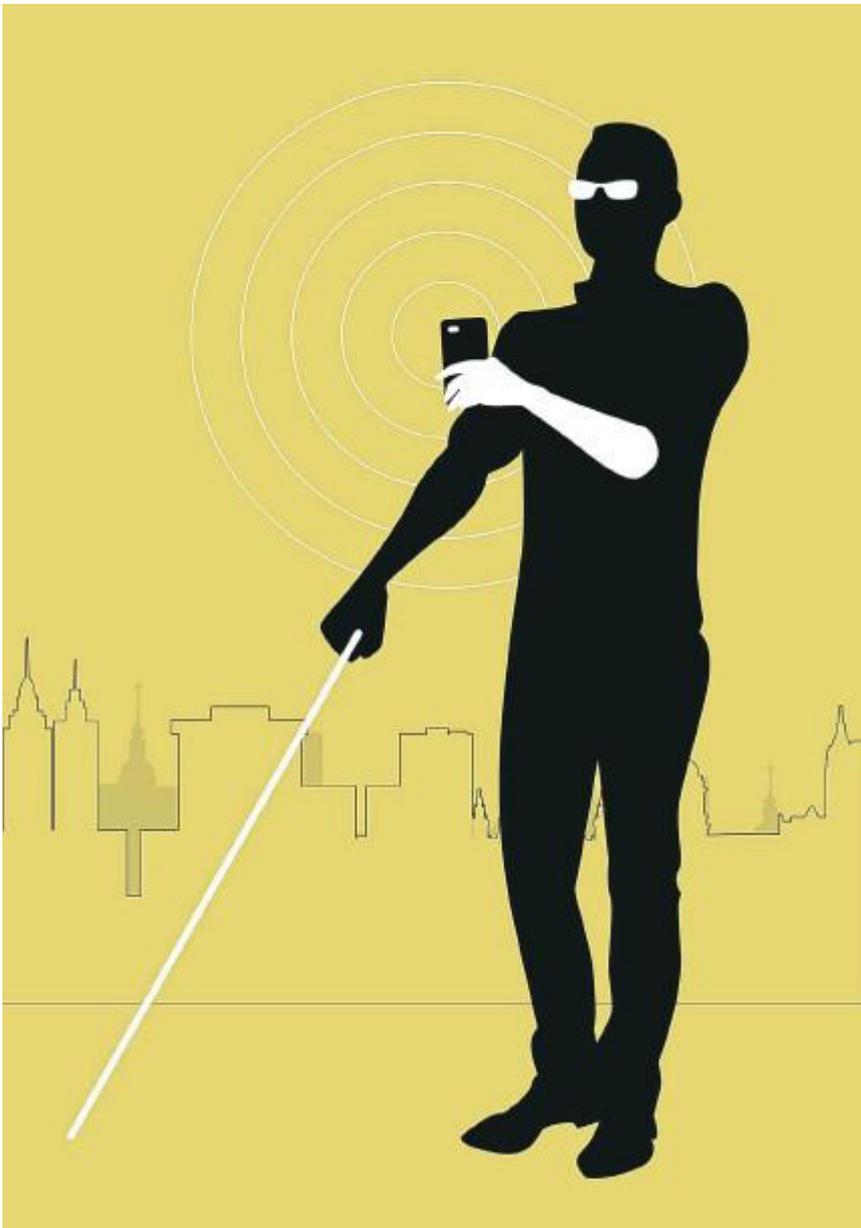


Handicap et IA : les voix de l'accessibilité

Grâce aux progrès de l'intelligence artificielle, de nouveaux outils numériques permettent aux personnes malvoyantes ou malentendantes de mieux percevoir leur environnement et de communiquer plus facilement.

Les Echos · 12 febr. 2019 · Rémy Demichelis RemyDemichelis

C'est une cuisine en longueur, tout ce qu'il y a de plus normal pour un appartement parisien, avec une table pour prendre le petit déjeuner ou offrir un café. Il n'y a pas de grande surprise à vue d'oeil, et il n'y en aura pas. Chez Christian Lainé, formateur en informatique adaptée à la Fédération des aveugles de France, la maison n'est pas « toute connectée ». Nous sommes encore loin des espaces de démonstration de Google où, de la machine à pop-corn à la lumière du salon, tout est relié à un seul assistant vocal – des vitrines qui donnent au concept de « domotique » un sacré coup de vieux. Mais on y vient.



Sur la table est posée une discrète enceinte Google Home qu'il appelle pour consulter la météo ou écouter la radio. « OK Google : mets franceinfo. » Les assistants vocaux sont en phase de devenir un outil essentiel pour Christian Lainé : « Ça ouvre de superbes opportunités aux aveugles, on est à égalité avec les voyants devant Google Home », se réjouit-il. Il peut acheter des billets de train via son enceinte, consulter son agenda. Il hésite encore sur la marque pour son thermostat connecté, l'américain Nest ou le français Netatmo, et attend avec impatience de nouvelles applications comme le four micro-ondes équipé d'un assistant vocal annoncé par Amazon. Chez lui, il peut aussi commander sa télévision Samsung avec la voix, mais elle n'est pas encore reliée à Google Home, ni ses plaques à induction ou sa balance de cuisine, qui donnent à l'oral le niveau de chauffe ou le poids. Fernando Pinto Da Silva, responsable numérique de la Fédération des aveugles de France, espère, quant à lui, pouvoir relier des ampoules à son assistant vocal : « Ma compagne et moi sommes aveugles, mais j'aimerais savoir quand les lumières sont allumées ou éteintes pour ne pas laisser notre fille, qui voit bien, dans le noir. » Seulement, il trouve leur coût encore trop élevé : 20 euros au minimum par ampoule, et toutes ne sont pas compatibles avec son système. Lui a opté pour Alexa d'Amazon.

Le boom des assistants vocaux

La voix devient une nouvelle interface pour accéder aux outils de tous les jours et à Internet. « Aujourd'hui, 20 % des recherches sur mobile sont réalisées par la voix », note Cécile Pruvost, manager stratégie et développement chez Google, lors d'une conférence à l'Institut national des jeunes aveugles (Inja), à Paris, organisée début décembre. Les ventes d'enceintes connectées sont estimées à 1,7 million en France pour leur première année de commercialisation, selon Médiamétrie. En 2018, le chiffre se situerait entre 62 millions et 75 millions dans le monde, au moins le double par rapport à l'année précédente. Un marché dominé par Amazon.

Le plus souvent, à chaque marque d'enceinte connectée ou de smartphone correspond un assistant : Siri d'Apple pour HomePod et l'iPhone, Google Assistant pour la Google Home et Pixel, et Alexa d'Amazon pour Echo. Mais bientôt les assistants vocaux seront peut-être comme des navigateurs, dissociés du hardware. Orange a déclaré que sur sa Livebox seront proposés le sien, Djingo, ainsi qu'Alexa. La fondation Mozilla de logiciel libre, à l'origine de Firefox, développe son propre assistant vocal. De quoi faciliter ce que l'on désigne par le concept barbare d'« interopérabilité des systèmes » ou plus simplement « tout-connecté ».

Ces outils ne sont pas sans soulever des questions. Lors de la conférence à l'Inja à laquelle participe Google, une personne aveugle du public s'inquiète du devenir de ses données personnelles, celles-ci devant être envoyées sur les serveurs du géant du Net pour être traitées : « Je n'ai pas envie que Google sache tout de moi. » Réponse de Cécile Pruvost : « Chaque utilisateur peut partager ou effacer son historique. » Les constructeurs insistent aussi sur le fait que seuls les moments où il est ordonné à l'enceinte de se réveiller sont traités.

Afin de court-circuiter le problème, la start-up française Snips a présenté, au CES (Consumer Electronics Show) de Las Vegas, des puces intelligentes à embarquer sur de l'électroménager : fours, aspirateurs, etc. Son credo est le traitement des informations en local : les données ne quittent pas le hardware. Une technologie qui répond toutefois à des commandes assez simples. Il faut donc apprendre à l'utiliser tandis que les géants du Web cherchent une interaction avec l'outil la plus intuitive possible : « Il y a 5.000 façons de programmer son réveil », indique Cécile Pruvost.

Systèmes de reconnaissance visuelle

Les avancées dans le domaine de l'intelligence artificielle (IA) ont été rendues possibles grâce au traitement des données de masse, dont celles de Google : plus il y a de données, plus la machine peut s'entraîner pour apprendre à reconnaître des motifs. Les techniques mises au point s'appliquent aujourd'hui autant au son qu'à l'image. Depuis 2015, les systèmes de reconnaissance visuelle sont plus pertinents que l'humain pour identifier des objets ou des animaux dans des photos.

Microsoft a été le premier à faire une telle annonce. La même année, l'entreprise israélienne OrCam, qui travaillait aussi sur cette technologie, lance la commercialisation d'une caméra dite intelligente pour malvoyants : MyEye. Montée sur des lunettes et dotée d'un petit haut-parleur, elle permet de lire un texte, de reconnaître des billets de banque ou des personnes et de les annoncer à haute voix. Avec ce dispositif, les données restent en local.

Son coût, pouvant être partiellement pris en charge par la MDPH (Maison départementale des personnes handicapées), s'élève de 3.750 euros à 4.750 euros.

Microsoft s'est aussi lancé sur ce marché avec l'application Seeing AI, disponible en France depuis décembre 2017, actuellement seulement en anglais, la version française devant arriver courant 2019. L'outil est gratuit, mais récupère les données pour les traiter dans le cloud. Il permet en plus de décrire des scènes (un aspect perfectible) et des personnes ou d'indiquer le niveau de luminosité. Il n'est disponible que sur l'AppStore d'Apple pour l'instant.

Ce n'est pas un hasard : « Je ne connais aucun aveugle qui n'a pas d'iPhone lorsqu'il a un smartphone », explique Fernando Pinto Da Silva. La raison en est simple, Apple est le premier constructeur à avoir inclus une synthèse vocale dans ses appareils pour lire l'écran, VoiceOver, en 2009, sur l'iPhone 3GS. « C'était pour accéder au marché de l'éducation américain », souligne Fernando Pinto Da Silva. L'accessibilité étant un prérequis pour les services au public aux Etats-Unis, et les puissantes fédérations d'aveugles veillent à faire respecter la loi.

En France, une loi de 2005, complétée par la loi pour une République numérique de 2016, garantit, en théorie, l'accessibilité des sites Internet des collectivités et des délégataires de service public, mais Fernando Pinto Da Silva déplore que les textes ne soient pas suffisamment appliqués. Aucune sanction n'est tombée, mais un décret prévu pour le premier semestre 2019 doit améliorer le suivi et étendre l'obligation aux grandes entreprises. Il y a aujourd'hui en France 1,7 million de personnes aveugles et malvoyantes. Elles sont environ 250 millions dans le monde, un chiffre qui pourrait être multiplié par deux ou trois d'ici à 2050 en raison de la croissance démographique et du vieillissement de la population. De quoi constituer un marché susceptible de susciter l'intérêt du privé.

Mais la réglementation semble rester la meilleure incitation. C'est en tout cas ce qu'a constaté Olivier Jeannel avec son application RogerVoice destinée aux personnes sourdes (comme lui) et malentendantes, qui sont 4 millions en France : « La loi a donné un coup d'accélérateur. » Depuis le 7 octobre, les grandes entreprises ont en effet pour obligation d'avoir un service téléphonique accessible à ce public. RogerVoice permet de retranscrire à l'écrit un dialogue par téléphone, et de répondre en tapant sur le clavier (une synthèse vocale s'occupe ensuite de lire le texte à haute voix). En plus, il est possible de faire appel à un interprète de langue des signes via la vidéo. Une heure est offerte par mois, puis les forfaits vont de 5,99 euros à 29,99 euros.

L'idée lui est venue quand il était aux Etats-Unis, justement parce que l'accessibilité y était plus développée, mais il précise que son outil est sans pareil. Son entreprise est d'ailleurs un opérateur téléphonique, ce qui permet à l'application d'être utilisée même quand l'interlocuteur ne la possède pas. Dernièrement, Olivier Jeannel a signé pas moins d'une vingtaine de partenariats avec des entreprises comme OuiSNCF, Allianz, GRDF ou Aéroport de Paris. De même, il aide d'autres opérateurs tels Orange, Bouygues et SFR.

Mais il n'a pas l'intention de les concurrencer : « On complète ce qu'ils font, nous n'avons pas de réseau, pas de carte SIM. » Avec 18 personnes, effectif plus qu'honorable pour une start-up créée en 2013, ce serait encore difficile. Ça ne l'empêche pas de voir grand. Son application est disponible (transcription et synthèse vocale uniquement) au Royaume-Uni,

en Allemagne, au Mexique et au Japon. Il vise maintenant les Etats-Unis et le Canada. Son coeur de métier est l'application, pas l'IA, fournie par un prestataire, mais il déclare quand même travailler sur un système de reconnaissance des émotions : « Quand j'appelle mon père, j'aimerais savoir s'il est content ou fâché. »

RogerVoice collabore avec Ava, une start-up destinée au même public cofondée notamment par le Français Thibault Duchemin. Comme Olivier Jeannel, il est passé par Berkeley, aux Etats-Unis, mais quelques années plus tard. Né en 1991, Thibault Duchemin est le seul qui entend bien de sa famille. Ayant eu l'habitude de jouer l'interprète durant son enfance, il a eu l'intuition de développer un outil pour faciliter les dialogues entre personnes sourdes et ceux qui ne maîtrisent pas la langue des signes : « Le problème, ce n'est pas l'audition, c'est la communication », défend-il.

Ava permet de retranscrire des discussions entre plusieurs personnes sur un écran de smartphone, mais il faut que chacun ait téléchargé l'application – un service gratuit pendant cinq heures par mois, puis à 29 euros en illimité. Elle a été lancée en 2016 aux Etats-Unis et cet été en France, Thibault Duchemin revendique 100.000 utilisateurs mensuels dans ces deux pays. Il ne veut « pas attendre que Google ou Apple développe ces logiciels et qu'on récupère les miettes ». Pour lui, son application ne doit pas être réservée aux sourds et malentendants, elle doit aussi être utile pour ceux qui veulent rattraper une réunion ou la suivre à distance.

Les outils d'accessibilité ont généralement pour but d'autres usages ; « Seeing AI est un outil de promotion du service cloud », explique Philippe Trotin, directeur accessibilité chez Microsoft. L'entreprise OrCam va, quant à elle, lancer en mars une caméra connectée de développement personnel pour tout un chacun. « L'assistant de Google n'a pas été créé dans le but premier d'aider les déficients visuels, mais c'est aujourd'hui un élément clef », indique Cécile Pruvost. L'accessibilité est devenue un terrain de choix pour l'élaboration des systèmes d'intelligence artificielle, à la fois en matière d'image, pour l'intérêt technique et pour perfectionner des outils qui doivent s'adresser à tous. Certains l'ont compris, d'autres attendent que la loi les y oblige. ■

Apple est le premier constructeur à avoir inclus une synthèse vocale dans ses appareils pour lire l'écran, VoiceOver, en 2009, sur l'iPhone 3GS.