de toute personne à l'information en ce qui concerne la liberté d'accès aux documents administratifs (art. L.300-1 code des relations entre le public et

¹⁰ J.-F. Kerléo, « la transparence en droit », recherche sur la formation juridique d'une culture juridique, mare&martin, bibliothèque des thèses, 2015, 995p

l'administration). Or, depuis la loi n°2016-1321 du 7 octobre 2016 pour la République numérique, les codes sources sont des documents administratifs¹¹.

Cette même loi fait peser sur les opérateurs de plateforme une obligation de délivrer du profit du consommateur « une information loyale, claire et transparente sur :

1° Les conditions générales d'utilisation du service d'intermédiation qu'il propose et sur les modalités de référencement, de classement et de déréférencement des contenus, des biens ou des services auxquels ce service permet d'accéder [...]» (art. L. 111-7 code de la consommation).

En l'absence de définition unique, il est néanmoins possible de relever que la notion de transparence en droit est liée au devoir d'information : délivrer une information claire et complète.

En informatique, c'est la notion de responsabilité qui prévaut. Un algorithme est responsable s'il respecte la confidentialité des données, la non-discrimination ET s'il se conforme à certaines règles éthiques comme la neutralité et la loyauté (*Nozha Boujema, DR/INRIA, Advisor to the CEO in Big Data*).

Si la transparence nécessite la diffusion d'informations claires afin de susciter la confiance ; nous verrons que le caractère protéiforme et évolutif d'un algorithme rend particulièrement complexe l'objectif de transparence des algorithmes.

B - Définition et typologie des algorithmes

Un algorithme est défini par la CNIL comme « la description d'une suite d'étapes permettant d'obtenir un résultat à partir d'éléments fournis en entrée ». De manière simpliste, une recette de cuisine est un algorithme permettant d'obtenir un plat à partir d'ingrédients.

Il est possible de classer les algorithmes en deux grandes catégories : les algorithmes déterministes et non déterministes.

¹¹ Article L300-2 code des relations entre le public et l'administration Modifié par LOI n°2016-1321 du 7 octobre 2016 - art. 2 « Sont considérés comme documents administratifs, au sens des titres Ier, III et IV du présent livre, quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions, codes sources et décisions ».

Un *algorithme déterministe* se comporte de façon prévisible en accomplissant un processus défini au préalable par le concepteur pour résoudre un problème donné. Il donne une « valeur unique » pour n'importe quelle entrée dans son « intervalle de définition ». De ce fait on dit qu'il calcule une fonction mathématique. On l'assimilera naturellement à un automate avec un nombre fini d'états (un état étant ce qui décrit l'action de la machine à un moment donné). En effet, dans le cas déterministe, la transition entre deux états se fait de manière discrète et précise ; l'état dit courant étant entièrement déterminé par le précédent. Par ailleurs ici, l'état initial est unique. Cependant, le déterminisme peut être contraignant et restrictif, d'où la mise en œuvre d'algorithmes non-déterministes.

Un *algorithme non-déterministe* se comporte, de façon non-prévisible : il peut produire des résultats différents pour une même entrée. C'est un automate fini non déterministe : le chemin d'états que l'algorithme va suivre pour une entrée donnée n'est pas prédéfini et est juste une possibilité parmi d'autres. Il y a en effet une part plus ou moins importante d'aléa. Certains algorithmes non déterministes ne sont pas basés sur l'aléa, mais leur complexité combinatoire rend difficile voire impossible la prédiction des résultats.

À partir de ces définitions, il est possible à dire qu'un algorithme déterministe est transparent. Lorsqu'il est rendu public la connaissance des données d'entrée et de l'algorithme permet d'en déduire son résultat. S'il est privé, des tests peuvent quand même « l'éclaircir » car le caractère reproductible des résultats (les mêmes entrées produisent les mêmes sorties) offre un certain degré de compréhension.

Concernant les algorithmes non déterministes les choses sont beaucoup plus compliquées. Pour les algorithmes Evolutifs, Génétiques, Chimiques, de type Machine Learning la structure même des algorithmes évolue, ce qui revient à dire que l'automate évolue en permanence.

On comprend que transparence et algorithme ne sont pas toujours compatibles. Pour autant peut-on se satisfaire de ce constat et accepter que certains algorithmes restent des boîtes noires alors même qu'ils calculent des résultats qu'il faut considérer vrais. La transparence n'est-elle pas la garantie contre l'arbitraire ?

II - Transparence et algorithmes : un mariage difficile

Ainsi, la structure de l'algorithme conditionne sa transparence rendant le degré de transparence très variable voire même nul dans certains cas (A).

En fonction du degré de transparence, il convient de s'interroger sur les garanties offertes au citoyen(B)..

A - Un niveau de transparence nécessairement lié à la structure de l'algorithme

Si l'on reprend notre classification initiale des algorithmes, nous avons les algorithmes déterministes et non-déterministes. Quand un algorithme déterministe est public, il est transparent à condition d'avoir accès à la fois au code, à l'automate et à l'algorithme lui-même. A contrario, un algorithme non déterministe et non public n'est pas transparent.

Il est alors possible d'établir un premier tableau :

	Public	Non Public
Déterministe	Transparent	
Non déterministe		Non transparent

Entre ces deux bornes, transparence acquise d'une part et transparence impossible d'autre part, une hiérarchie de transparence peut être établie entre les différents types d'algorithmes.

Pour les algorithmes non déterministes, il est nécessaire de faire plusieurs distinctions qui vont déterminer un degré d'opacité. Il faut distinguer les algorithmes évolutifs (également appelés évolutionnistes) et probabilistes.

Un algorithme évolutif ou dit évolutionniste s'inspire de la théorie de l'évolution de Darwin. En partant d'une population déterminée par avance ou en partant de données indéterminées, on définit des opérateurs de variation qui vont engendrer de nouvelles configurations. Un opérateur de sélection élimine la configuration la moins bonne. Les opérateurs de sélection sont variables. Ils peuvent être fondés sur l'élitisme (on garde le meilleur individu d'une population), sur la probabilité (on calcule un coefficient de probabilité permettant une sélection d'individus), sur le hasard ou encore sur le rang (ranking).

Le caractère évolutif de l'algorithme s'explique par la redéfinition de la population opérée à partir des opérateurs de sélection ; ce qui influence sur l'algorithme lui-même, c'est-à-dire sur son fonctionnement intrinsèque.

À cela, il convient d'ajouter que l'algorithme peut être supervisé permettant plus ou moins de contrôler les évolutions de l'algorithme lui-même afin d'éviter les dérives non souhaitées.

Concernant les algorithmes probabilistes, deux grandes familles existent :

- Las Vegas : pour une entrée donnée, le résultat attendu doit être correct ; la probabilité porte sur le temps de calcul. L'algorithme calcule une réponse correcte en un temps a priori aléatoire.
- Monte Carlo : pour une entrée donnée, le résultat peut différer et même être faux (pour une certaine probabilité)

Les deux algorithmes ont un caractère probabiliste par définition. Cependant pour le premier, il apparaît clairement qu'il peut être comparé à un algorithme déterministe dans le sens où pour une donnée fournie, le résultat est toujours le même. Le second ne peut l'être car pour une même entrée, le résultat peut différer.

On obtient alors le tableau de transparence suivant :

Non déterministe/Las Vegas	Non déterministe/Monte Carlo	Non déterministe/Evolutif Supervisé	Non déterministe/Evolutif non supervisé
Assez transparent	Peut être transparent	Difficilement transparent	Non transparent

À partir de ce tableau, on comprend que la transparence dépend principalement de la typologie de l'algorithme et il apparaît clairement que le degré de transparence de l'algorithme a un impact direct sur les garanties envisageables.

B - Un niveau de garantie nécessairement lié au degré de transparence des algorithmes

Quand l'algorithme peut être transparent, il faut s'interroger sur les critères qui garantissent l'effectivité de la transparence et lorsque l'algorithme n'est pas transparent, il convient de s'interroger sur les garanties envisageables.

1. Les critères garantissant l'effectivité de la transparence

Rappelons en préalable que la recherche de transparence a une vocation protectrice des droits de la personne concernée par la mise en œuvre d'un algorithme. La transparence peut aussi servir l'intérêt général en favorisant la concurrence et l'innovation. Néanmoins, il ne faut pas occulter que

la transparence peut s'opposer aux intérêts économiques de la personne ou la société détentrice de l'algorithme.

En droit, la transparence peut s'opérer à plusieurs niveaux :

1. Être informé de l'usage d'un algorithme ;
2. Avoir accès aux « modalités de fonctionnement ».

La première étape prend la forme d'un droit à l'information ; la deuxième étape est un droit d'accès. On comprend aisément que le second est conditionné par le premier. Étant informés de la mise en œuvre d'un algorithme, les individus concernés sont à même de demander un droit d'accès.

Être informé de la mise en œuvre d'un algorithme nécessite de définir les renseignements pertinents à transmettre. Connaître la finalité de l'algorithme est notamment un critère déterminant.

Avoir accès à l'algorithme mis en œuvre suppose d'identifier les informations pertinentes permettant d'assurer la transparence d'un algorithme.

Pour un programme informatique donné, si l'algorithme sur lequel il se base est déterministe, il possède d'un point de vue informatique toutes les caractéristiques pour être transparent. Néanmoins, cela n'est pas suffisant car il existe différents niveaux de description d'algorithme, des plus généraux aux plus détaillés. Ainsi, pour atteindre une transparence effective, la publication du code et d'une version détaillée et/ou explicite de l'algorithme est une nécessité pour « comprendre » ce que réalise et comment fonctionne le programme informatique. Un algorithme trop général ou un code trop complexe ne sont pas suffisants. En effet, la complexité peut être telle qu'une absence de détail (ou au contraire d'abstraction) rend l'algorithme (ie. le programme) incompréhensible et donc...non transparent.

Dans certains cas, l'absence de publicité de l'algorithme n'est pas un obstacle à sa compréhension. Ainsi, Qwant, le moteur de recherche Européen est un algorithme neutre et impartial, ne tenant pas compte des historiques, du profil psychologique de l'utilisateur, des annonceurs, des opinions, etc. Son code (pour une grande partie) est accessible à tous mais les algorithmes ne sont pas publics. Néanmoins la grande communauté fédérée autour de Qwant ainsi que le nombre de publications relatif à son fonctionnement donnent autant de renseignements que la publication de l'algorithme lui-même, ce qui le rend...transparent.

En résumé, le simple accès à l'algorithme n'est pas une garantie de transparence. Ainsi pour garantir une transparence effective les informations pertinentes sont l'accès au code, à l'algorithme dans sa version détaillée.

Cela correspond à la description d'une situation idéale mais comme nous l'avons vu précédemment un certain nombre d'algorithmes n'offre pas les critères de transparence. Quelles sont alors les garanties envisageables ?

2. Les garanties envisageables en l'absence de transparence

La finalité de la transparence face aux algorithmes est la protection du droit des personnes : garantie d'égalité face à un traitement algorithmique, absence de discrimination.

Or nous l'avons vu certains types d'algorithmes ne sont pas transparents.

D'une part, le caractère évolutif d'un algorithme rend impossible la compréhension du résultat fourni. Ainsi, dans le domaine du « morpho engineering », bien que l'algorithme de duplication des cellules soit connu (incluant la réplication d'ADN avec des erreurs statistiques), le nombre de possibilités est tel qu'il est impossible de prévoir les chimères résultantes. Il s'agit ici d'algorithmes évolutifs non supervisés. D'autres algorithmes liés au Big Data et relatifs à l'identification de traitements contre le cancer, intègrent des milliers de médicaments, étudient les interactions entre eux afin de proposer de nouveaux traitements contre le cancer [Marr, 2015]. Ils utilisent la plateforme d'IA Watson d'IBM utilisant le Big Data et ses algorithmes statistiques, sans pour autant en maîtriser les algorithmes utilisés.

D'autre part, certains algorithmes font l'objet de brevet ; ils sont donc couverts par le secret des affaires.

Dans les cas où la transparence n'est pas possible (ou difficilement atteignable) il faut imaginer des jeux de tests permettant de vérifier le respect du droit des personnes. Cela ne rendra certes pas l'algorithme transparent mais permettra de diminuer le degré d'opacité de ce dernier, et par là même d'augmenter le degré de confiance que nous pouvons en avoir.

Conclusion : Quel avenir pour la transparence ?

La « tendance » actuelle est à l'utilisation de données massives (« Big Data »). Il s'agit de récupérer tout ce que l'on peut récupérer comme données (structurées ou non), d'y appliquer des formules mathématiques, statistiques ou algorithmes, de constater le résultat puis d'éventuellement l'exploiter. De très nombreux logiciels – tous domaines confondus - aujourd'hui utilisent ces « Big Data ». Face à la volumétrie, les algorithmes développés sont pour la plupart statistiques, à base de réseaux de neurones et/ou d'apprentissage rendant leur transparence difficile voire impossible (algorithme évolutifs supervisés ou non).

Associé à ces masses de données se pose un autre problème lié à l'anonymisation. En effet, la plupart des traces numériques que nous laissons peuvent être exploitées, décuplant les risques de discrimination.

Le défi des années à venir se situe très certainement autour de ces algorithmes et de ces données. D'ores et déjà de nombreux travaux de recherches visent à faire en sorte que les données « laissées » sur le web soient « anonymisées » afin d'éviter les recoupements et ainsi de faciliter la loyauté. Mais cela ne résout qu'un seul problème laissant la transparence s'éteindre de plus en plus.

Il semble urgent pour garantir la protection du droit des personnes et peut-être même la démocratie de réfléchir à une gouvernance des algorithmes dans laquelle le poids des pouvoirs publics soit fort afin d'éviter la main mise (si c'est encore possible) de sociétés dont les intérêts économiques sont évidents.

Il est certain que l'utilisation de données massives (« Big Data ») associée aux algorithmes évolutifs rendra aigue la question de la transparence des algorithmes.

L'algorithme a une dimension technique mais les choix techniques posent toujours des questions de sens, c'est-à-dire des questions politiques et ce serait une grave erreur de voir dans un algorithme la simple mise en œuvre d'une solution technique.

Sitographie

<https://www.scriptol.fr/programmation/algorithmes-classification.php>

<https://www.scriptol.fr/programmation/liste-algorithmes.php>

<http://www.meta-media.fr/2017/01/31/algorithmes-lheure-de-la-grande-regulation.html>

<http://www.internetactu.net/2016/03/16/algorithmes-et-responsabilites/>

<http://internetactu.blog.lemonde.fr/2016/12/03/comment-rendre-les-algorithmes-responsables/>

<https://www.labri.fr/perso/duchon/Teaching/ENSEIRB/AlgoProba/Poly.pdf>

http://ufrsciencestech.u-bourgogne.fr/master1/mi1-tc5/CM2009/Genetiques/Genetique_1.pdf

Bibliographie

S. Arpin, « vers le règne de la transparence », revue-le banquet, n°25, fev. 2008

O. Cognasse, O. James, H. Meddah, M. Moragues, J. Thoin-Bousquié « intelligence artificielle, cerveau à saisir », L'usine nouvelle, n°3517, 25 mai 2017

J.-M. Sauvé, « Culture du secret contre transparence sans limite : quel équilibre pour garantir l'intérêt général » exposé d'ouverture, colloque organisé par Transparence International France, Assemblée Nationale, 5 Jull. 2011

L. Cohen-Tanugi, « le clair-obscur d'internet », pouvoirs, n°97, 2001

G. Carcassonne, « le trouble de la transparence », pouvoirs, n°97, 2001

D. Kessler, « l'entreprise entre transparence et secret », pouvoirs, n°97, 2001

B. Marr - How Big Data Is Transforming The Fight Against Cancer -

<https://www.forbes.com/sites/bernardmarr/2015/06/28/how-big-data-is-transforming-the-fight-against-cancer/#3230fdd31d4f>