

Les *glitches*, ces moments où les algorithmes tremblent

[Axel Meunier](#), [Donato Ricci](#), [Dominique Cardon](#), [Maxime Crépel](#)

Dans [Techniques & Culture 2019/2 \(n° 72\)](#), version courte en ligne

Imaginons un utilisateur qui voudrait comprendre comment sont ordonnés les résultats des recherches sur Google. Il fait une requête sur Google. Le 26 mars 2019, à sa question "how does google page rank work ?" — disons que notre utilisateur est anglophone — Google fournit une liste de liens hypertextes, mais aussi d'autres informations arrangées sur la page de résultats: la première ligne indique qu'il y a environ 281 000 000 de résultats récupérés en 0,59 secondes. La deuxième ligne : "Did you mean: how does google pagerank work" propose une reformulation de la demande de l'utilisateur. Toujours au-dessus des résultats proprement dits, se trouvent deux encadrés éditorialisés par Google. Le premier est intitulé "Why is the Google algorithm so important ?" et présente les premières lignes d'un article du site *How stuff works* ? extrait automatiquement par Google parce qu'il "contient une réponse possible à la question de l'internaute". Le deuxième s'intitule "People also ask" et propose une liste de questions supposées proches de celle de notre utilisateur (comme "How do I increase my Google page rank?") auxquelles il peut également avoir des réponses, etc.

Bien sûr, l'utilisateur va généralement consulter directement les résultats à sa requête sans prêter trop attention à tous ces détails qui racontent un autre versant des calculs que le moteur de recherche de Google effectue : il procède non seulement à un ordonnancement des résultats mais également à des calculs sur la personne qui fait la requête : qui est-elle ? que cherche-t-elle ? A qui ressemble-t-elle ? Comment la satisfaire ? Au lieu de faire face à une procédure computationnelle, l'utilisateur est pris dans un système algorithmique, assemblage complexe d'acteurs mobilisés dans l'activité banale de faire un requête : une entreprise de services numériques, une interface épurée, plusieurs algorithmes qui récupèrent, sélectionnent et classent des informations de manière à présenter des résultats pertinents pour l'utilisateur, un réseau d'autres utilisateurs par rapport auxquels cette pertinence est calculée, une économie publicitaire etc. Les opérations de calcul effectuées par le système algorithmique concernent non seulement la gestion d'une grande quantité d'URLs mais aussi les personnalités et les attentes des utilisateurs humains, dans un couplage dont nous souhaitons trouver une prise pour en réinterroger l'*évidence*.

Vue comme la métaphore d'une recherche empirique qui chercherait à questionner le fonctionnement d'algorithmes du quotidien, la description de cette scène déplace le cadre de l'attention vers l'*expérience d'être calculé*, et mobilise une figure d'*utilisateur-chercheur*, sujet de l'expérience et potentiellement porteur d'un type de connaissance singulier. Mais il lui manque un *objet* à partir duquel une enquête permettrait de déployer le système algorithmique à l'œuvre.

Le 16 décembre 2017, une journaliste du New York Times nommée Rachel Abrams publie un article intitulé “Google Thinks I’m Dead (I know otherwise.)”. Le point de départ de l’article est une capture d’écran du résultat de la requête “Rachel Abrams” sur Google. La journaliste s’est *googlée*, et constate sur la page de résultats la présence d’un *knowledge panel*, notice biographique éditorialisée automatiquement par Google si une requête concerne une personne dotée d’une certaine visibilité publique. Rachel Abrams a annoté à la main cette capture d’écran. Pointant “Rachel Abrams, american writer”, elle ajoute : “could be me...?”, puisqu’elle est journaliste. La photo ne fait aucun doute : “Me”. Les autres détails concernant “Rachel Abrams”, par exemple sur sa famille, sont en revanche déconcertants (“Not me”), en particulier sa date de naissance... et de décès (“Definitely not me”). Dans l’article, la journaliste raconte sa bataille pour rapporter et faire corriger l’erreur, qui pourrait plonger dans la peine des personnes la connaissant et découvrant ces informations marquées du sceau de l’autorité de Google. Cette bataille consiste en une enquête pour qualifier l’erreur et en attribuer la responsabilité, mais aussi en une exigence de renégocier l’identité de Rachel Abrams, ici en la distinguant d’une homonyme dont Google a tiré les détails biographiques de son entrée Wikipedia.

Cette fois, et par contraste avec la première histoire, l’utilisatrice se rend compte qu’elle a été mal calculée. Le dysfonctionnement du calcul rend visible que la notion de *pertinence*, construite ici en agrégeant automatiquement des informations éparses à propos d’un nom et en inventant de toutes pièces une personne (**fig 1**), est problématique. La victime rend compte de son expérience et nous interroge : que devrait faire Google ? Comment traiter les incertitudes liées à l’homonymie par exemple ? Faut-il sacrifier la possibilité d’identifier une personne par égard pour celles avec qui elles pourraient être confondues ? La force de cette histoire est finalement sa banalité, car il est bien sûr des situations beaucoup plus problématiques liées à la confusion d’une personne avec une autre (**fig 2**).

Cette banalité, mise en regard avec la banalité d’être calculé sans s’en rendre compte, est la prise que nous cherchions pour pouvoir aiguïser notre sensibilité aux effets des calculs dans une vie de tous les jours où ils ne brillent pas toujours par leur performance et connaissent des ratés, des erreurs, des glitches, des perturbations si mineures que les utilisateurs de services numériques ajustent souvent leur comportement pour les corriger. Dans ces situations sont convoqués autant des principes éthiques que des limitations techniques ou des variations imprévisibles des contextes d’usages dans lesquels les algorithmes opèrent. Une des spécificité des systèmes algorithmes est de continuer coûte que coûte à fonctionner, de s’obstiner à ne pas tomber en panne, ce qui crée une fracture entre le monde des calculs et le celui de l’utilisateur, qu’il ne faut pas seulement réparer mais aussi explorer et renégocier.

FIGURES

Figure 1 Photo de personne inexistante sur laquelle un glitch est visible. Le site thispersondoesnotexist.com produit automatiquement des photos de visages humains grâce à des réseaux antagonistes génératifs, une technologie de machine learning utilisée pour créer des images photoréalistes à partir d’un ensemble de photographies collectées sur Flickr. (©Tero Karras et Ian Goodfellow).

Figure 2 Capture d'écran du film Brazil (1985). Un bug dans la chaîne d'impression de documents où figure le nom de famille d'un terroriste recherché entraîne le remplacement de la lettre T par la lettre B et mène à l'arrestation d'un innocent.