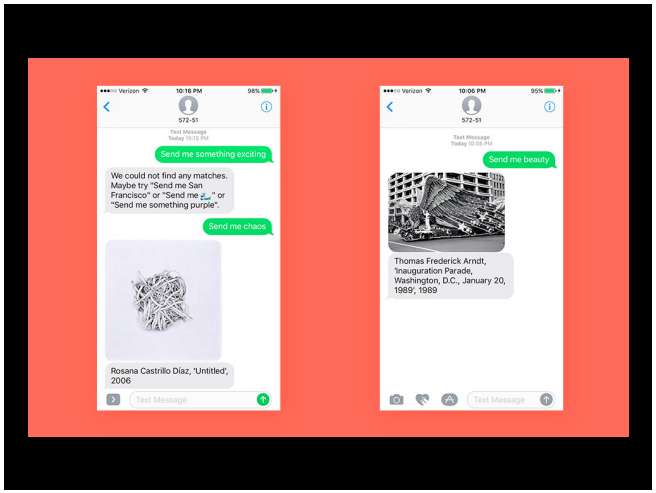
# *Quelles questions posent les assistants vocaux au design d'interface ?*

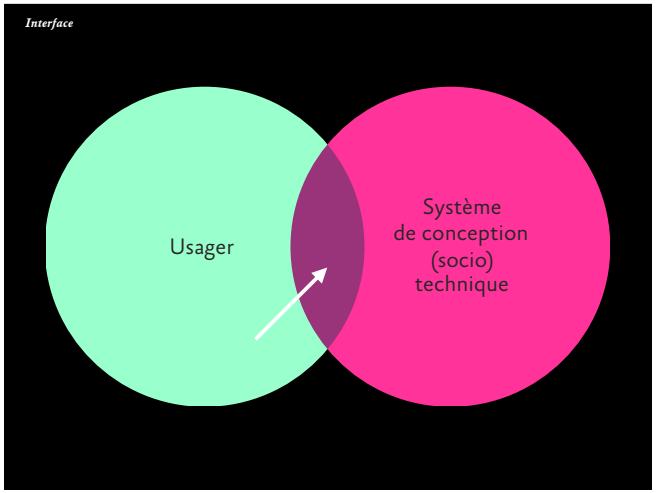
Julien Drochon
2019




3

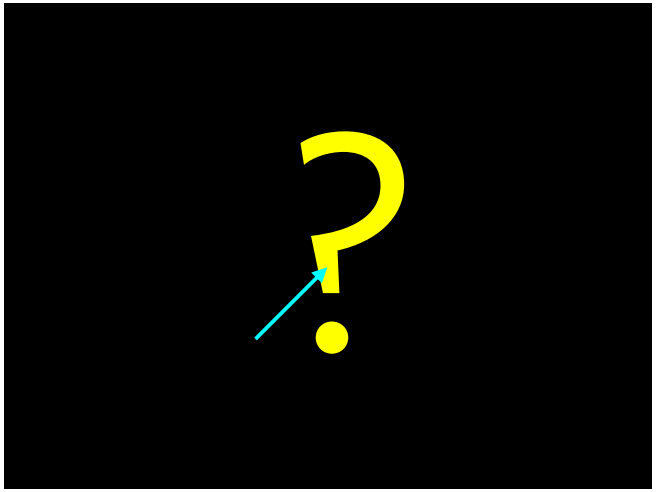
chatbots, agents conversationnels… sont autant de terminologies pour les interfaces conversationnelles. Les assistants vocaux tels que nous les connaissons s'inscrivent dans une histoire d'objets technologiques qui déjouent les effets d'annonce d'innovation disruptive des industriels de l'informatique.

---




4

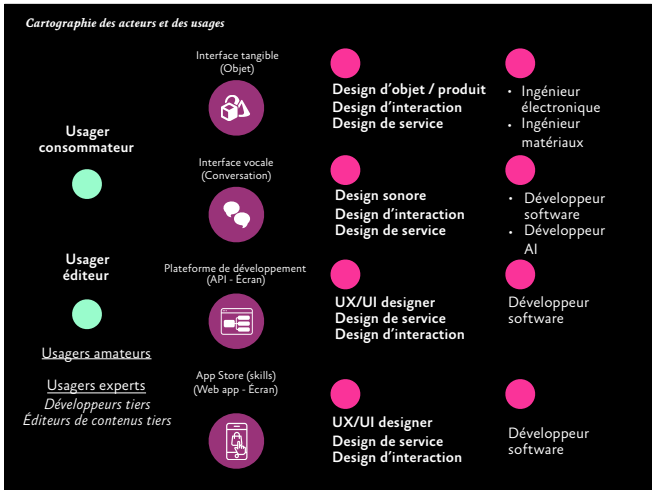
Une interface est un élément qui détermine une surface commune entre 2 milieux. Le design a toujours essayé de faire exister l'interface comme une surface sensible entre l'humain et l'objet. Sensible comme donner un sens et comme solliciter les sens. Le designer objet passe par la matière pour y arriver, le designer graphique par l'image et la typographie, …



5

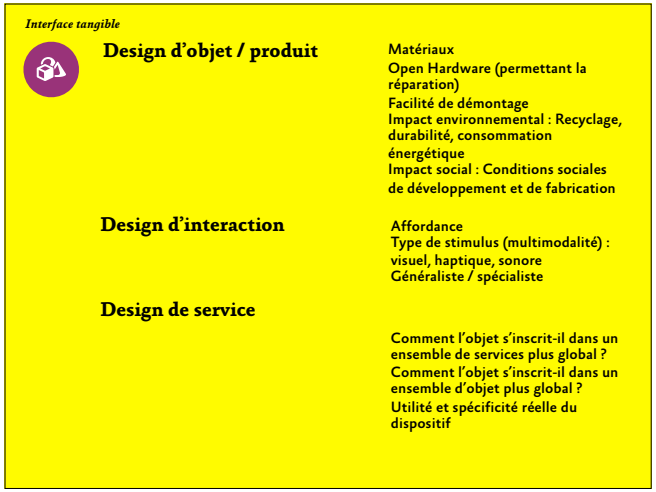
En terme de design, l'utilisation de la voix seule déplace les enjeux de conception des interfaces. Si les interfaces graphiques relèvent principalement de structuration de la surface, les interfaces tangibles de l'espace, les interfaces vocales relèvent principalement d'une structuration du temps.

Quel est alors l'endroit des interfaces vocales ?




6

Les interfaces vocales n'ont pas d'endroit commun entre la machine et l'humain (contrairement à la souris, au clavier, à l'écran). L'endroit commun se trouve être celui de la conversation. Le dispositif socio technique dans lequel existe le développement d'une interface vocale mêle de nombreux aspects de design, qu'il s'agisse du design d'objet, design d'interaction, design sonore,… Le design d'interface vocale relève à l'heure actuelle d'un grand nombre de spécialités du design lorsqu'on le considère dans sa totalité. Ces disciplines a priori périphériques en dessinent cependant une cartographie en creux.



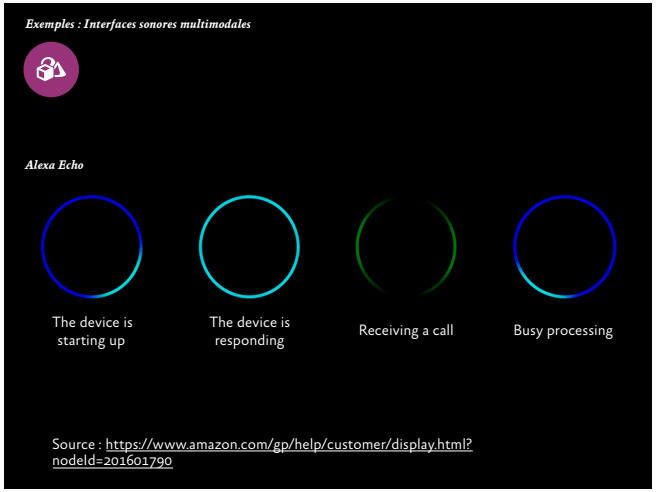

7

Si on considère son interface tangible, l'assistant vocal relève du design d'objet. Et donc des problématiques qui lui sont contemporaines : préoccupation marquée pour la gestion des ressources des matériaux, durabilité, inscription la production dans une économie circulaire.

---

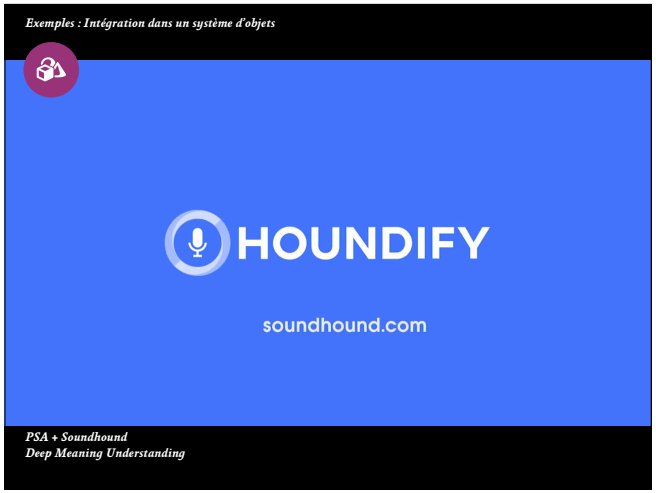
8

Certains fabricants d'objets électroniques, plus précisément de haut-parleurs, inscrivent le cycle de leurs objets dans une économie circulaire en concevant un objet dont les matériaux sont issus du recyclage. Combien d'assistants vocaux relève-t-ils d'une démarche d'éco conception ?



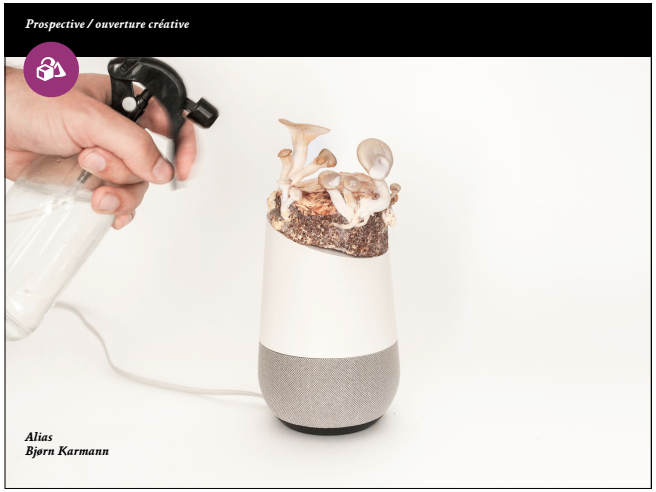

9

Cette diapositive montre quelques animations du cercle lumineux présents de l'interface tangible de l'assistant vocal d'Amazon, Echo Dot. Ces animations s'activent lors des phases d'interactions silencieuses de l'objet. Elles affichent par un procédé multimodal l'état actif du dispositif lors de ses phases d'inaction vocale. L'objet doit montrer continuellement son activité. L'objet interactif est avant tout un objet hyperactif.




10

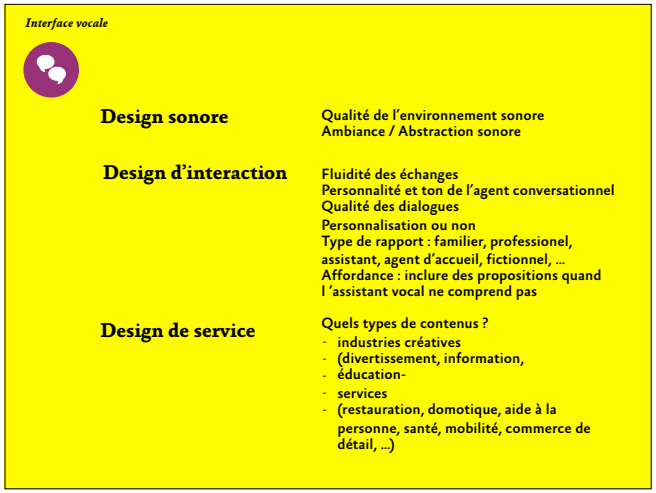
Nous considérons principalement les assistants vocaux intégrés à nos smartphones où reposant sur nos tables sous des formes plastiques ramassées. À terme, certains tendent à s'inscrire dans des systèmes d'objets selon les principes de l'informatique pervasive. PSA a récemment annoncé son partenariat avec Soundhound pour équiper ses véhicules neufs en 2020.




11

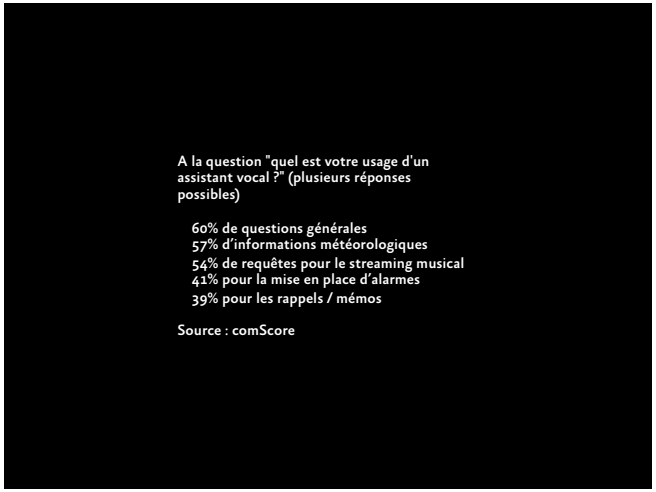
Devant l'envahissement de l'espace par la parole synthétique, le designer Bjoer Karmann lui oppose un parasite. Le projet Alias est une extension qui se glisse sur la partie supérieure des interfaces tangibles des assistants vocaux. À l'intérieur un dispositif électronique émet un signal parasite qui brouille toute reconnaissance vocale tant que l'usager n'indique pas le contraire. Proposition de design biomimétique, projet de design manifeste démontrant comment la culture des makers redéfinit la délégation passive que l'usager autorise au designer.

---




12

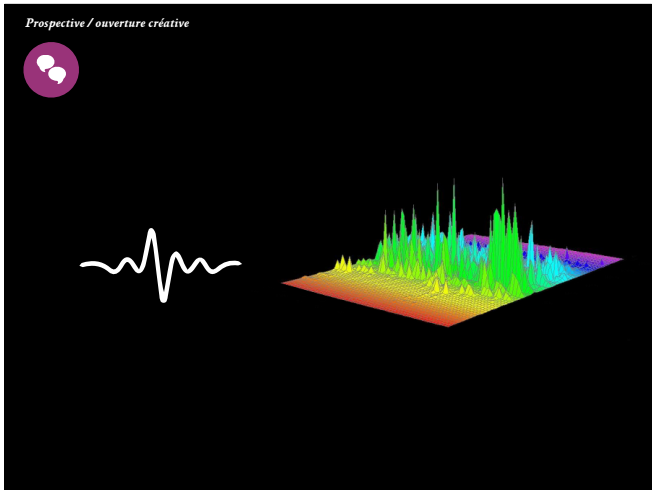
J'ai pu expérimenter depuis 3 mois, 2 assistants vocaux dans mon appartement. Principalement utilisés comme radio, comme alarme, aucun usage particulièrement adapté à ce type de périphériques.




13

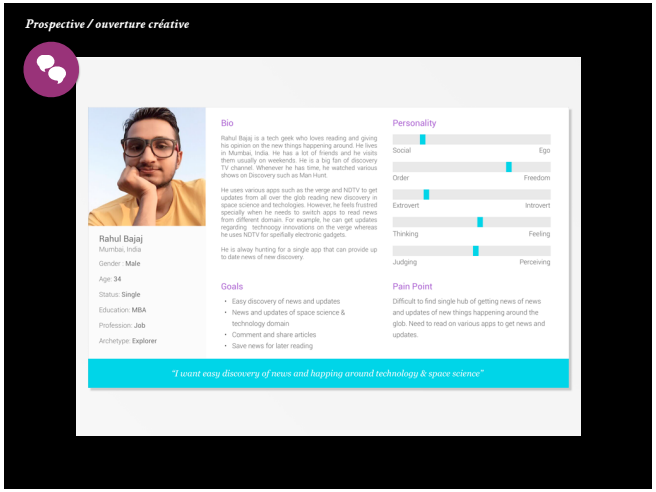
En regardant ce sondage, les usages montrent que l'objet a du mal trouver des spécificités d'usage. Quels seraient les usages possibles, inenvisagées par les ingénieurs d'Amazon ?

---




14

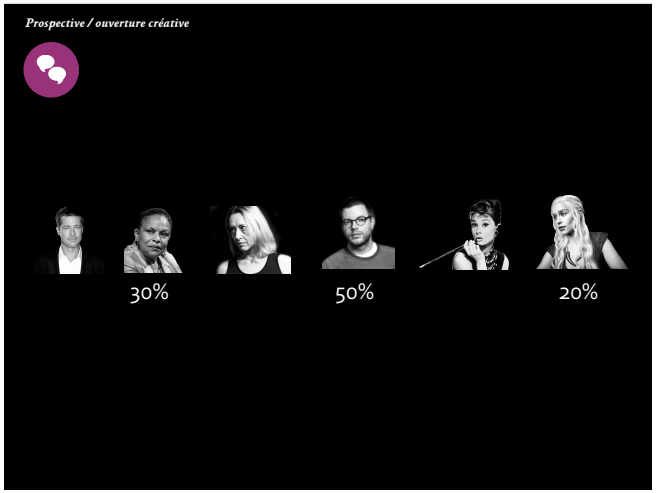
Au delà du jingle et des caractéristiques de la voix, peut-on imaginer un design sonore moins pauvre accompagnant les interactions, proposant des paysages sonores en fonction de l'heure, de la saison, du temps, en fonction de ce que l'usager dit…




15

La méthode visant à réaliser des Persona est souvent utilisé dans le design d'interaction pour projeter la conception dans un cas de figure en situation. Le persona peut désormais constituer une méthode pour la conception d'agent conversationnel : la personnalité est ce qui définit le style de ce type d'interface.

---




16

Plutôt que d'être une interface de télécommande ou de contrôle à distance, l'assistant pourrait peut être devenir une personnalité à découvrir au fur et à mesure des utilisations. Le persona permet de donner un caractère, s'inspirant de personnes fictionnelles, réelles, intellectuelles ou familières. En complément du design sonore, il s'agit de penser un design de la personnalité de l'interface qui passe par la voix tant sur le point de la tonalité (tonalité, intonation, grave aigu, tessiture de voix) que sur celui du ton (tonalité, intonation,…) ou de l'humeur (grave aigu,
est elle drôle, sérieuse, narquoise,…).

17

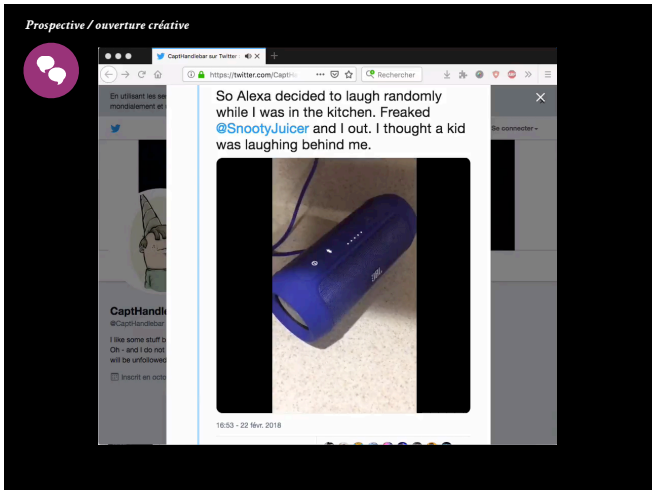
On peut imaginer Un assistant vocal qui devient désagréable quand on passe trop de temps avec lui…

Imaginez un haut-parleur intelligent qui est gêné lorsqu'il ne trouve rien, fatigué après une journée chargée, attristé par de mauvaises nouvelles, excité par les visiteurs, ou qui a froid ?

Peut-être que des dispositifs d'expression émotionnelle pourraient fournir des indices plus utiles aux utilisateurs - si c'est le cas, cela ouvre la voie à un large éventail d'applications possibles. Il pourrait fournir une forme d'interaction de soutien pour améliorer les cas d'isolement chez les

18

Qu’est ce que pourrait être un souvenir sonore, et quelles seraient les modalités de sa consultation ? Tirés du son d’une vidéo souvenir d’une bibliothèque d’archives personnelles à la demande de l’utilisateur, cette ambiance sonore pourrait exister par instant, comme un écho ou une photo souvenir.



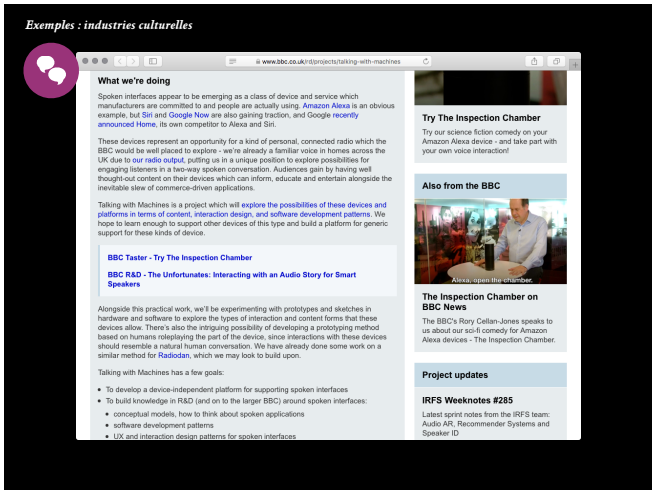

19

La réalité pour l'instant est que les assistants vocaux basculent de caractères neutres à des états quasi psychotiques…

---

20

L'expérience du jeu vidéo pourrait fournir les éléments de méthode de création de narration interactives vocales et sonores. Les jeux vidéos prennent aujourd'hui la voix comme mode d'interaction entre les personnages, de manière de plus en plus marquée. Par exemple Firewatch est un jeu qui privilégie l'interaction entre le personnage incarné par le joueur et sa supérieure, moteur du récit.




21

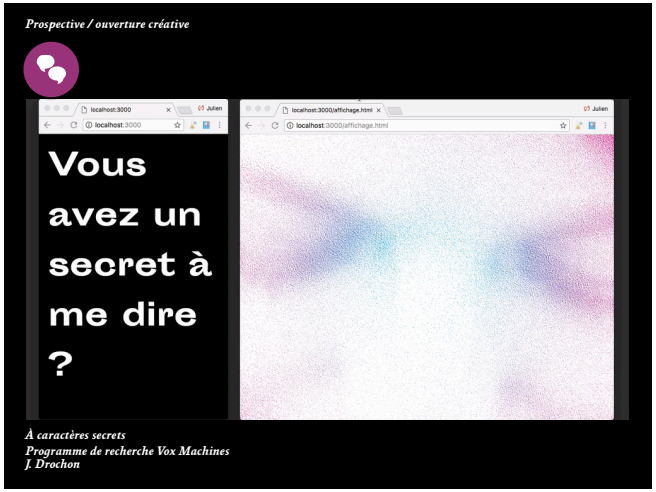
Devant l'accroissement de la possession d'assistants par les ménages anglais, la BBC a ouvert un département R&D consacré à la recherche et développement liés aux machines parlantes. Ils explorent les possibilités des assistants vocaux en terme de contenus, design d'interaction et outils logiciels. Il est nécessaire que les éditeurs de contenus s'interrogent des possibilités de création offertes par ces périphériques s'ils en souhaitent augmenter la qualité.

---

22

Un des projets réalisés par cette équipe est The Unfortunates.

L'une des parties les plus difficiles de la création d'une application pour haut-parleurs intelligents - certainement plus difficile que l'écriture du logiciel lui-même - est la conception du flux de conversation à travers l'application, la cartographie de ce que l'appareil va dire et de prévoir ce que les utilisateurs pourraient dire en réponse et à faire des demandes de fonctionnalité. L'équipe de BBC Research a donc découpé les contenu en 3 types : l'histoire principale, The Advance Guard of the Avant-garde (un documentaire sur la littérature expérimentale des années 1960), et une discussion en coulisse (comme un commentaire du réalisateur sur un DVD).

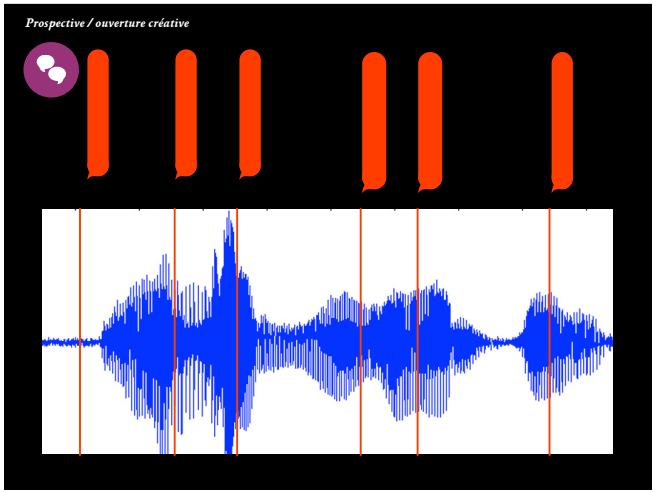



23

Ce prototype repose sur une conversation secrète ou 2 secrets se rencontrent pour se dévoiler : celui de l'utilisateur/énonceur actif et celui d'une énonciation précédente.

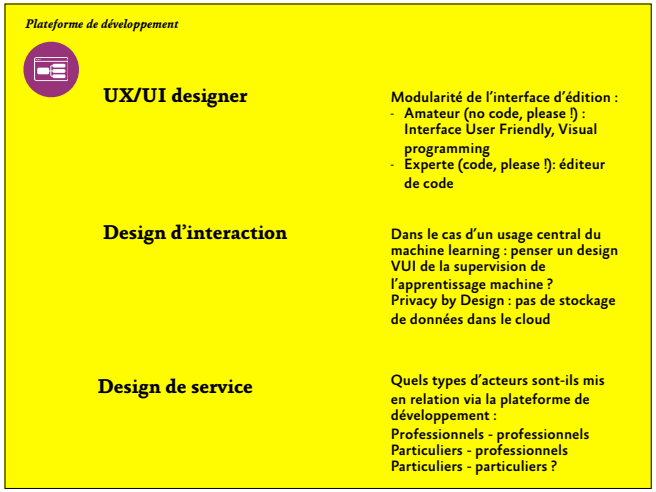
La résolution appartient à la participation des utilisateurs ou un principe de don/contredon du secret permet son affichage. De cet espace de conversation oscillant entre affichage et dissimulation, se crée la possibilité de la lecture.

---




24

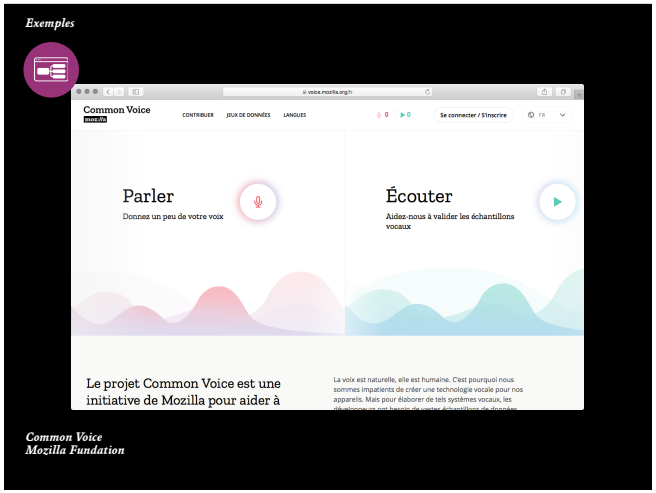
Les assistants vocaux sont très rarement envisagés comme comme des éditeurs de contenus. Particulièrement adaptés à l'écoute de podcasts, il serait pourtant intéressant de se poser la possibilité d'une édition (indexation) de contenus de références, d'entretiens, … avec la voix.




25

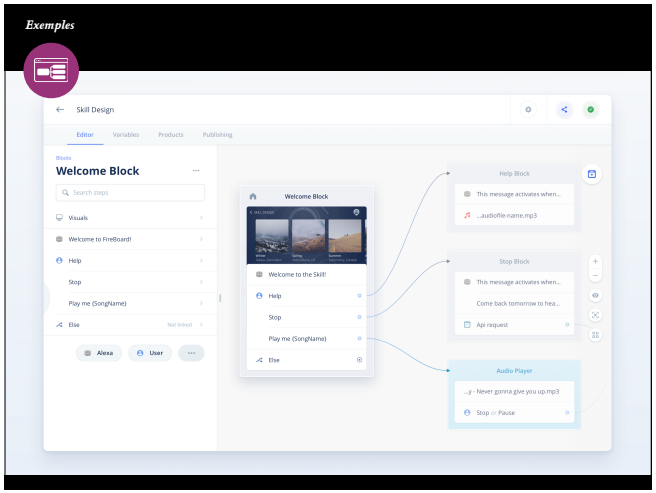
La plateforme de développement, dans la cartographie de l'assistant vocal, occupe une place centrale. Espace de la reconnaissance vocale, de développement des fonctions logiciels

---




26

Un élément central des plateformes, est la reconnaissance vocale (ASR - Automatic Speech Recognition) et de la compréhension des requêtes (NLU - Natural LAnguage Understanding). Le machine learning a particulièrement amélioré la capacité de transcription automatique de la voix. Alors que les GAFAM ont bénéficié de jeux de données publics (transcription et enregistrements audio de la commission européenne), ils ne mettent pas en libre accès les modèles de reconnaissance tirés de ces données massives, freinant l'existence de plateformes de reconnaissance vocale alternatives. Une initiative comme Common Voice de Mozilla, vise à corriger cet état de fait. Elle repose cependant sur la participation active et volontaires des internautes pour atteindre ses objectifs, qui, pour l'instant pour le français



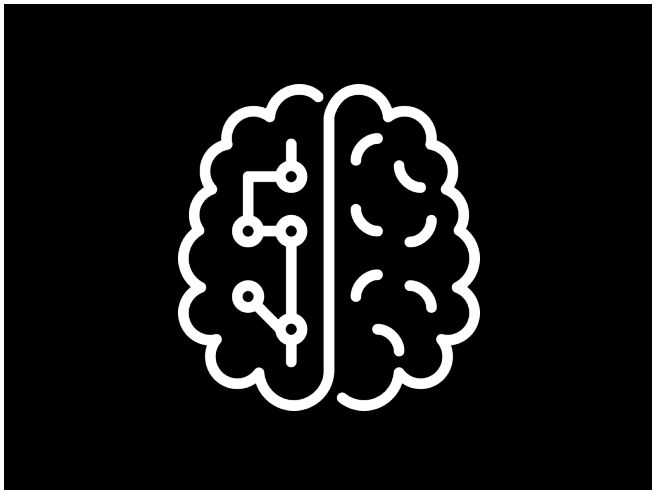

27

La participation des usagers est un élément central pour corriger la sur-puissance des GAFAM, que ce soit pour le traitement de la reconnaissance vocale ou dans sa capacité à expérimenter des usages. La mise en place d'interfaces permettant à l'utilisateur d'accéder aux différentes fonctionnalités de l'interface vocale quelquesoit son niveau d'expertise, sous la forme d'interfaces de programmation visuelles peut faciliter l'adhésion

---

28

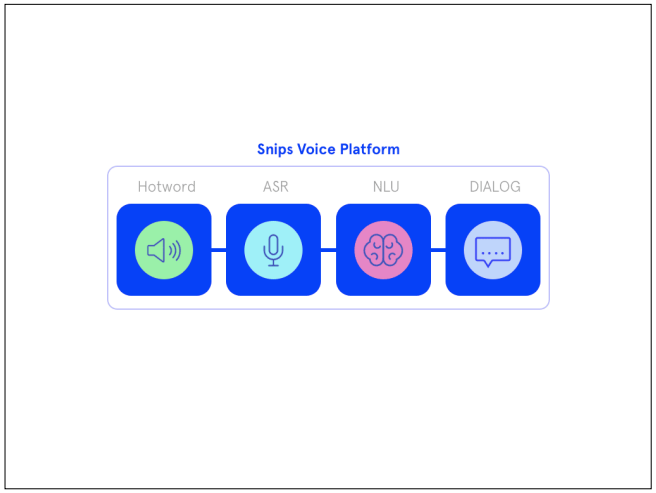
Un des enjeux au cœur de ces plateformes, en ce qui concerne le respect de la vie privée des utilisateurs, et de restreindre voir de supprimer le traitement des requêtes utilisateur par ce type de plateforme. Certains acteurs proposent des alternatives, centrant la conception de leur service sur la privacy by design, et le traitement des requêtes de l'utilisateur en local, sans passer par les serveurs du cloud.



29

La personnalisation passe-t-elle par l'automatisation du traitement d'un jeu de données non consentie de la part de l'utilisateur, son identification par un motif statistique, où passe-t-elle notamment par la participation de l'usager à son élaboration ? Plutôt que de construire des modèle de développement favorisant le digital labor et son automatisaton, quelles solutions s'offrent à la co-création ?

---




30

En dehors de la reconnaissance vocale, pour ce qui concerne la qualité des dialogues, il semble difficile de faire confiance à l'apprentissage automatique. L'apprentissage automatique est fait pour générer des stéréotypes, ce qui semble être incompatible avec la création d'interfaces vocales singulières. Le place de l'humain et de l'auteur est capitale dans l'élaboration de profil d'interface et de scripting inattendus et originaux, pour favoriser l'adhésion des utilisateurs.

31

Si les interfaces vocales proposées par les Gafam visent alternativement à renforcer un lien direct avec le consommateur, s'assurer de l'approvisionnement en données massives,  où préserver des parts de marchés, elles peuvent néanmoins, sous certaines conditions d'usages et sous certaines condition d'existences d'acteurs tiers, apporter un intérêt. Peu gourmandes en consommation électrique (pas d'écran), limitant la surcharge cognitive (par la capacité à ne traiter qu'une requête à la fois), par une simplicité d'usage a priori (plus d'utilisation d'interface tangible tactile) elles pourraient faciliter l'accès à des services numériques (« 100 % de services publics dématérialisés en 2022. »)  aux publics concernés par l'exclusion numérique. Cependant dans l'état actuel des usages émergents

32