

Au secours ? Amazon veut rendre Alexa intelligente en traitant nos données personnelles



Rohit Prasad, chef scientifique du développement d'Alexa, a partagé avec la MIT Tech Review les futures capacités qu'Amazon envisage pour Alexa. Alexa aura plus de capacités de prédiction, de prise de décision, et sera capable de raisonnements de plus haut niveau.

Avec Benoît
Berthelot

Atlantico : Comment Alexa est-elle devenue un produit phare d'Amazon?

Benoît Berthelot : Amazon est certainement en train de devenir un acteur très puissant grâce à Alexa.

Ce n'était pas nécessairement planifié au début du développement de cette technologie. Toutefois, comme Amazon est une entreprise très pragmatique, elle développe ce qui marche. Amazon a donc mis beaucoup de moyens pour développer Alexa, qui était, avant de devenir un produit en soi, une fonctionnalité créée pour le Fire phone. Ce smartphone a connu un échec retentissant, mais Jeff Bezos a trouvé que la fonctionnalité Alexa était très intéressante. Il a donc décidé de l'intégrer dans une enceinte connectée du type Echo. C'est un peu comme l'ordinateur dans Star Trek : on peut lui parler, et l'ordinateur opère à partir de ces interactions. C'est donc la genèse d'Alexa, qui a été lancée en 2015, aux Etats-Unis.

Amazon a compris que c'était un produit qui trouvait très bien sa place sur le marché, alors que la technologie proposée existait en fait déjà depuis un certain temps, avec Siri. Le fait qu'Alexa était dans le foyer, dans la maison, libérait les utilisateurs du complexe de parler à une machine : chez eux, personne ne les juge, ce qui n'était pas le cas de l'interface vocale sur téléphone portable, un désavantage pour Siri.

Amazon a donc créé des applications, et a fait en sorte que les gens commencent à utiliser Alexa. L'entreprise a ensuite amélioré la technologie de façon spectaculaire année après année. Plus de 10 000 personnes travaillent désormais pour Alexa chez Amazon. Ces équipes ajoutent des nouvelles langues tous les six mois. Amazon est en train de s'assurer qu'Alexa devienne l'interface vocale préférée des utilisateurs partout dans le monde, et qu'Alexa soit partout.

Le responsable France d'Alexa l'a dit : il veut qu'Alexa devienne un ordinateur ambiant, auquel on peut s'adresser dans n'importe quel contexte, chez soi, dans sa voiture, sa chambre d'hôtel, ou dans la rue. Amazon a donc lancé les lunettes Alexa, la bague Alexa, et des milliers d'appareils qui sont compatibles avec Alexa. C'est leur vraie force.

Désormais, pratiquement tous les constructeurs peuvent intégrer Alexa à leurs produits, à condition que les données soient récupérées par Amazon. Free, SFR, Orange ont intégré Alexa dans leur box internet et autres produits : beaucoup de foyers français vont donc être concernés.

Doit-on s'inquiéter de la collecte de données personnelles que les nouvelles capacités d'Alexa

impliquent?

Ce qui peut en effet susciter l'inquiétude, c'est l'opacité du traitement qui est fait des données récoltées par Alexa. Car contrairement à Google Home ou l'Apple Homepod, Alexa collecte les données par défaut. Il n'y a pas de possibilité de désactiver la collecte des données par Amazon. Vous pouvez dire à Alexa : « efface ce que je viens de dire », ou « efface ce que j'ai dit pendant les trois derniers mois ». Vous pouvez donc a posteriori supprimer des données, mais par défaut Alexa les collecte et les stocke.

Ces données nourrissent le profil de chacun des utilisateurs, et on ne sait pas ce qui sera fait de ces données dans les années à venir.

Sait-on déjà si Amazon utilise ces données à des fins de marketing pour d'autres produits?

Amazon dit que ces données ne sont collectées pour être ajoutées au profil global du consommateur Amazon seulement quand ce dernier utilise des applications tierces, comme l'application Amazon. Si vous êtes abonné à Amazon et que vous dites à Alexa : « Alexa, achète un paquet de croquettes », cet achat rejoint le stock de données d'Amazon et la connaissance qu'Amazon a du profil de ce consommateur.

Amazon est aujourd'hui un acteur très puissant dans la publicité précisément car ils disposent d'une base de données sur leurs clients qui est extrêmement riche. Ils détiennent la vérité de ce que consomment les gens. Ils le commercialisent très largement à des entreprises, qui peuvent alors acheter les profils des utilisateurs d'Amazon et faire de la publicité à ces clients d'Amazon où qu'ils aillent sur internet.

Amazon dit qu'elle ne vend pas les données récoltées par Alexa à des entreprises tierces. Mais dès que les données sont récoltées par Alexa, elles enrichissent bien le profil global des consommateurs, et Amazon n'a aucune raison de ne pas s'en servir.

On peut dire la même chose également de nos recherches Google. Google est très précis, et c'est bien parce qu'on donne nos données et qu'elles sont très personnelles. Amazon veut rendre Alexa beaucoup plus personnelle, contextuelle, pour nous proposer des choses pertinentes pour notre prise de décision. Alexa va beaucoup mieux nous connaître pour mieux fonctionner, de manière personnalisée.

Rohit Prasad citait quelques mesures prises pour pouvoir protéger la vie privée des utilisateurs. Il citait l'effacement périodique possible des données et les options de retrait, c'est-à-dire la possibilité de demander à ne pas être inclus dans certaines collectes de données faites par Alexa. Que pensez-vous de ces mesures de protection?

Je pense que ces mesures ont été prises en réaction aux scandales qui ont été provoqués à la suite de certains articles. Ces articles démontraient que des données avaient été stockées et que les interactions avec Alexa étaient écoutées par des personnes bien réelles pour améliorer l'algorithme de la reconnaissance vocale.

Ce qui est incompréhensible, c'est pourquoi Amazon ne permet pas aux utilisateurs, par défaut, de modifier leurs profils pour empêcher la collecte des données. Cela révèle une certaine hypocrisie dans leur approche.

Amazon devrait donc proposer à ses utilisateurs des options « opt-in », de choisir quelles données ils veulent bien livrer, plutôt que des options « opt-out », où les utilisateurs s'excluent uniquement de quelques catégories définies par Amazon?

Exactement. Ce qui serait vraiment juste envers le consommateur, ce serait de pouvoir choisir de partager ou non les données avec Amazon pour améliorer Alexa. A minima, il faudrait pouvoir décocher le stockage des données par Alexa.

Le stockage des données par Alexa est assez extensif. J'ai utilisé Alexa pour la tester dans le cadre de l'écriture de mon livre *Le Monde Selon Amazon*, pendant deux ans. J'ai découvert que 1 112 interactions avaient été stockées par Alexa en 2 ans. Cela comprenait non seulement le texte de mes interactions mais aussi l'audio, ma propre voix. Par moment, Alexa se déclenchait par erreur à cause de bruits dans l'appartement, émis par la télévision, ou par une parole qu'elle avait entendue comme un ordre. Ces dysfonctionnements sont inquiétants, à cause de l'opacité et de l'incertitude de cette collecte de données.

Certains critiquent Amazon pour son approche binaire dans les options de retrait : il est demandé aux utilisateurs de choisir entre ne rien partager ou tout partager dans des catégories aux périmètres très larges. Pensez-vous qu'il s'agisse d'une stratégie pour pouvoir justifier d'un choix laissé aux utilisateurs, quand bien même ce choix rend intentionnellement très peu attractif le choix de refuser la collecte des données?

Amazon ne veut pas avoir à s'embarrasser du tri des données, ils veulent donc que l'on accepte toutes les conditions d'utilisation. La quantité de données que nous cédon à Alexa, à Google ou à d'autres, devient tellement importante que l'on pourrait envisager un système dans lequel les données de santé, les données financières et les données d'ordre privé, devraient être exclues des collectes.

Mais ce serait pour eux très compliqué à mettre en place, et ce n'est pas du tout dans leur intérêt, donc il n'y a pas de raison qu'ils le proposent.

Au Royaume-Uni, beaucoup de données de santé sont collectées par Alexa. En effet, un partenariat avec les services de santé britannique permet aux utilisateurs d'Alexa de poser des questions au National Health Service. Nous ne sommes qu'au début de ce genre d'interrogations, et de ce type de demandes de la part de certains utilisateurs.

Les ambitions qu'Amazon a pour Alexa vont-elles accentuer véritablement les problèmes de collecte de données que vous évoquez?

De plus en plus de questions vont se poser, effectivement. Surtout avec le développement des « conversations » avec Alexa . Amazon la rend de plus en plus intelligente. On va de plus en plus se rendre compte de la masse de données qu'Amazon a stockée au fur et à mesure qu'Amazon va utiliser ces données pour améliorer Alexa. Alexa va devenir de plus en plus proactive dans ses propositions. Si demain Alexa vous rappelle de prendre vos médicaments parce que le distributeur de médicaments affiche que vous les avez oubliés, si elle vous dit que vous avez consommé trop de calories pour aller au restaurant, ou vous propose de commander un Uber pour la soirée de cinéma que vous avez prévue, on va se rendre compte du stockage de données.

Alexa est au fond une nouvelle illustration, très tangible, de ce que font tous les GAFA. Comme c'est une interface vocale, qui a vocation à être omniprésente au travail, dans la voiture, à la maison, ces pratiques seront toutefois dans ce cas très prégnantes, très visibles.

Il faut ajouter aux problématiques posées par Alexa celles posées par la maison connectée qu'Amazon développe. Par exemple, Amazon a créé les sonnettes vidéos, connectées à Alexa. Ce système peut vous dire qui frappe à votre porte. C'est une facette du modèle de surveillance, de sécurité, qu'Amazon développe beaucoup. Cette problématique de la surveillance va s'ajouter à la problématique de la collecte des données et à son utilisation.

Vous mentionnez le partenariat entre les services de santé britannique et Alexa, qui n'existe pas en France. Est-on plus alerte sur ce débat?

Nous sommes en effet plus prudents et plus sceptiques en France. Par exemple, les sonnettes vidéos ne se développent pas à la même vitesse chez nous. Nous sommes loin d'avoir des partenariats entre Google et la police, comme c'est le cas aux Etats-Unis. Mais il n'y a pas de raison pour que ces usages n'arrivent pas en France.

Je ne crois pas que demain la Sécurité Sociale va faire un partenariat avec Alexa. Toutefois, je crois que le site Doctolib pourrait le faire.

Les interfaces vocales ont vocation à remplacer le smartphone, qui est pour l'instant omniprésent. Les usages des technologies comme Alexa vont donc se développer.