Communication et organisation

Revue scientifique francophone en Communication organisationnelle

54 | 2018

Pratiques de la communication et *Big Data*

Entretien avec Stefaan Verhulst, GovLab. Collective data, empowerment citoyen, data activisme : des preuves réelles que le *Big Data* fait partie d'une intelligence collective

Propos recueillis par Nayra Vacaflor, à New York, octobre 2018.

Traduction et adaptation : Anne-Marie Cotton, Nayra Vacaflor

Nayra Vacaflor

Édition électronique

URL : <https://journals.openedition.org/communicationorganisation/7308>

DOI : 10.4000/communicationorganisation.7308

ISBN : 979-10-300-0341-3

ISSN : 1775-3546

Éditeur

Presses universitaires de Bordeaux

Édition imprimée

Date de publication : 1 décembre 2018

Pagination : 243-250

ISBN : 979-10-300-0340-6

ISSN : 1168-5549

Référence électronique

Nayra Vacaflor, « Entretien avec Stefaan Verhulst, GovLab. Collective data, empowerment citoyen, data activisme : des preuves réelles que le *Big Data* fait partie d'une intelligence collective », *Communication et organisation* [En ligne], 54 | 2018, mis en ligne le 01 décembre 2021, consulté le 04 janvier 2023.

URL : <http://journals.openedition.org/communicationorganisation/7308> ; DOI : <https://doi.org/10.4000/communicationorganisation.7308>

Tous droits réservés

ENTRETIEN

Entretien avec Stefaan Verhulst, GovLab. Collective *data*, *empowerment* citoyen, *data* activisme : des preuves réelles que le *Big Data* fait partie d'une intelligence collective

Propos recueillis par Nayra Vacaflor, à New York, octobre 2018

Traduction et adaptation : Anne-Marie Cotton, Nayra Vacaflor

Stefaan Verhulst est co-fondateur et directeur recherche et développement de GovLab, laboratoire gouvernemental et centre de recherche action basé à l'Université de New York. Avant de rejoindre la NYU à plein temps, Verhulst a passé plus de 10 ans en tant que chef de recherche à la fondation Markle, où il continue d'être conseiller principal. En outre, il est professeur auxiliaire au département de la culture et des communications de l'Université de New York, chargé de recherche au Centre d'études sur les médias et la communication de l'Université d'Europe centrale à Budapest. Chercheur affilié au Centre for Global Communications Studies de la Annenberg School for Communication de l'Université de Pennsylvanie. Il est Senior Fellow affilié à PCMLP - Université d'Oxford qu'il a cofondé en 1996. Il siège également au conseil d'administration d'Orbicom, le réseau des chaires UNESCO.

Quels sont les objectifs de GovLab et quel rôle y jouent les *Big Data* ?

L'objectif de GovLab est d'améliorer la vie des gens en modifiant la façon dont nous gouvernons, en utilisant les nouvelles technologies. Nous nous concentrons particulièrement sur l'apport des nouvelles technologies en ce qui concerne les décisions prises dans l'intérêt public, la conception des services publics et la résolution des problèmes publics. Nous nous rendons compte de la portée ambitieuse de cet objectif. C'est pourquoi nous nous sommes concentrés sur deux domaines importants pour changer notre façon de prendre des décisions, de résoudre les

problèmes du public, de changer notre façon de gouverner et d'améliorer la vie de la population, à savoir d'une part les *datas* et l'intelligence artificielle, et d'autre part, les personnes connectées, ou l'intelligence collective.

Une grande partie de notre travail consiste à déterminer comment exploiter les *datas* pour améliorer la gouvernance et comment tirer parti des compétences des personnes et de leur expertise pour améliorer la nouvelle gouvernance.

Je me concentrerai plus particulièrement lors de cet entretien sur nos d'activités concernant la *data*, le travail que nous effectuons en faveur des citoyens et sur l'intelligence collective.

En ce qui concerne les *datas*, nous étudions comment les déverrouiller et les exploiter de manière à améliorer la prise de décision et l'attitude responsable.

Nous avons deux grands domaines d'intervention: l'un d'eux est l'ouverture des *datas* gouvernementales communément appelés «*Open Data*». Nous collaborons ici avec les gouvernements du monde entier pour les ouvrir, de sorte que les ressources générées par les gouvernements, elles-mêmes souvent financées par des personnes ou des fonds publics, puissent être libres d'accès auprès du grand public.

L'*Open Data* peut servir à améliorer la politique gouvernementale, à autonomiser les individus, à augmenter les opportunités économiques, ou encore à résoudre certains problèmes publics. Quatre domaines sur lesquels elles peuvent avoir un impact.

Nous avons également analysé et effectué de nombreuses recherches sur la manière dont l'*Open Data* peut-être utilisée pour changer les prises de décisions, mais également pour concevoir autrement les services publics en termes de mobilité, de logement ou d'éducation.

En ce qui concerne l'autonomisation des individus, nous nous sommes beaucoup concentrés sur la manière dont les données peuvent améliorer les choix que font les gens. Par exemple, comment choisir l'école où envoyer leurs enfants ou encore l'hôpital où aura lieu une intervention sur la base du classement de l'école ou de l'hôpital.

Que développe GovLab pour favoriser l'autonomisation des individus ?

Nous permettons aux individus d'être plus intelligents en réagissant à leur situation par le biais de l'activisme par exemple.

Qu'entendez-vous par plus intelligent ?

En divulguant des informations, les personnes peuvent être informées des performances des écoles dans leur région, ville ou localité et ont ainsi une meilleure idée de la manière dont leurs droits ou leur situation sont indûment assimilés. Cela peut les conduire à un activisme plus performant ou à de meilleures formulations de demandes d'amélioration de leur situation. C'est ce que nous appelons « *data activism* », lorsque les données peuvent être utilisées par des personnes pour plaider et exiger un changement. Nous voyons ce nouveau mouvement social issu de nouvelles mobilisations de données.

Nous touchons ici un sujet capital : le pouvoir politique que les citoyens peuvent avoir avec « *l'Open Data* ». Lorsque vous parlez d'activisme, référez-vous à un changement par la force ? *L'Open Data* est-elle une sorte de force ?

Il ne faut pas utiliser la force. Je vais vous donner un exemple : pendant de nombreuses années, dans différentes régions des États-Unis, il y a eu une discrimination croissante entre les communautés noire et blanche. Et la discrimination a été particulièrement ressentie au niveau de la fourniture de biens et de services. Ainsi, dans certaines villes, tous n'avaient pas accès à l'eau qui était clairement séparée pour cause de discrimination raciale. Mais il n'y avait pas de données pour prouver que c'était le cas. Maintenant, en considérant les données relatives à la fourniture d'eau, nous avons accès à la preuve que l'approvisionnement de l'eau s'arrête à la limite de la ville là où se trouve la communauté noire. Elle a maintenant la preuve qui lui permet de former de meilleurs activistes pour changer cette discrimination.

Le déverrouillage de ces *datas* peut créer un mouvement social autour des données : pareil pour les données relatives au logement, aux schémas de mobilité, etc.

Les données qui se transforment en un outil permettant aux individus d'exiger le changement : comment pensez-vous que cela contribue au processus de communication des citoyens ? Les données transforment-elles ces citoyens qui transmettent des messages plus forts, en une sorte des nouveaux professionnels de la communication ou de nouveaux types de journalistes ?

Absolument. Un des aspects de *l'Open Data* concerne la manière dont on informe les individus sur les décisions, et elle passe souvent par la médiation des données. Il y a l'urgence et le partage dans la médiation

de données : données narratives ou données visuelles. Le format quel qu'il soit a un impact sur le journalisme traditionnel, y compris sur le journalisme d'investigation. Les *datas* n'ont jamais un impact en soi. C'est leur traitement qui est aussi impactant : comment nous en extrayons les idées, comment nous les partageons, avec qui et dans quel but.

Nous parlons de groupes de personnes qui pourraient se réapproprier l'information, déclencher un processus de démocratisation. Pensez-vous que cela pourrait être une nouvelle stratégie de revendication sociale, ou cela pourrait-il aussi être utilisé à des fins moins honorables ?

Chaque pièce a deux faces. C'est également le cas des données ouvertes. Les données vous permettent de voir le monde différemment. Traditionnellement, les données étaient verrouillées. Seuls quelques élus pouvaient voir le monde tel qu'il était, en se basant sur ces données. Si vous les déverrouillez, leur plus grand intérêt est de pouvoir traiter les asymétries informationnelles dans les sociétés. Aujourd'hui, les élus sont plus nombreux à être capables de percevoir la réalité de manière plus factuelle, ce qui leur permet d'exiger certains types de changements.

Mais la question n'est-elle pas de savoir si les citoyens peuvent réellement «lire» et comprendre ces données ? Comment les gouvernements communiquent-ils ces données ? Comme ils l'ont toujours fait ou est-ce popularisé ?

Oui. Mais c'est là que la médiation commence. Les journalistes, les médias... c'est leur rôle de faciliter les choses, de consulter les « tableaux de bord » des villes pour voir comment sont utilisés leurs revenus, ou quelle part du budget de la ville est utilisée à quelles fins dans la région où ils habitent. Avant, vous n'y pensiez même pas, car ces données n'avaient jamais été partagées. La manière dont cet élément d'habilitation individuelle est traité déterminera son effet car, comme toute médiation, il peut polariser la société. Bien souvent, les données n'ont de valeur que si elles sont mises en contexte. Et le plus grand défi de la publication des données est que vous n'avez pas le contexte dans lequel elles ont été collectées et qu'elles ont été analysées automatiquement. C'est là que réside le danger de l'utilisation de ce type d'information.

Le flux est constant et comporte de grands risques : celui qui en est propriétaire est puissant car il peut en faire ce qu'il veut, même de fausses informations...

Plutôt que de parler de fausses nouvelles, je préférerais parler de données fabriquées. Ne connaissant pas l'objectif initial de la recherche qui a mené aux résultats, il se pourrait simplement qu'ils aient des données, en ignorent une partie et ne disposent pas de l'équation réelle pour les croiser. Beaucoup de gens croient qu'avec des statistiques et des données, tout le monde peut argumenter et c'est ce que les lobbyistes ont fait à l'époque.

La clé d'une bonne lecture des *Big Data* réside donc dans leur analyse. Pensez-vous qu'il existe une articulation entre les données ouvertes et les médias en général ? Les données ouvertes créent-elles un nouveau type de média à côté des médias payés-détenus-acquis ?

Si nous parlons de nouveaux médias, les données peuvent générer quatre types d'informations.

Le premier niveau répond à l'analyse situationnelle où les données peuvent fournir un meilleur aperçu de la réalité. Du point de vue journalistique au moins, on entend la réalité médiatisée au sens de la description de la réalité. Les données peuvent devenir un nouveau facteur de médiation qui montrent la réalité, car nous avons essentiellement des chiffres. La relation peut être positive lorsque les chiffres s'ajoutent à la narration, au *storytelling*. Les chiffres fournissent une certaine représentation scientifique (en fonction des données) de la réalité.

Le deuxième niveau qui suit la description de la réalité, est l'explication de la réalité. Ainsi, nous pouvons utiliser les données croisées et les corrélations pour dire « nous sommes dans cette situation et c'est pour cette raison ». Les données peuvent débloquent des liens de causalité et indiquer où se trouvent les causes de problèmes fondamentaux tels que : la migration, les problèmes de santé publique, l'économie et l'(in)égalité. Nous pouvons avoir une corrélation ou une causalité entre le phénomène et la source. Ceci est extrêmement important pour les médias. Les médias ont tendance à décrire la réalité plutôt qu'à l'expliquer. Mais les données peuvent permettre aux médias de devenir capables non seulement de décrire le problème, mais également d'expliquer pourquoi nous avons ces problèmes.

Le troisième niveau est celui des prévisions. Les prévisions sont basées sur ce que nous savons de ce qui s'est passé, de ce que nous présumons et avons prédit de ce qui se passera dans le futur. C'est un facteur de médiation important car, en tant qu'êtres humains, nous évaluons assez souvent la réalité en tenant compte de l'avenir, à la lumière de ce qui va se passer ou de ce qui pourrait en sortir.

Enfin, le quatrième point concernant les données, en simplifiant, concerne l'évaluation de l'impact. Par exemple, nous avons un ouragan, quel en a été l'impact ; nous avons changé de politique en France, quel en est l'impact ; et, là encore, nous aurons différentes fonctions de médiation. C'est l'impact sur les médias, les nouveaux médias et le journalisme d'investigation.

Pouvons-nous nous concentrer sur le dernier point ? Vous avez parlé d'impact et celui-ci décidera peut-être des choix stratégiques non seulement pour les médias, mais pour toute organisation. Si vous introduisez ces niveaux dans une organisation publique ou privée, comment pensez-vous que les décisions stratégiques vont changer ?

Je pense que tous les niveaux que j'ai mentionnés sont liés même s'ils sont distincts. Pour réaliser le quatrième, vous avez besoin du premier. Pour avoir un impact ou des évaluations, vous devez disposer d'une base de référence de la situation et si vous souhaitez l'analyser, vous devez le prédire. Mais ce que je pense vraiment, c'est que le potentiel réel des *Big Data* consiste à aller au-delà de la description de la réalité. Il s'agit d'expliquer la réalité ou d'expliquer l'impact de cette réalité. Vous avez toujours besoin d'une deuxième vision.

La discussion portera sur l'impact de ce que vous lisez dans les nouvelles. Un produit lancé, une nouvelle loi adoptée... quel est l'impact réel ?

Je pense que tout est traité de la même manière. Il existe un meilleur moyen de valoriser la valeur des développements. Parce que chaque développement n'est pas considéré comme tout aussi important, ni traité comme tout aussi important. Si vous disposez d'un meilleur moyen de comprendre quel est le potentiel et l'impact existant, nous aurons peut-être pour objectif réel de rediriger l'attention sur les éléments importants.

Lorsque vous créez vos données ici à GovLab, vous les analysez également ?

Nous avons des projets en partenariat avec des organisations spécialisées dans les *datas scientists* et, par conséquent, une grande partie de nos activités, outre les données du gouvernement ouvert, table sur ce que nous appelons la « *data collaborative* ». Il s'agit d'une nouvelle manière d'organiser les analyses de données lorsque nous nous impliquons dans le secteur marchand pour créer des référents que le secteur public peut utiliser. Nous analysons des *datas* d'entreprise, telles que les appels ou les enregistrements de recherche d'entreprises de téléphonie ou les flux de satellites, afin de générer des informations sur la mobilité. Telefonica de Santiago du Chili analyse la mobilité dans cette perspective. Les femmes voyagent-elles différemment des hommes et si oui pourquoi et comment ? Il est important de savoir que le transport urbain est axé sur les hommes.

Nous avons un projet en Inde avec Microsoft et le moteur de recherche Bing, qui vise à mieux comprendre les idées qui mènent au suicide au moyen de mots-clés utilisés dans une recherche. Ainsi, en identifiant les mots-clés qui peuvent inscrire l'idée de suicide, nous pouvons travailler sur la prévention.

Lorsque vous parlez de « *data collaborative* », je voudrais savoir s'il est éthique de tout communiquer à propos de votre analyse. Quelle est votre définition de l'éthique lorsque vous informez de vos résultats ?

Les données répondent essentiellement à une question. Ce que nous avons tendance à faire, c'est d'essayer de comprendre qui s'intéresse à cette question. Très souvent, nous développons la question avec ceux qui se soucient de la question. La conception d'un projet *datas* a déjà des implications éthiques.

Une implication éthique est de définir à quelle question souhaitez-vous répondre et qui décide que cette question est celle à laquelle vous répondez avec les données. En ce qui concerne la conception de la recherche, décider de la question à laquelle vous souhaitez répondre à l'aide de données constitue déjà une réflexion éthique. Nous nous engageons donc afin de comprendre la question qui importe et de prévoir les conséquences que peuvent avoir la ou les réponses auprès du ou des publics visés.

Tout au long d'un projet, nous avons des stratégies d'engagement qui s'adaptent aux connaissances que nous collectons. C'est pourquoi nous avons tendance à ne pas publier directement un dossier finalisé, mais à « socialiser » nos résultats afin de récolter les premières réactions qui nous permettront d'identifier ce que nous avons omis ou quelles pourraient être les applications.

Passé ce stade, nous les disséminons à travers des *workshops* et des groupes de réflexion. Une grande partie des *datas* que nous recueillons ont de la valeur pour ceux qui peuvent agir sur ces données. Nous ne ciblons jamais le grand public, mais nous nous concentrons sur ceux qui peuvent agir d'une manière ou d'une autre sur des données.

Les *datas* ne sont pas une chose. Ce n'est pas quelque chose qui émerge. Les données sont un processus. En travaillant avec les données, nous devons comprendre comment elles sont collectées et quelles sont les implications éthiques de cette collecte. Nous avons donc une éthique de l'inclusion. Nous avons une composante éthique du processus, en analysant et en utilisant les données...

Tout cela n'est pas axé sur l'éthique de l'exploitation des données. Il y a aussi une implication éthique de ne pas utiliser certaines données. Il y a beaucoup de points de vue en ce qui concerne l'abus ou le non-usage des données... Notre société a tendance à mettre l'accent sur l'éthique de l'utilisation mais je pense qu'il est tout aussi important de se concentrer sur l'éthique de la non utilisation des données.