

Personne interviewée : Geoffrey Delcroix
Interviewers : Paul Fournier et Louis Zigrand
Date de l'interview : Jeudi 01 Juin 2017, entre 15h et 16h
Lieu de l'interview : CNIL, PARIS

►(00:40) Pour commencer, est-ce que vous pouvez nous dire en quelques mots, préciser votre travail ?

Bien sûr. Donc, en ce qui concerne. Donc, donc je m'appelle Geoffrey Delcroix. Je suis chef de projet innovation et prospective, ici à la CNIL. Peut-être d'abord deux mots sur la CNIL. Alors vraiment très rapide si vous avez des questions, je peux préciser. Mais juste en deux mots, donc on est une autorité administrative indépendante ; La CNIL en fait, c'est une commission de 17, en fait 18 avec le président de la CADA, mais commissaires qui sont des hautes personnalités désignées par la cour de cassation, le conseil d'Etat, le gouvernement, le président de l'Assemblée Nationale. Donc évidemment ce sont des personnalités qualifiées, des magistrats qui sont notre organe, qui a la légitimité par la loi, donc c'est notre organe qui, entre guillemets de décision. Derrière il y a un président de la CNIL, en ce moment c'est une présidente qui s'appelle Isabelle FALQUE-PIERROTIN. Et donc des services, on est 200 agents ici. Sur un certain nombre de missions, généralement les gens nous connaissent, éventuellement par deux aspects, ont l'idée de déclarations ou de trucs du genre, et peuvent avoir une idée de nos missions de sanctions éventuellement, le côté gendarmes des données, comme les journalistes aiment bien le dire. Donc c'est effectivement deux missions qui existent mais en fait notre mission est plus large que ça. Donc c'est, heu... la CNIL à été créée par la loi informatique et liberté en 1978. Elle est chargée de l'application de cette loi qui donc protège les données personnelles et la vie privée des individus. Et le cadre dans lequel on agit est très européen car il y a eu une directive en 95 et un règlement européen à la protection des données. Qui s'appelle le RGPD, le règlement général à la protection des données, qui a été adopté définitivement en 2016 et qui sera applicable l'année prochaine en 2018 donc pile 1 ans, en fin mai 2018. Et donc qui est vraiment une loi européenne sur la protection des données. Et donc dans chaque pays européen il y a un équivalent de la CNIL avec des pouvoirs qui vont se rapprocher avec les différentes autorités. Dans la salle juste derrière il y a justement aujourd'hui un séminaire du groupe des CNILs européennes qui s'appelle le G29, autour des questions de télécommunication. Ça c'est vraiment. Donc les missions de la CNIL sont essentiellement de deux ordres. Donc on a une mission d'accompagnement, de conseil, à la fois du grand public pour le côté éducation du numérique, explications des enjeux, tutoriels, informations sur comment faire pour sécuriser les données, utiliser des mots de passe. On a fait, il n'y a pas longtemps pas mal de recommandations autour de ce du sujet des mots de passe sûrs. Et aussi accompagner les responsables de traitement. Ça c'est le vocable de la CNIL pour parler de toutes les entreprises ou administrations qui vont collecter ou traiter des données à caractère personnel. Donc on les accompagne beaucoup en fait. On ne fait pas que les embêter, on est aussi là pour les accompagner. Et donc c'est de plus en plus important car le règlement met ça en avant. Et l'autre fonction, c'est effectivement une fonction plus de régulation traditionnelle. C'est-à-dire, contrôle sanction lorsqu'il y a des manquements à la loi. Donc en fait on a, on reçoit beaucoup de plaintes, il faut savoir, de citoyens. Ça peut être des citoyens qui ont des problèmes avec des entreprises ou des employés. Il y a des caméras dans les bureaux. Donc c'est très, très large, très varié. On reçoit plus de 8000 plaintes par an. Je ne suis même pas sûr que ce sont les derniers chiffres que je vous donne, enfin plusieurs milliers de plaintes par an. On fait des contrôles sur pièces et sur place, ou en ligne auprès d'entreprises. On peut, on a des gens dans la maison qui passent leur temps à aller vérifier, c'est un système d'information. Ils peuvent prendre des copies, ce qui est les contenus. Voilà.

Les logs de serveur les choses comme ça. Et quand on relève des manquements on fait en sorte que ça soit réglé. Si c'est réglé on se quitte bons amis. Et si ça ne se passe pas bien et bien on a des pouvoirs de sanction, qui sont, qui étaient relativement limités jusqu'à maintenant du point de vu de la sanction pécuniaire. Parce que l'on pouvait aller au maximum à 150 000€ et 300 000€ en cas de récidive. Ce qui est très peu évidemment car quand vous donnez une sanction de 150 000€ à Google, ou à Facebook, déjà pour eux ce n'est pas beaucoup, puis ça veut dire que quand vous après vous sanctionnez des entreprises plus petites, jusqu'aux PME, vous adaptez. Donc toutes les sanctions pécuniaires sont trop petites en gros. En revanche la sanction de communication est douloureuse car on publie nos sanctions, et généralement ça, ça écorne on va dire l'image des entreprises en question. Et surtout dans le cadre du règlement européen ça changera car le pouvoir de sanction pourra atteindre, au niveau européen, pas pays par pays heureusement quand même faut pas exagérer, mais jusqu'à éventuellement 4% du chiffre d'affaire mondial consolidé. Je vous laisserais faire le calcul si ça vous intéresse mais c'est, ça commence à faire beaucoup plus d'argent. Donc ça c'est sur les missions globalement de la maison, la maison est avant tout une maison de juristes. Le cœur du métier c'est des juristes. Maintenant le, le périmètre dans lequel on travaille étant donné, comme c'est celui du numérique au sens large, et des données, on a de plus en plus une compétence technique. Donc nous, moi je suis dans une direction qui s'appelle la Direction de Technologies et de l'Innovations. Dans laquelle en fait il y a 3 services. Il y a le service de l'informatique interne. Et ensuite il y a un service de l'expertise technologique, qui est un service dans lequel les gens ont en fait un fort background technique, donc ingénieurs, docteurs en sciences informatiques, il y a une docteure en crypto, il y a des spécialistes de sécurité en système d'information. Donc voilà. Et puis l'équipe à laquelle j'appartiens qui est l'équipe du Pôle Innovation Etudes et Prospective, où là la logique est de compléter. Beaucoup de juristes, un regard très juridique, un regard technique qui est de plus en plus présent, et nous on est un peu entre guillemet le reste. C'est-à-dire le regard, par les usages, par les sciences humaines et sociales, par les modèles économiques, par les aspects non concurrentiels, le design, fin tout ce qui peut être le contexte très important dans le domaine et qui sont ni purement technique, ni purement juridique. Et en fait notre mission dans ce pôle, en gros très résumé là aussi si ça vous intéresse je peux décliner un petit peu mais en vraiment résumé, on a 3000, le nom du pôle est pas très joli, Innovations Etudes et Prospective c'est un peu lourd. Mais en fait il décrit bien ce que l'on essaye de faire. La partie prospective ça veut dire que l'on regarde, ça ne veut pas dire que l'on regarde forcément à 50 ans dans le futur car pour nos sujets c'est un peu difficile, l'objectif c'est de prendre des sujets qui sont assez émergents, qui ne rentre pas dans la maison par les voix traditionnelle, c'est-à-dire pas des sujets où il y a encore des plaintes, pas des sujets où on a des dossiers très bien ficelés matures. C'est des dossiers assez émergents. On va essayer de comprendre comment ça fonctionne, de réfléchir aux enjeux et puis donc de, de préparer, on va dire d'anticiper pour que, pour que lorsque les choses deviennent un peu urgentes on soit en mesure de réagir. Le côté étude c'est parce que, c'est ce que j'ai déjà dit, c'est parce que on regarde le sujet avec vraiment une vision kaléidoscopique, on ne rentre pas par le toit. Si il faut que l'on parte de, la manière de... Oui, on part souvent des usages en fait. C'est vraiment le point de départ souvent de nos analyses. Et le dernier élément, le côté innovation ça marche dans les deux sens en fait c'est, on essaye d'être le point de contact des systèmes d'innovation. Donc on va dire, ça c'était la manière chique de le dire, la manière un peu plus directe c'est de dire : nous notre rôle c'est d'aller discuter avec des gens qui ne savent pas qu'ils ont un intérêt à discuter avec nous, ou qui n'ont pas envie même, pour être encore plus... . C'est-à-dire parler avec des startups, parler avec des chercheurs, parler avec des équipes d'entreprise ou d'innovations. On va dire que quand les gens sont déjà en train de se dire, on a vraiment besoin de la CNIL, c'est pas nous qu'ils viennent voir. Ils ont une question précise et tout. Nous on est là plutôt pour accroître la compréhension, monter des enjeux. Voilà. Et puis ça marche dans l'autre sens aussi. On est une équipe d'innovation en interne. C'est-à-dire que l'on essaye de

travailler avec des méthodes un peu innovantes, de changer la manière dont la CNIL produit de la connaissance. On a des publications, papier, qui sont des cahiers d'innovation et de prospective, où on aborde de manière assez transversale un sujet, en ayant la voix qui n'est pas la voix du régulateur qui dit « ça c'est légal, ça ce n'est pas légal », mais qui essaye de poser les enjeux éthiques et de créer, de catalyser du débat. On a un site à l'intérieur du site de la CNIL qui s'appelle LINC.fr. Donc c'était pour faire Laboratoire d'Innovations Numérique de la CNIL. Ça faisait un acronyme récursif donc ça nous plaisait bien. Et puis c'est CNIL à, l'envers donc ça nous plaisait bien aussi. Et sur ce site on produit de la veille, des articles un peu de réflexion, et on a un travail d'exploration, de laboratoire avec nos collègues de l'expertise, qui eux sont plutôt ceux qui, sont capable de développer, monter les projets mais nous on va les, aller piloter avec eux dans une logique vraiment de prototypage ou de R&D à notre niveau. Voilà un peu nos , nos différents actes d'activité dans l'équipe.

► **(10:10) Pour entrer concrètement dans le sujet, si on vous parle de bulle filtrante ou de “Filter Bubble”, comment vous le définiriez ?**

Alors, bon c'est vrai que l'on a pas mal, on a un peu travaillé sur le sujet. On avait, on avait, on avait repéré assez tôt les travaux d'Elis Pariser sur le sujet. Alors si je, j'essaye de le définir un peu à ma manière, comme ça spontanément, je dirais : c'est la capacité que peuvent avoir outils et des services numériques à accroître l'enfermement d'une personne, on va dire surtout de manière peu visible, fin de manière, de manière peu sensible pour la personne en elle-même. C'est-à-dire que c'est, c'est, **les plus grosses bulles de filtre c'est nous qui les créons**. Fin je veux dire on travaille beaucoup avec Dominique Cardon qui a pas mal écrit là-dessus, qui fait que c'est pas nouveau, c'est pas Facebook qui à créer le fait que les gens de, si on prend la politique, que les gens de droite ont tendance à discuter plus souvent et à trouver des informations plutôt de droite. Voilà c'est pas nouveau. La différence là c'est que l'on est vraiment face à des bulles dans le sens où, c'est créé par des outils éventuellement algorithmique, ou essentiellement algorithmique d'ailleurs, qui ne sont pas très compréhensibles, visible par l'individu. Et qui donc rendent cette bulle transparente. Alors du coup une bulle c'est transparent et justement ça permet de voir au travers donc c'est pas le bon, c'est pas ce sens-là de transparence. C'est plutôt transparent dans le sens on ne le voit pas, on ne se rend pas compte qu'il y a une paroi et que il y a des choses qui sont déformées.

► **(11:56) D'après votre expérience et vos connaissances, quels seraient les signes visibles des “Filter Bubbles” dans l'utilisation d'internet au quotidien ?**

Les effets les plus visibles ils sont autour de l'accès à l'information par les réseaux sociaux, ou par les moteurs de recherche et les services de vidéo comme YouTube. Enfin c'est ces choses-là qui l'incarnent le plus. Mais moi j'ai, nous on a tendance, étant données nos missions ici, nous on a tendance à regarder ça en s'intéressant au fait que les services numériques s'appuient sur les données personnelles collectées, de manière d'ailleurs là aussi très peu visible, pour personnaliser et améliorer en termes de ces gens-là, améliorer l'expérience d'utilisation en personnalisant les choses et en utilisant des outils très puissant algorithmique en particulier et avec en fait un effet qui est... Nous on était très embêté par une expression qui était à un moment. On l'entend peut-être un petit peu moins, peut-être parce que maintenant tout le monde parle d'intelligence artificielle, ça fait encore plus chic. Mais c'était de parler de “magie des algorithmes”. Moi j'entendais beaucoup dans la presse, j'ai un exemple en tête en fait, c'est le patron de, c'est pas pour stigmatiser particulièrement l'entreprise mais je me souviens qu'il avait répondu ça. C'était le patron de Deezer qui était interrogé par un journaliste sur les articles de recommandations de Deezer et pour plein de bonnes raisons il n'avait pas envie, c'est pas très média, c'est pas très presse, enfin il allait pas se mettre à rentrer, à parler

de des algorithmes, donc il a dit : « ça c'est la magie des algorithmes ». Je trouve cette expression elle est hyper dangereuse car ce n'est pas magique un algorithme. C'est pas de la magie derrière. Il y a des choses qui sont finalement fort concrètes. Et on a aujourd'hui en fait un phénomène qui est que l'on veut tellement simplifier l'expérience d'utilisation qu'il faut faire disparaître tous les moments où l'utilisateur est en contact avec la quincallerie, le BackOffice du service. Et donc on infantilise un peu les utilisateurs en leur donnant l'impression que les choses sont toujours neutres, simples, évidentes. Voilà. Alors que, en fait, elles ne le sont pas, il y a des choix, il y a . Moi j'ai l'impression que c'est ça, au-delà du sujet que tout le monde aborde de la question de, est-ce que Facebook ou Google influence sur nos opinions et tout ça, plus globalement on a une espèce d'euphémisation de ces outils, on donne l'impression que parce que ce sont des outils techniques ils sont neutres. Et donc on a pas besoin d'expliquer, on a pas besoin d'expliquer le modèle économique, on a pas besoin de ça. Et donc moi je considère que c'est ça qui incarne le plus, en lien avec nos questions à nous qui sont liées aux droits des personnes, aux libertés individuelles et à la protection des données, c'est vraiment cette question de d'euphémisation donnant l'impression que voilà c'est. Je ne sais plus qui avait dit : « L'innovation n'est ni bonne, ni mauvaise, ni neutre (*insiste sur ni neutre*) ». Mais il ne faut vraiment pas oublier le dernier bout quoi, parce que ce n'est jamais neutre de mettre en place un outil technique de tri d'information, de personnalisation, de recommandation, de priorisation, peu importe après il y a plein de, plein de choses.

► (15:41) Si on parle des algorithmes, tels que le PageRank de Google ou les recommandations YouTube, quelles parties de cette personnalisation du Web pensez-vous qui ont été sciemment programmées à la base ?

Alors je pense que c'est une question importante celle de l'intention des créateurs de ces outils. Moi, je, l'important c'est de ne pas oublier les modèles économiques qui sont derrière. C'est pas grave, enfin je veux dire c'est pas sale, c'est pas des méchants. Voilà. On peut pas en vouloir à Google, à Facebook, de, de... il y a une, il est normal que les entreprises alignent leur choix avec leurs intérêts économiques, c'est pas scandaleux. C'est pas à Facebook de garantir la neutralité du débat public. C'est bien si ils ont envie de participer à ça, mais ça n'a de la valeur que ce que l'on, enfin voilà c'est. Il n'y a rien qu'il les y oblige en tant que tel, sauf si on les, les y oblige voilà par la loi ou si on crée des outils mais c'est. On peut pas juste se contenter de compter sur leur bonne volonté parce que voilà c'est pas obligatoire. C'est une entreprise son objectif c'est quand même, en théorie la maximisation du profit. Si à côté de ça ils arrivent à faire du bien commun et à avoir des valeurs c'est très bien, mais ça ne remplace pas la maximisation du profit. Donc moi je pense que... il y a toujours une intention, ou qu'il y a toujours un moment où lorsque l'on va améliorer, accroître les performances, évaluer les performances de ces outils, l'aspect, le lien avec ces les modèles économiques va être au moins aussi important que le côté neutre, rigoureux, scientifique et pas généralement beaucoup plus. Et comme les domaines du numérique sont quand même beaucoup basés sur la gratuité, donc sur les marchés que l'on appelle les marchés bifaces, avec un utilisateur et un client qui ne sont pas la même personne. Un client payant. Un consommateur et un client payant qui ne sont pas la même personne. Ben on a des effets de déformation qui peuvent quand même être liés à ça. Je pense, on le voit bien avec YouTube le fait que l'algorithme de YouTube à quand même intérêt à vous maintenir sur la plateforme le plus longtemps possible, c'est comme l'intérêt de Facebook, vraiment que vous y passiez le plus de temps possible. Parce que le temps est un bon proxy de votre valeur sur le marché publicitaire. C'est le patron de TF1 qui avait dit « le temps de cerveau disponible ». C'est moins brutal quand on fait de la personnalisation, de la publicité ciblée. Justement l'idée c'est que l'on est censé pouvoir en mettre moins. Mais on est là quand même pour vendre ça, du temps d'accès à l'attention de la personne. Donc, oui, il y a un effet, un effet sciemment pour répondre à votre question, lié à un, à des modèles

économiques. Des modèles économiques qui sont de plus en plus variés quand même, mais des modèles économiques qui sont tout de même centrés sur la gratuité de l'accès en échange de la valorisation publicitaire qui est concurrentiellement intéressante pour s'enrichir.

►(19:26) De votre point, quelles peuvent être les influences que peuvent avoir les “Filter Bubbles” sur les individus et la société ?

Alors, je pense que bon, d'abord il y a des vraies questions dans l'accès à l'information et les risque effectivement de manipulation des opinions qui peuvent exister mais à la limite moi je suis peut-être moins spécialiste là-dessus enfin c'est évident. Après, moi je pense qu'il y a un vrai risque de forme d'infantilisation de l'utilisateur par rapport à ces services. Sur le côté expliquer quelque chose ça ennuie tout le monde, comment ça fonctionne. En plus faut connaître que c'est pas simple essayer de résumer l'algorithme du fils de Facebook ça demande du travail pour l'expliquer, c'est pas d'une simplicité complète comment dire. Mais je crois que c'est vraiment très dangereux que de. On risque de produire chez les gens un très faible esprit critique en vue des outils techniques en eux même avec cette impression que ils sont ambiant, voilà, ils ont pas de. L'expression de neutralité est très dangereuse voilà, un algorithme ne peut pas être neutre. Ne serait-ce que parce que on ne peut pas coder la neutralité, c'est des critères et ces critères ils sont pas aussi objectifs, ils sont le plus objectifs possibles, mais ils ne le seront jamais complètement. On le voit bien, je ne sais plus qui a écrit cet article, ah si c'est avec Dana Boild, qui a écrit un article en disant que les intelligences artificielles avaient un problème de ... d'homme blanc. Je ne sais plus quel est le titre de l'article, ça doit être le “White men problem”. Ce qui est de dire qu'aujourd'hui tous les gens qui font des travaux sur les intelligences artificielles et même toutes les données sur lesquelles ont les entraînent sont extrêmement concentrés sur des homme blanc d'un certain âge. Et il y a des effets fort de ce point de vu là. Je crois que c'est elle dans son article qui donnait un exemple assez génial. C'est de taper dans Google des mots très générique, comme ‘homme’, ‘femme’ et de voir ce qui sort en premier. Et je crois que c'était ça, elle disait « allez taper ‘man’ ou ‘men’ dans Google Image et les premières images c'est des hommes blanc américains entre 30 et 50 ans. Et la moitié c'est Donald Trump en plus en ce moment. C'est encore plus flippant, il est encore plus âgé. Mais en fait, c'est vraiment, il y a un effet lié aux concepteurs qui est très, très fort et qui est difficile à venir casser. De non, de non, de faible diversité donc ça renvoie à un deuxième danger, je pense, qui est le fait de recouvrir d'un voile technique des discriminations et des problèmes qui existent et donner l'impression que en fait quelque part il y a une forme de blanchiment, comme on blanchit de l'argent, de blanchiment de discrimination à travers ces outils. L'exemple qui avait été donné, c'est une chercheuse qui s'appelle Latania Souiné, qui avait travaillé sur les algorithmes de résultat suggérés de Google et qui avait montré que si on tapait dans Google des prénom typés afro-américain on tombait très souvent sur des suggestions sur le fait de faire des recherches sur le casier judiciaire. Je crois que c'était sur l'auto-complétion mais ça il faudrait regarder sur l'article de. Voilà si vous tapiez un prénom qui est plutôt associé à la communauté afro-américaine vous allez plus facilement tomber sur “recherche casier judiciaire”. Et quand vous tapiez des prénoms plutôt associés aux blancs c'étaient d'autres suggestions qui arrivées. Et en fait ce que Google avait répondu, ce qui a était aussi leur argument dans pas mal d'autre cas, c'était de dire qu'eux ils montraient juste, fin c'était ce que les gens faisaient dans la majorité des cas. Une des raisons qui est triste mais qui est réelle, c'est qu'il y avait énormément de propriétaires ou d'employeurs, avant d'employer un noir ou de lui louer un appartement, qui faisaient cette recherche. Voilà. Et ils le faisaient pour les noirs et pas pour les blancs. Et donc ça crée massivement dans les résultats de Google voilà. Google disait « On a pas créé cette discrimination[pf2] .[pf3] ». Mais ils ne l'ont pas créées mais il la renforce. Fin ça la rend, ça la multiplie. C'est pas de leur responsabilité, en tant que entreprise de lutter contre ce genre de discrimination. Ils ne les ont pas créées et ce

n'est pas à eux de les détruire. Mais si on intègre vraiment, s'il n'y a aucun outil juridique, aucun outil technique mis en place face à ça il y a un vrai effet, oui de blanchiment comme pour l'argent sale. Ça les transforme en quelque chose de complètement acceptable en fait. Ça je pense que c'est un vrai risque de ces outils. Enfin c'est moins directement lié aux bulles de filtre mais c'est quand même lié aux bulles de filtres. C'est des bulles de filtres dans le sens où des discriminations vont être associées à des catégories de types de résultats de recherche de manière beaucoup plus forte. On le voit aussi sur des choses moins graves, on a notre. On publie un de nos cahier régulièrement et le troisième c'était « la place de l'industrie dans les données culturelles ». Donc lecture, musiques en ligne, vidéos en ligne et jeux vidéo. On essaie d'expliquer que là où il n'y avait pas beaucoup de personnalisation dans ces secteurs-là, aujourd'hui il y en a énormément. Et un des sujets était « Est-ce que des outils comme Spotify ou Deezer pour la musique peuvent (ou pour la vidéo mais les catalogues sont notamment compliqués c'est moins facile à voir), peuvent appauvrir la diversité culturelle d'accès à la culture ? ». Et en fait, oui, ça peut être le cas selon comment sont fait les algorithmes et les outils, c'est pas obligatoire non plus et en particulier quelqu'un nous a dit quelque chose d'assez intéressant qui est, c'est entreprises ont souvent pas mal d'outils, pas un algorithme mais souvent plein en concurrence les uns avec les autres, et en particulier l'un des éléments c'est que ces plateformes essaient d'identifier justement l'appétence pour la diversité de l'utilisateur. C'est-à-dire que quelqu'un qui va écouter uniquement les grands hits à la mode ben l'algorithme va se dire « bon ben aller je ne lui sers que ça » ; et il n'a plus vraiment besoin de se dire je vais lui rajouter d'autre type de musique pour voir si. Donc il y a un effet qui peut être ici très fort car ça correspond à ce que l'utilisateur souhaite qui va le faire rester voilà. Et en revanche les outils sont plutôt capable de reconnaître des gens qui sont très diversifiés et en recherche de sérendipité, de nouveautés, de découvertes tout ça, et donc leur servir beaucoup de nouveautés. De nouveautés pas dans le sens ça vient de sortir mais de choses différentes en fait. En fait casser énormément la bulle. Et ça rejoint également ce que disait Dominique Cardon par rapport à Facebook, qui disait « si vous voulez casser votre bulle de filtre sur Facebook, ouvrez-vous à, vous êtes de droite suivez les infos de pas seulement du Figaro mais de Libération et pas seulement et inversement si vous êtes plutôt de gauche suivez la page de valeurs actuelles vous verrez au moins arriver d'autres informations ». C'est peu la même chose c'est que les algorithmes peuvent filtrer dans le sens enfermer les gens, mais peuvent surtout reproduire les inégalités du capital social dans l'accès à la diversité. On le voit aussi, sur les outils de géolocalisation. Ça peut même être sur des applications de rencontre. On sait que la mobilité géographique est très liée à la mobilité sociale. Voilà. Donc quand vous utilisez des outils de ce genre ou entre guillemets quand vous, l'algorithme au départ c'est juste de la géolocalisation. Ben en soit ça introduit déjà des biais et des bulles de filtre parce que vous avez des gens qui vont avoir une mobilité géographique extrêmement faible, donc vont être dans un environnement où ils voient les mêmes personnes. Et donc des gens qui vont avoir plutôt une mobilité géographique forte et qui auront accès à des profils différents. Et on le sait depuis longtemps : mobilité géographique et mobilité sociale sont très liées.

► (29:35) Si l'on voulait prendre l'exemple de médias ou de journaux, qu'est-ce que la personnalisation du Web pourrait avoir comme effet(s) sur une ligne éditoriale ? Ou sur la manière de communiquer en fait ?

(29:55) Je pense qu'il y en a déjà beaucoup alors hum... il faudrait vraiment interroger des journalistes, peut-être que vous en avez mais... eux, ils peuvent expliquer l'impact que ça peut avoir par exemple sur les choix des titres, ou le format des articles, c'est... c'est déjà très fort. Heu... sur les lignes éditoriales, c'est ça votre question de... ? (*grande respiration, puis silence d'une dizaine de secondes*) Heu... (*rire un peu gêné*) pour... ouais, je pense qu'il y a vraiment deux choses différentes. Il y a heu... le rôle que peuvent avoir des intermédiaires des

plateformes heu... dans l'accès au, au... au contenu médiatique, au contenu de presse on va dire pour être plus.... On voit bien que, qu'aux États-Unis, aujourd'hui, 45 % des américains, heu, je ne sais plus exactement le chiffre, mais il y avait une étude qui montrait que 40, lors de la campagne présidentielle, que 45 % des américains qui, qui déclaraient que leur première source d'accès à des articles sur la présidentielle, c'était Facebook. A partir du moment où vous avez des chiffres de ce genre, et ce qui est la réalité, hein, sur les gens, sur les plateformes qui vous apportent des heu... visiteurs, en notant l'importance de ces, de ces visiteurs dans les modèles économiques, vous avez un... une relation de dépendance qui se crée, qui est très forte, à ces outils, à Twitter, à Facebook, heu... ou aux résultats de Google, donc il y a déjà cet effet-là, qui sont en fait que les... les médias sont tentés d'optimiser par rapport à ce que recherchent les... enfin, ce que je veux dire, c'est, c'est... c'est classique, c'est... comme il y avait une sorte d'optimisation pour les sites internet, les journalistes font ça aussi maintenant pour... pas mal pour les articles. Donc ça, c'est déjà on va dire une question, parce que ça a des conséquences directes sur un certain nombre de choix éditoriaux. Hum... après, moi je... on peut le voir dans l'autre sens, c'est que ça devient un sujet en fait heu... enfin que... on a... dans notre comité de la prospective, on a Bruno Patino qui est le... le, le, le... directeur des rédactions de heu... ARTE, je ne sais plus qui est le directeur, je ne sais plus quel est le titre exact, mais bon, c'est un des... un des dirigeants d'ARTE et il était un des dirigeants de France Télévision avant, surtout sur les sujets numériques, qui... il dirige l'école du journalisme de Science Po Paris, qui dit que maintenant, les, les... les algorithmes sont devenus un sujet et un outil de travail par... enfin, c'est devenu un élément du travail du journaliste. Donc ça peut aussi montrer le sujet dans l'autre sens, on voit quand même aujourd'hui pas mal de, de... de médias qui heu... qui utilisent ces, ces... ces outils-là pour heu... enfin, qui, qui... sur les bulles de filtre, les travaux les plus impressionnants qui aient été faits, ont été fait par des journalistes qui ont montré, il y avait je sais plus, c'était le Wall Street Journal ou New York Times qui avait fait le truc « Blue Feed, Red Feed » où ils montraient deux feeds Facebook, un républicain, un démocrate, et heu... donc ça c'est un peu heu... c'est positif quand même, parce qu'on voit des choses, on peut... il y aura peut-être une autre question qui est une question que vous aborderez peut-être plus tard, mais qui est « qu'est-ce qu'on peut faire », mais faut-il de la transparence dans les activités et tout et c'est pas facile mais il y a quand même des, des... des approches comme ça qui peuvent permettre de... de, d'apporter des choses. J'ai pas trop d'autres idées.

►(33:34) **Pour reprendre un peu plus justement dans la sphère publique, en quoi pensez-vous que les « Filter Bubbles » sont sources de controverses et pourquoi vous pensez que ça a resurgi un peu ces dernières années ?**

(33:55) Moi je pense c'est parce que heu... je pense qu'il y a une controverse parce que... il y avait un problème de maturité globale, à la fois du public, des... heu... des acteurs étatiques, gouvernementaux, des médias, et aussi des entreprises de ce secteur-là et, pendant très longtemps, tout le monde a essayé de tenir ce discours, enfin, les entreprises du numérique ont tenu ce discours de neutralité pendant très longtemps, de neutralité même bienveillante, le « don't be evil » de Google, et puis cette idée que, en fait, heu... le numérique étant essentiellement, ce qui est vrai, une mise en place d'une infrastructure et d'un intermédiaire technique heu... l'effet est en fait un effet d'optimisation, un effet de... de simplification d'interface, voilà, mais il y a peu d'impact... donc c'est ce discours sur le fait que... heu... Facebook ne serait pas un média, que Uber n'est pas une entreprise de transport, que RB&B n'est pas un hôtelier ou un acteur du tourisme... qui sont des discours qui sont complètement faux. Heu... ils sont vrais sur le fait qu'ils ne sont pas heu... Uber n'est pas un acteur traditionnel du transport, Facebook n'est pas un média traditionnel, et donc leur appliquer des régulations traditionnelles serait peut-être non seulement inefficace mais probablement absurde

et contre-productif : vouloir appliquer la régulation des taxis à Uber n'est peut-être pas la meilleure des choses à faire. Mais le, le... le discours inverse, qui est de dire que nous on est juste une espèce de surcouché éthérée, intermédiaire et, et, et... et neutre, est très, très dangereuse et je pense que c'est ça qui a créé un peu aussi la controverse, c'est qu'il y a un moment où, en fait, c'est, c'est... la réalité de ce qu'on observe, et des impacts, et des conséquences énormes, étant donné la... la vitesse d'adoption dans les usages de ces services-là, et ce discours de heu... « nous ne sommes quasiment rien, nous ne sommes, nous sommes... nous sommes de tout petits acteurs » c'est, c'est... c'est créer une telle divergence, ça ne pouvait que créer de violentes polémiques et de violentes controverses ; ce qui fait qu'aujourd'hui, je pense... enfin, je suis peut-être un peu optimiste, que les choses sont en train de changer des deux côtés, c'est-à-dire que... on voit peut-être un peu moins des réactions heu... un peu absurdes des, des... acteurs publics ou des, des... responsables en mode heu... « il faut taxer, il faut interdire » et tout ça et plus dans une logique de régulation ; de l'autre côté, un discours qu'est plus de prise de conscience des responsabilités qui, quand même, me paraît plus fort, on voit que... on voit que côté Facebook par exemple, sans dire « on est un média », on voit que les discours ont quand même changé pour reconnaître le fait que, ben, ils ne pouvaient pas juste se... s'extraire complètement de ce genre de choses. C'est pareil pour Google, pareil pour voilà... et les, les, les, les polémiques sont fortes et à mon avis elles vont le rester encore un bon moment sur ce domaine, parce que, parce que ces acteurs sont très puissants et gagnent beaucoup d'argent et heu... et ont acquis beaucoup de maturité aussi. Regardez les dépenses de lobbying d'un compte de ces entreprises au niveau américain ou à Bruxelles, c'est... c'est très impressionnant. Ils sont nombreux hein ? Ça devient très, très impressionnant. Leur puissance de feu en termes de, d'éléments de langage, de politique, de heu... ce n'est plus... c'est plus le, le, les, les geeks avec des sweats à capuche qui font sembler de rester des geeks sous la capuche quoi, maintenant, on voit bien quand même les avocats, les costumes-cravates qui tiennent de beaux discours politiques, enfin, c'est... ça y est, c'est, c'est du sérieux quoi. Ça fait longtemps que ça l'est hein, mais je veux dire, ça, maintenant c'est assumé un peu plus qu'auparavant.

► (38:01) En quoi vous pensez que la société civile pourrait avoir des moyens d'actions envers les bulles filtrantes ?

(38:16) Alors je pense qu'il y a... après, ça dépend de la définition de société civile, déjà, deux choses : du côté des citoyens, des militants et tout ça, hum... ben, je pense que... ça rejoint un peu ce que je disais sur l'exemple des travaux des journalistes... il y a plein de choses qui peuvent être heu... bricolées, testées de manière un peu hum... d'ailleurs ouvertes pour essayer de, de mesurer les choses. C'est très difficile de, en fait, d'avoir les outils pour mesurer ces idées, les bulles filtrantes et tout ça ; mais finalement, si on doit le faire, peut-être qu'il est plus réaliste d'essayer de le faire en partant des individus, en comparant des choses, en comparant des choses un peu « bottom-up », plutôt que d'essayer d'imposer des sortes d'audits heu... très centralisés où, honnêtement, c'est quand même pas d'une simplicité folle à mettre en place... même un équivalent de ce qui existe sur les comptes, les commissaires aux comptes et tout ça, il y a plein de gens qui proposent des commissaires aux algorithmes, des audits et tout ça, il y a des domaines pour lesquels c'est probablement possible, mais c'est quand même pas simple à mettre en place, une espèce de certification de ce genre. Alors qu'en revanche, produire des outils de heu... qui permettent de comparer les entrées et les sorties sur plein de gens différents et de voir du coup s'il y a des choses qui apparaissent, ça semble heu... faisable, il y, il y a un projet intéressant dont vous avez peut-être entendu parler autour de YouTube, hum... il faudrait que je retrouve le nom, il est sur... il essaie de créer de la transparence autour des algorithmes de recommandation en prenant l'exemple de YouTube, où l'idée c'est que, heu, ils ont développé un, un... un petit outil qui est capable de... heu, de

scraper les recommandations en fait, et l'idée est que si plein de gens l'utilisent et puis commencent à noter eux-mêmes un certain nombre de choses, ça peut permettre de créer une base de données partagée, pour essayer de voir quelle, de, de, de, de... faire une sorte de rétro-engineering mais pas basée sur heu... une atteinte à la propriété intellectuelle, plutôt sur une idée de tester des entrées et des sorties et voir si, massivement, ça dit quelque chose. Moi je pense que ça, ça peut être une... ça peut être une manière de faire. Et après, heu... je pense que... et là je prêche un peu pour ma paroisse, mais, en même temps, je pense quand même que, il faut pas négliger le rôle tout simplement du monde des institutions démocratiques et des institutions heu... de la loi, d'un certain nombre de, de, d'obligations en... modernisant, en essayant de trouver des modes modernes pour faire ça, pas forcément des déclarations, formalités et des autorisations, mais des outils, enfin des, des outils de droit qui permettent d'encadrer heu... ces, d'encadrer ces acteurs et de leur imposer à eux aussi l'application du droit, parce que c'est vraiment un sujet qui a été très fort ici, parce que pendant très longtemps, on avait un problème, c'était la question de l'applicabilité territoriale de la loi ; il y avait des acteurs qui nous disaient « moi je suis installé en Californie, la loi française, elle s'applique pas », c'est des choses qui, qui... je pense que c'est un combat qui a duré dix quinze ans, et qu'on est en train de gagner, heu... le, les juges généralement ont plutôt donné raison aux autorités et le règlement dont je parlais tout à l'heure, le règlement général de la protection des données, pour, pour vraiment vous faire la version rapide, sur ce sujet-là, dit que, il s'applique ce règlement pour toutes les entreprises, enfin les responsables de traitements, ça peut être des acteurs publics, qui soit ont un établissement principal quelque part sur le territoire européen, que ce soit en France, en Irlande ou en Espagne, enfin, peu importe, du moment qu'il y a un endroit un peu installé, donc heu... pour toutes ces grandes entreprises là, ils finissent quand même, enfin, ils ne peuvent pas, ils peuvent pas, ils ont, ils ont... des établissements, pas forcément pays par pays, mais justement comme là c'est à l'échelle de l'Europe, ça marche quand même normalement bien, il prend un deuxième critère pour les entreprises qui ne seraient même pas installées physiquement sur le territoire, on considère, un peu comme on le fait dans le droit commercial, en fait, que si vous ciblez spécifiquement et, on va dire, de manière un peu évidente, la population européenne pour lui fournir un service heu... ça veut dire que vous devez appliquer à cette catégorie de vos clients le droit européen. En gros, la question après est, maintenant, il y a des juges et des avocats et des juristes, et les juges vont pouvoir statuer là-dessus dans les temps qui viennent... mais l'idée, c'est que, vous diffusez votre service sur des plateformes de distribution qui sont à destination des européens, vous avez un certain nombre de contenus qui sont spécifiquement ciblés sur heu dans les langues des pays, vous vendez des produits, des services, en euros ; tout ça, ce sont des indices qui montrent que vous vous adressez de manière assez directe et évidente à un... à un public, à des clients européens, et donc les droits de ces clients impliquent que vous acceptiez, en gros, la, la... l'application de, de... de ce règlement. Et ça, je pense que c'est quand même quelque chose d'intéressant parce qu'on a, finalement, si on veut vraiment positiver, faut aussi être optimiste des fois, et je pense que, on a vraiment passé beaucoup de temps à voir des acteurs du numérique et d'internet qui expliquaient, à juste titre, aux... aux acteurs publics que Internet était sans frontière et que vouloir faire une règle en France, une règle en Belgique, une règle en... ce n'était pas acceptable, c'était, c'est pas viable. Là j'ai l'impression qu'on a vraiment une logique de dire « effectivement, vous avez raison, ce qui est important, c'est de savoir si le droit des individus est protégé, essayer d'avoir une logique quand même de coordination et d'harmonisation européenne pour ne pas imposer des règles différentes dans tous les pays, mais heu... effectivement, ce qui compte, ce n'est pas de savoir si vous êtes installés à Kuala Lumpur, à Tel Aviv, à, à... San Francisco ou à Paris, mais heu... qui sont vos clients, où est-ce qu'ils sont et qu'est-ce qui protège leurs droits. ».

► (44:22) Éventuellement, pour finir, un peu dans le même style, est-ce qu'il y a des mesures précises qui pourraient ou qui ont déjà été prises vis-à-vis des bulles filtrantes ?

(44:45) (*long silence*) alors... qu'est-ce qu'on pourrait... alors, oui... je vais aussi rester dans à notre, à, à, à mon domaine hein ? Parce que vous allez avoir plein d'autres entretiens avec plein d'autres gens, heu... dans notre domaine, il y a deux choses qui me paraissent intéressantes à relever, c'est la question de heu... quel est le droit des personnes, quel est le... quel droit ont les personnes de comprendre le mécanisme d'une décision automatisée qui est prise. Heu... et ce qui est intéressant, moi, je trouve, c'est que, historiquement, la, donc la loi informatique des libertés, elle date de 1978, donc heu... vous êtes jeunes mais vous pouvez vous rendre compte que, moi je n'étais pas né en 1978 donc heu... vous voyez, l'informatique en 78 heu... voilà, c'était pas les mêmes, on parlait pas des mêmes outils, et pourtant, dès cette époque, le législateur a eu une vision de ce sujet, parce que l'article 10 de la loi de 78 parle de ce droit à, à... l'article 10 disait, enfin, dit toujours, parce qu'avec le règlement ça change, que heu... hum... on ne peut pas prendre une décision automatisée portant atteinte aux droits des personnes, ou alors il faut être en mesure soit d'expliquer, enfin, il y a une interdiction de la décision automatisée portant atteinte aux droits des personnes qui était un peu posé globalement dans cet... dans cet article 10, alors, qui était peu applicable, qui a été peu appliqué en tant que tel, mais je pense que c'est intéressant de voir que dès cette époque, le législateur s'est dit « une décision purement automatisée, c'est problématique parce que ça empêche de comprendre ce qui s'est passé quoi ». Heu... ça se traduit aujourd'hui en une, d'une manière un peu plus moderne, dans le règlement par un droit à demander des explications sur des, heu, sur le « profiling », en fait, le règlement prévoit, dans un de ses articles, heu... j'ai plus le numéro en tête, 22 peut-être... enfin bon, prévoit un droit à heu, un individu peut dire « j'ai été profilé, je heu... je demande que vous m'expliquiez » heu... alors je ne sais pas si dans le droit ça s'appliquerait totalement ce que je vais dire, mais pour vous donner un exemple un peu concret, on pourrait imaginer que vous demandiez qu'on vous explique pourquoi vous avez été classé dans une catégorie de risques pour un crédit et l'entreprise doit vous expliquer « ben, on a pris en compte tel critère, tel critère, tel critère, de telle manière ». Donc c'est, c'est relativement, cet article dans le règlement a fait pas mal parler au niveau international, en bien et en mal, heu avec des caricatures dans les deux côtés d'ailleurs, disant que ça crée une sorte de droit général à l'explication d'intelligence heu de décision d'intelligence artificielle et tout... alors que le périmètre de l'article est quand même un peu plus limité, et d'un autre côté, les articles qui disaient que ça allait empêcher tout essor de justement l'intelligence artificielle en Europe parce que c'était impossible d'expliquer, même pour les concepteurs, un certain nombre de choses, parce que quand on est dans les, dans les... dans les outils et cetera, mais heu... et donc je pense que dans cet angle-là, c'est est-ce que le droit peut créer des outils, des mécanismes de heu... de, de, de... d'explication à destination des personnes, et le droit européen essaie vraiment de le faire, je pense que c'est un peu balbutiant mais c'est... c'est intéressant. Le deuxième truc qui est un peu un dada personnel mais je pense qu'il est quand même en lien avec le sujet, c'est que le règlement aussi prévoit un deuxième heu... crée un droit, qui s'appelle le droit à la portabilité des données personnelles. Donc l'idée, c'est que toute personne pourra demander heu... la récupération, à récupérer et éventuellement pouvoir transmettre à quelqu'un d'autre les données personnelles qui le concernent et qu'il aura fourni à l'entreprise dans le cadre d'un service heu... et que cette récupération, elle se fait en, dans un format, les données doivent être restituées dans un format heu... communément utilisé, structuré, communément utilisé et lisible par machine. Je trouve que c'est intéressant parce que, on a quelque chose qui s'appelle le droit d'accès qui existe depuis 1978, le droit d'accès, c'est l'idée que si vous savez qu'une entreprise ou une administration traite des données sur vous, vous avez le droit de demander à accéder à ces

données. C'est très heu... c'est un droit de savoir. C'est on doit vous donner et c'est lisible par l'humain pour le coup, ce qui est demandé, c'est-à-dire qu'on doit vous remettre quelque chose en disant « voilà, je sais heu... je connais votre adresse, je connais... » voilà, normalement, c'est sensé... s'appliquer et ça s'applique plus ou moins bien, mais c'est vraiment cette idée, c'est un droit d'accéder à l'information ; et là, on a un droit en plus qui vient s'ajouter, qui est un droit de portabilité, c'est pas un droit d'accéder à l'information, c'est de la récupérer et de la réutiliser ; et pourquoi ça peut avoir un lien avec le sujet des bulles filtrantes, ça pourrait être une manière d'imaginer heu... que les individus puissent heu... ne pas être coincés dans un environnement de services heu... et tester différents, différents... à partir de leurs données, tester différentes solutions techniques, algorithmiques, voir les différences, voilà, peut-être voir, par rapport à ce que j'ai évoqué tout à l'heure, l'idée un peu de « crowdsourcer » les, heu, les résultats, enfin, au moins d'avoir une base de données qui permette peut-être de faire ça, enfin, ça peut, ça peut vraiment créer... là aussi, c'est très nouveau, ça rentre en application l'année prochaine, personne sait trop comment ça va, ça va... ça va évoluer, mais je pense que c'est des outils qui sont intéressants parce qu'ils sont assez modernes quand même, pour, pour heu... pour des outils juridiques, on va dire, qui avancent moins vite que les outils techniques.

►(50:21) Et bien parfait, ça sera tout pour nous, à moins que vous vouliez ajouter peut-être quelque chose ?

(50:27) Non, non, c'est juste bien. C'est très bien qu'on parte tout de suite et puis c'est même mieux, que ça évite que vous m'expliquiez trop de choses et que ça biaise justement mes réponses, mais juste, est-ce que vous pourriez m'expliquer en deux secondes comment se passe votre travail ? *(Nous expliquons alors le principe du projet ETIC ainsi que comment l'interview sera rendue visible sur le site web lorsque celui-ci sera terminé.)*

►(53:05) Fin de l'interview.