

# Algorithmes

## Généralités

### Définition minimale et vidéo didactique

Algorithme : « une opinion enfermée dans une formule mathématique » (Cathy O'Neil).

À travers la métaphore de la préparation d'un repas destiné à ses enfants, Cathy O'Neil, mathématicienne états-unienne, explique comme l'algorithme naît de la combinaison de deux éléments : des données organisées et une définition de « succès ». Cathy O'Neil est l'auteure de *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown Publishers, [2016].

VENERUSO, Valerio, 2018. La verità sugli algoritmi. Un'animazione didattica su un tema attuale. *Artribune* [en ligne]. 1<sup>er</sup> novembre 2018. [Consulté le 1<sup>er</sup> novembre 2018]. Disponible à l'adresse : <https://www.artribune.com/television/2018/11/video-the-truth-about-algorithms/>

### Historique

Les premières traces d'algorithmes ont été retrouvées chez les Babyloniens (l'actuel Irak) au deuxième millénaire avant notre ère et étaient principalement des méthodes de calcul pour le commerce et les impôts. Il faut attendre le troisième siècle avant J.-C. en Grèce pour l'apparition du fameux algorithme d'Euclide pour le calcul du PGCD : on peut considérer qu'il s'agit du premier algorithme « moderne » et il est tout à fait remarquable qu'il soit toujours utilisé de nos jours.

Mille ans plus tard, au IX<sup>e</sup> siècle ap. J.-C., Al Khuwarizmi, un mathématicien perse (actuel Iran), publie un ouvrage consacré aux algorithmes : l'étymologie du terme « algorithme » vient du nom de ce mathématicien. On commence en effet à étudier les algorithmes en tant que tels, mais il faudra encore 800 ans avant que l'Occident continue cette étude.

PERIFEL, Sylvain, 2014. *Complexité algorithmique* [en ligne]. [Consulté le 2 novembre 2018] Disponible à l'adresse : <https://www.irif.fr/~sperifel/complexite.pdf>

## Quelques effets des algorithmes sur nos vies

### Les algorithmes sentimentaux, vers la quatrième génération

L'essentiel des algorithmes actuels sont centrés sur la prescription<sup>25</sup> et ne se soucient plus qu'à la marge – publicitaire – de répondre correctement à nos requêtes. Leur ADN algorithmique est de s'arranger pour que la prescription sur ordonnance (des résultats) l'emporte aussi souvent que possible sur l'improbable satisfaction de voir s'afficher des réponses à des questions que de toute façon nous ne posons même plus<sup>26</sup>. En quelque sorte des algorithmes « d'imposition ». L'EdgeRank de Facebook<sup>27</sup> est emblématique de cette troisième génération algorithmique puisqu'il s'agit, dans un écosystème presque entièrement dénué de requêtes, de déterminer quels contenus sont les plus susceptibles d'intéresser tels types de profils.

C'est alors qu'apparaissent les problématiques des techniques dites « d'opinion mining » et de « sentiment analysis »<sup>28</sup>. Une quatrième génération d'algorithmes qui ont pour but non plus de croiser (matching), non plus de prescrire, non plus d'imposer en dehors de tout requêtage, mais de détecter dans des instanciations de discours divers et épars, des « opinions » et des « sentiments ». Des algorithmes « sentimentaux ». Qui tentent pour l'instant encore maladroitement de singer des phénomènes interpersonnels de connivence<sup>29</sup>. Des algorithmes sentimentaux dans la roue d'un web de l'émotion, de l'affect, du pathos, de la pulsion<sup>30</sup>. Aujourd'hui ces algorithmes, ceux en tout cas qui permettent de travailler et de visualiser ces « analyses de sentiment » en les quantifiant, sont réservés à quelques professionnels de l'information et sont principalement mis en place dans quelques outils logiciels à destination des mêmes professionnels.

Mais de la même manière que l'on fut initialement incapable de détecter que le PageRank était déjà un algorithme de prescription (2<sup>e</sup> génération) inaugurant un nouvel ordre documentaire en n'y voyant qu'un super algorithme de recherche (1<sup>e</sup> génération), les environnements numériques de demain seront à leur tour régis par des algorithmes sentimentaux (4<sup>e</sup> génération) capables de fabriquer de l'opinion et du (res)sentiment sous couvert de simplement nous imposer un enchevêtrement complexe « d'opinions » et « de sentiments » (3<sup>e</sup> génération). Et ce pour une raison simple :

« Aujourd'hui, ce qui préoccupe les grands acteurs de l'économie de l'attention, c'est la manière dont des motifs récurrents (patterns) peuvent être extraits d'un ensemble disparate de singularités (d'individus), placés *au cœur* de l'écosystème informationnel et non à sa sortie ou à sa périphérie, pour

optimiser la capacité de satisfaction immédiate que ledit système (le medium) peut apporter aux individus sous forme de récompense cognitive<sup>31</sup>, aliénant ainsi leur attention tout en lui assignant un coût cognitif quasiment nul. »<sup>32</sup>

ERTZSCHEID, Olivier. Usages de l'information numérique : comprendre les nouvelles enclosures algorithmiques pour mieux s'en libérer. *Revue française des sciences de l'information et de la communication* [en ligne]. 2015. N° 6. Usages et usagers de l'information à l'ère numérique. [Consulté le 22 juillet 2018]. Disponible à l'adresse : <https://journals.openedition.org/rfsic/1425#bodyftn5>

## AI et subconscient

Un team di ricercatori del MIT di Boston ha annunciato la creazione della "prima intelligenza artificiale psicopatica". Il software in questione si chiama Norman, in omaggio al protagonista di *Psycho*. (...) Nel caso del povero Norman, esposto volontariamente a una serie di contenuti violenti provenienti da un forum online dedicato al tema della morte, il risultato è un sistema che "vede" uccisioni, incidenti e sangue in qualsiasi immagine, anche astratta, che gli venga mostrata. Dove un altro software, allenato con dati diversi, vede "un gruppo di uccelli seduto sulla cima di un ramo", Norman vede "un uomo che muore fulminato".

(...)

Un simile esperimento era stato condotto nel 2016 su Tay, un Twitter bot lanciato da Microsoft. (...)

Google Deep Dream (...)

Oggi, però, non si discute soltanto dell'ipotesi che i computer possano sviluppare il dono della coscienza e prendere decisioni autonome, ma ci si inoltra direttamente nei territori del subconscio. Se infatti il compito principale dell'intelligenza artificiale è quello di simulare il funzionamento della mente umana, questa simulazione implica necessariamente l'esistenza di un livello inconscio, inaccessibile ai controlli della parte cosciente. Ed è proprio su queste dinamiche che la ricerca sull'AI sta lavorando.

Paul Dolan, professore di Scienze comportamentali alla London School of Economics and Political Science, ha così commentato un recente studio sul tema commissionato dalla Huawei in occasione del lancio dello smartphone Mate 10 Pro, il primo dotato di un chip AI: "I risultati sono affascinanti perché dimostrano che la nostra mente processa molte cose a livello non cosciente". Questi meccanismi che lavorano "in background", tanto utili per velocizzare le operazioni di routine, abitano gli stessi territori da cui provengono incubi e visioni. Proprio come nei romanzi di Philip K. Dick, dove l'attività che rende gli androidi più simili agli umani è quella onirica, il punto di contatto tra uomo e macchina potrebbe trovarsi sepolto, da qualche parte, nelle profondità dell'inconscio.

TANNI, Valentina, 2018. Il subconscio artificiale. *Artribune* [en ligne]. 13 août 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : <https://www.artribune.com/progettazione/new-media/2018/08/subconscio-intelligenza-artificiale/>

## Nous, travailleurs bénévoles pour algorithmes

On sait depuis longtemps que ces Captcha, loin de leur finalité initiale (distinguer utilisateurs humains de robots malveillants), sont aussi et surtout un moyen d'entraîner des technologies ... diverses. Technologies linguistiques principalement où, reconvertis en travailleurs numériques bénévoles, nous "aidons" à affiner d'immenses corpus, nous permettons aux technologies de reconnaissance optique de caractère (OCR) de progresser et donc aux acteurs économiques disposant desdites technologies d'être plus performants. → Il y a quelques années de cela, de nouvelles "Captcha" sont arrivées. Sous forme de tableau découpé en 9 ou 12 cases, on y voit des paysages, des immeubles d'habitation, des panneaux de signalisation, et l'on nous demande de "cliquer" sur les cases qui comportent ou ne comportent pas tel ou tel élément.

ERTZSCHEID, Olivier, 2018. I'm a digital worker, killing an arab. Chronique de la guerre algorithmique. *Affordance.info* [en ligne]. 13 mars 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : [https://www.affordance.info/mon\\_weblog/intelligence-artificielle/](https://www.affordance.info/mon_weblog/intelligence-artificielle/)

GUILLAUD, Hubert, 2018. Concrètement, comment rendre les algorithmes responsables et équitables ? *InternetActu* [En ligne]. 19 juillet 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : <http://www.internetactu.net/2018/07/19/concretement-comment-rendre-les-algorithmes-responsables-et-equitables/>

**Difficile maîtrise d'algorithmes opaques, impénétrables, trop nombreux et mouvants pour être répertoriés et capables d'influencer le comportement humain<sup>1</sup> :**

<sup>1</sup> « Les nouvelles "approches algorithmiques" ou "l'algorithme" ne désigne plus simplement la trivialité d'une série d'instructions logico-mathématiques mais un environnement complexe d'automatismes décisionnels sur la base de gigantesques jeux de

Dresser l'Atlas des algorithmes contemporains, de tous les domaines où ils ont une influence, est déjà devenu quasiment impossible.

Les inconvénients des effets de ces algorithmes sont bien sûr nombreux. D'abord, il y a l'opacité. C'est-à-dire que nous ne savons pas ce qu'ils font.

Autre inconvénient : l'impénétrabilité, c'est-à-dire le fait que les algorithmes ne soient pas lisibles et ne correspondent pas à la manière dont les humains réfléchissent.

Désormais, les mathématiques disent qu'ils peuvent influencer le goût et la préférence des gens. « 60 % des films loués sur Netflix sont des films qui ont été mis en avant pour vous par l'algorithme de Netflix en se basant sur ce que d'autres et vous-mêmes avez déjà apprécié ». Pire, une société comme [Epagogix](#) permet d'évaluer un film avant même qu'il soit fait : on introduit un scénario dans l'ordinateur et celui-ci peut vous dire le profit qu'il devrait générer...

Trop tard pour en sortir ?

GUILLAUD, Hubert, 2011. Kevin Slavin : « il nous faut dresser l'Atlas des algorithmes contemporains ».

*InternetActu* [En ligne]. 24 février 2011. [Consulté le 3 novembre 2018]. Disponible à l'adresse :

<http://www.internetactu.net/2011/02/24/kevin-slavin-il-nous-faut-dresser-latlas-des-algorithmes-contemporains/>

### Failibilité (mortelle) des algorithmes

La mort d'[Elaine Herzberg](#), première humaine tuée par une voiture autonome en mars 2018 ([voir l'enquête qu'en dressait l'écrivain Thierry Crouzet](#)).

La machine apprend à faire ce qu'on lui demande de faire, comme obtenir le meilleur score possible à un jeu en y jouant des millions de fois. Mais ce qu'elle en apprend n'est pas transférable d'un jeu à un autre. Et elle ne sait expliquer comment elle y arrive. Le problème, c'est que quand l'algorithme apprend, nous ne savons plus quels sont les règles et les paramètres qu'il utilise puisqu'il les transforme. Ces algorithmes-là ne sont plus prévisibles et peuvent se mettre à produire des résultats erratiques. Pour Smith ce sont des « algorithmes Frankenstein », des machines qui créent leurs propres règles, comme ceux qui dominent désormais les marchés financiers via [les transactions à haute fréquence](#). (...) Pour la sociologue [Lucy Suchman](#) de l'université de Lancaster, ces problèmes sont partout. [Une étude sur les attaques par drones au Pakistan entre 2003 et 2013](#) a ainsi montré que 2 % des personnes tuées par drones étaient des cibles présentant une menace. 20 % étaient considérés comme des non-combattants... Et 75 % étaient des cibles... « non identifiés ». Pour elle, nous sommes face à une technologie d'identification très grossière.

GUILLAUD, Hubert, 2018. De l'imbrication algorithmique<sup>2</sup>. *InternetActu* [En ligne]. 12 septembre 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : <http://www.internetactu.net/a-lire-ailleurs/de-limbrication-algorithmique/>

### Les différents coûts de la gouvernance algorithmique

Pour le dire autrement, à l'échelle de l'hypothèse d'une gouvernementalité algorithmique qui serait ou se verrait souveraine, le risque de faux-positifs est perçu et traité comme un biais nécessaire pour optimiser la gouvernance ; alors que dans une société démocratique le risque de faux-positif est un signal qui permet de repérer et d'éviter des dysfonctionnements législatifs, politiques et institutionnels. Une démocratie non-algorithmique n'acceptera jamais de mettre 10 innocents en prison si elle a la certitude qu'un seul coupable figure parmi eux, une gouvernance algorithmique considèrera que c'est un mal nécessaire pour optimiser statistiquement la mise en sécurité de la population. Mais tout cela a déjà été très bien décrit dans nombre d'ouvrages de science-fiction. Même si la science-fiction d'hier a laissé place à des réalités politiques (en [Chine notamment avec le système de "Social Credit"](#)) ou technologiques (avec, parmi tant d'autres, [le récent score de fiabilité mis en place par Facebook](#)).

ERTZSCHEID, Olivier, 2018. Les algorithmes du dr Frankenstein (ou pourquoi faut pas pousser la créature algorithmique dans les datas). *Affordance.info* [en ligne]. 14 septembre 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : [https://affordance.typepad.com/mon\\_weblog/2018/09/algorithmes-frankenstein.html](https://affordance.typepad.com/mon_weblog/2018/09/algorithmes-frankenstein.html)

### Quelques institutions participant au débat sur l'éthique et l'intelligence artificielle

---

données (Datasets) que plus personne n'est en mesure de comprendre ou d'englober dans leur totalité, à commencer par ceux-là mêmes qui développent lesdits algorithmes et jeux de données.» ERTZSCHEID, Olivier, 2018. Les algorithmes du dr Frankenstein (ou pourquoi faut pas pousser la créature algorithmique dans les datas). *Affordance.info* [en ligne]. 14 septembre 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : [https://affordance.typepad.com/mon\\_weblog/2018/09/algorithmes-frankenstein.html](https://affordance.typepad.com/mon_weblog/2018/09/algorithmes-frankenstein.html)

<sup>2</sup> Cet article propose une synthèse de : SMITH, Andrew, 2018. Franken-algorithms: the deadly consequences of unpredictable code. *The Guardian* [en ligne]. 30 août 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : [https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger?CMP=share\\_btn\\_tw](https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger?CMP=share_btn_tw)

CERNA (Commission de réflexion sur l'Éthique de la Recherche en sciences et technologies du Numérique d'Allistene. [Consulté le 2 novembre 2018] Disponible à l'adresse : <http://cerna-ethics-allistene.org>

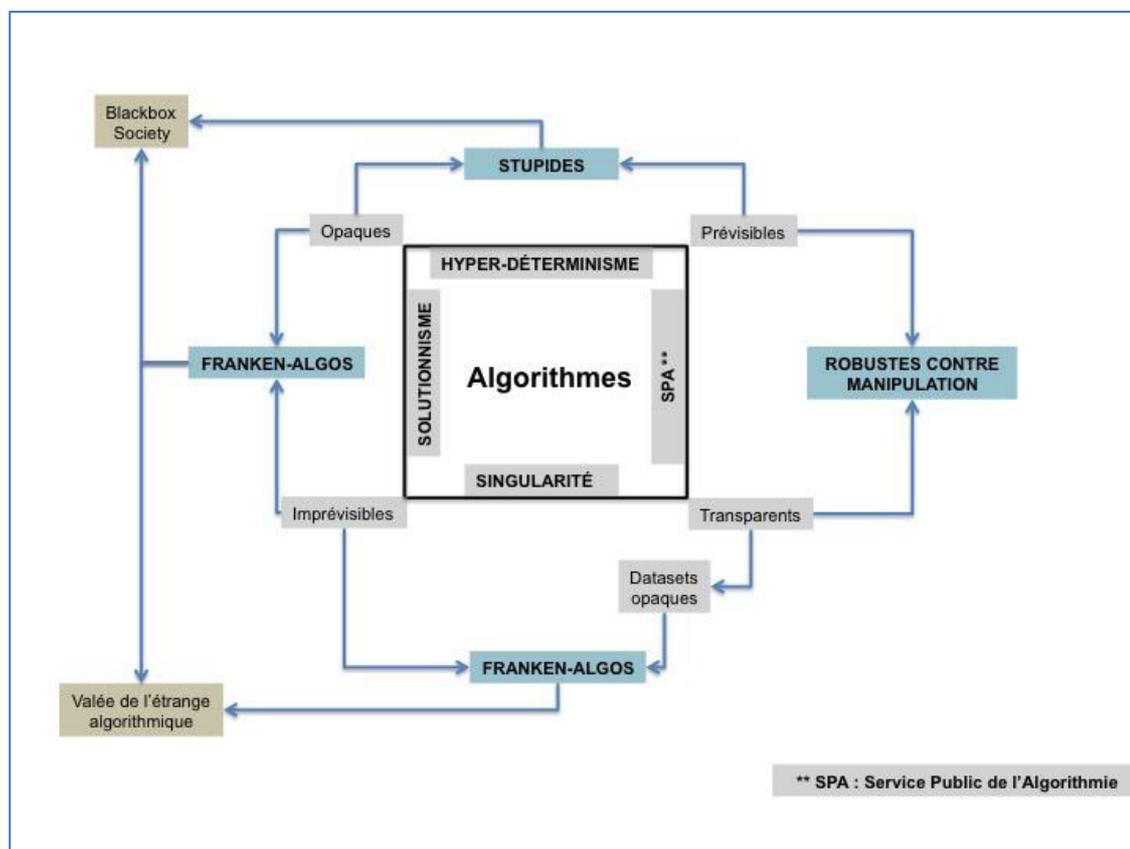
CNIL (Commission Nationale de l'informatique et des Libertés), 2017. Éthique et Intelligence artificielle. CNIL [site officiel, en ligne]. [Consulté le 2 novembre 2018]. Disponible à l'adresse : <https://www.cnil.fr/fr/ethique-et-intelligence-artificielle>

**Présentation du colloque « Gouvernance des algorithmes »**, Paris, 1<sup>er</sup> février 2016. [Consulté le 2 novembre 2018]. Disponible à l'adresse : <http://cerna-ethics-allistene.org/page+CERNA+gouvernance+algorithms/>

Les algorithmes sont conçus par des humains et au cœur même de tous nos systèmes d'information. Ils utilisent, sans que souvent nous le réalisions, les données produites massivement dans tous les domaines de nos sociétés : des données économiques, écologiques, scientifiques, techniques, personnelles, financières, ... A base de traitements statistiques et qui peuvent être largement distribués de par le monde, ces algorithmes infèrent des corrélations à caractère prédictif voire autoréalisateur mais non nécessairement explicatif, qui orientent les choix et les comportements individuels ou collectifs : publicité ciblée, déplacements intelligents, préconisations de santé, détermination de tarifs, de politiques publiques ou privées. La gouvernance des algorithmes devient un sujet majeur d'attentions, notamment éthiques, pour toute la société mais tout particulièrement pour celles et ceux qui conçoivent et déploient ces technologies, c'est-à-dire les chercheurs et développeurs du numérique.

### Carré sémiotique des algorithmes

Olivier Ertzscheid, maître de conférences à l'université de Nantes et à l'Institut universitaire de technologie de La Roche-sur-Yon propose un carré sémiotique construit sur ces notions d'opacité et de prévisibilité et leurs contraires (transparence et imprévisibilité donc) :



**Les algorithmes à la fois opaques et prévisibles sont donc en effet "stupides"**. Parmi les exemples on pourra citer le processus qui, sur Facebook, classe systématiquement comme "pornographie" toute photo ou représentation d'organes génitaux masculins et féminins, incapable ainsi de reconnaître des œuvres d'art ("L'origine du monde" de Courbet) ou des photos à caractère historique (comme celle du prix Pulitzer de la petite Kim fuyant les bombardements au Vietnam).

Ces formes algorithmiques, certes "sophistiquées" dans leur programmation mais essentiellement primaires et grégaires dans leur(s) application(s), participent d'un hyper-déterminisme qui permet aux plateformes de s'abriter derrière la barrière de leurs CGU tout en "fabriquant" de nouvelles normes sociales.

**Les Franken-Algorithmes sont imprévisibles et le plus souvent opaques** mais peuvent également être "transparents" à partir du moment où ce sont les jeux de données (Datasets) leur permettant de "fonctionner" qui, eux, basculent dans l'opacité de régimes propriétaires (comme je l'expliquais en Mai 2016, [tous les GAFA ont en effet basculé l'essentiel de leur coeur algorithmique en Open Source](#)).

Imprévisibles et opaques ils alimentent l'**idéologie solutionniste** (= à tout problème une solution techno-algorithmique). Imprévisibles et "transparents" mais reposant sur des jeux de données propriétaires ils nourrissent l'**idée une singularité à venir** qui s'empare des avancées *réelles* dans le domaine de [l'apprentissage par renforcement](#) ou des algorithmes génétiques pour alimenter et nous vendre le fantasme d'une IA forte et toute-puissante (et dangereuse).

**La conjonction d'une vision solutionniste et d'un hyper-déterminisme algorithmique** caractérise à son tour la ["Black Box Society" décrite remarquablement par Frank Pasquale](#).

**Et au croisement du solutionnisme et de la singularité** on entre, *de facto*, dans une forme de [vallée de l'étrange](#) algorithmique.

Dans un monde régulé par la technologie, ou en tout cas par ces formes technologiques particulières que sont les algorithmes et les architectures techniques toxiques et propriétaires dans lesquelles ils se déploient, le danger vient de l'**incapacité à penser simultanément l'ensemble des processus à l'œuvre** tant sur le plan des internalités logico-mathématiques (qui sont devenues imprévisibles) que des externalités (les interactions avec le monde "réel", également imprévisibles et potentiellement chaotiques).

**Reste, heureusement, la possibilité d'algorithmes à la fois prévisibles et transparents**, et reposant également sur des jeux de données également non-propriétaires, qui sont alors la seule garantie d'être réellement robustes contre toute manipulation ou contre toute dérive dès lors qu'ils ont vocation à intervenir dans des interactions sociales de nature régaliennne (transports, santé, éducation, etc.)

Comme le rappelaient en effet Bostrom et Yudowsky (deux théoriciens de l'intelligence artificielle) dont je cite souvent ici l'article de 2011, ["The Ethics of Artificial Intelligence"](#) :

"Les algorithmes de plus en plus complexes de prise de décision sont à la fois souhaitables et inévitables, tant qu'ils restent **transparents à l'inspection, prévisibles pour ceux qu'ils gouvernent, et robustes contre toute manipulation.**"

Voilà ce qui, sur mon petit schéma donne naissance au "SPA", une sorte de "Service Public de l'Algorithmie" que mon côté Marxiste idéalise peut-être mais que je crois nécessaire et urgent de bâtir si nous voulons éviter de prendre le risque d'un Hiroshima technologique désormais plus que probable à l'échelle de nos formes démocratiques de gouvernance qui ont déjà en partie basculé dans les formes assumées et délétères de surveillance.

ERTZSCHEID, Olivier, 2018. Les algorithmes du dr Frankenstein (ou pourquoi faut pas pousser la créature algorithmique dans les datas). *Affordance.info* [en ligne]. 14 septembre 2018. [Consulté le 3 novembre 2018]. Disponible à l'adresse : [https://affordance.typepad.com/mon\\_weblog/2018/09/algorithmes-frankenstein.html](https://affordance.typepad.com/mon_weblog/2018/09/algorithmes-frankenstein.html)

### **Modification de la ville par les algorithmes (financiers)**

Il existe d'autres algorithmes qui surveillent les petites transactions pour les analyser, prévoir les évolutions du marché et comprendre ce qu'ils masquent. 70 % des échanges à Wall Street passent par des algorithmes qui tentent soit de les rendre invisibles soit de les repérer... (...) Dans le monde financier, la vitesse est une arme. Si vous faites des échanges financiers à haute fréquence ou que vous essayez de voir comment ceux-ci se comportent, il vous faut de bons algorithmes, de bons processeurs et surtout un très bon réseau. Et pour gagner quelques millisecondes, il faut être proche de l'internet. Or, tout le monde pense que l'internet est un réseau également et équitablement distribué. Mais ce n'est pas si vrai. A New York, à proximité du Carrier Hotel, il y a un endroit où arrivent les principaux tuyaux de l'internet mondial et qui permet d'obtenir des temps de latence quasiment nuls, explique encore Kevin Slavin. Et cet endroit est justement très proche de Wall Street. Ces quelques millisecondes valent des millions de dollars et les immeubles des environs se remplissent de serveurs plutôt que de gens. Les bâtiments et les villes sont structurellement façonnés par les algorithmes. Dans ce quartier de New York, on vide les bureaux et les appartements pour y mettre des tonnes de serveurs, parce que cette localisation peut donner un avantage stratégique en étant au plus près d'un des cœurs de l'internet. « La ville de New York est en train d'être optimisée pour fonctionner comme une puce électronique. » Et tout ce que nous pourrions faire en tant qu'être humain dans ces mêmes locaux, vaut bien moins que ce que ces machines font et gagnent.

GUILLAUD, Hubert, 2011. Kevin Slavin : « il nous faut dresser l'Atlas des algorithmes contemporains ». *InternetActu* [En ligne]. 24 février 2011. [Consulté le 3 novembre 2018]. Disponible à l'adresse : <http://www.internetactu.net/2011/02/24/kevin-slavin-il-nous-faut-dresser-l-atlas-des-algorithmes-contemporains/>

