

La révolution des assistants vocaux industriels passera par le cols-bleus !

André JOLY, SPIX industry – septembre 2022
andre.joly@spix-industry.com

Plusieurs articles annoncent l'arrivée de l'utilisation de la voix et de l'assistance vocale au travail dans les mois à venir. Les arguments mis en avant sont pertinents, mais insuffisants pour SPIX industry. La voix est utilisée dans l'industrie quand les fonctions proposées sont contextualisées et jugées utiles. C'est ce que nous allons voir !



L'assistant vocal Spix utilisé dans un atelier, avec un support visuel.

Un point sur les assistants vocaux grand public

Il y a quelques mois, les agences PwC et Bloomberg ont analysé l'usage des assistants vocaux grand public, afin de comprendre les motivations des utilisateurs et les tendances à venir. L'intérêt de comparer leurs études réside dans la prise en considération des assistants vocaux installés sur des smartphones et des enceintes connectées pour l'un, et uniquement sur des enceintes connectées pour l'autre. Les différences sont majeures et pertinentes pour notre analyse industrielle : dans un cas le dispositif matériel est mobile et propose un retour visuel, dans l'autre une interface purement vocale.

Les résultats de l'analyse de PwC en 2019¹ ne créent pas la surprise. Les utilisateurs jeunes adoptent la technologie rapidement. Les plus âgés sont plus lents, mais confiants dans leur utilisation future. L'utilisation de ressources d'assistance vocale reste principalement sur téléphone mobile (57%), l'utilisation des enceintes connectées arrive après (27%).

L'étude de PwC vient confirmer l'utilisation attendue de cette technologie dans un contexte privé pour des tâches simples. Dans l'ordre pour les 5 premières :

- Faire une recherche sur un moteur de recherche
- Poser une question rapide
- Vérifier la météo ou les dernières nouvelles
- Jouer de la musique
- Lancer un minuteur

L'étude de Bloomberg² publiée en 2021 (réalisée 1 an après celle de PwC) nous apprend que les utilisateurs de l'assistant vocal Alexa d'Amazon exploitent principalement le dispositif équipé d'un écran, donc avec un retour visuel. Ces dispositifs sont néanmoins fixes, la majorité du temps dans le contexte du domicile. Les utilisateurs de Alexa à la maison utilisent dans l'ordre :

- Jouer de la musique
- Lancer un minuteur
- Allumer ou éteindre des lumières

¹ <https://www.pwc.com/cisvoiceassistants>

² <https://www.bloomberg.com/news/articles/2021-12-22/amazon-s-voice-controlled-smart-speaker-alexa-can-t-hold-customer-interest-docs>

En parallèle, l'éditeur lui-même sur le site promotionnel d'Amazon-Alexa³ propose l'utilisation de l'assistant vocal pour les fonctions suivantes : « *Alexa est toujours prête à répondre à vos questions, vous raconter des blagues, jouer de la musique, trouver le trajet le plus rapide pour aller au travail, ajuster le thermostat grâce à vos appareils connectés compatibles, et bien plus* ». Répondre aux questions, se distraire, jouer de musique : la proposition est cohérente de l'utilisation constatée par PWC et Bloomberg.

Mais alors à quoi pourrait bien servir un assistant vocal au travail ? Effectivement, les fonctions citées précédemment ne donnent pas très envie à un industriel d'investir sur cette technologie.

³ <https://www.amazon.fr/>

Une première réponse généraliste pour l'entreprise

Une étude réalisée par le groupe Manutan⁴ en 2022 présente les utilisations possibles des assistants vocaux au travail, ainsi que les gains attendus. Parmi ces fonctions, les suivantes sont invoquées comme « *champs d'application des assistants vocaux en entreprise* » :

- La dictée vocale qui est l'élément le plus utilisé;
- La prise de notes (compte-rendu de réunion);
- Le travail d'équipe;
- La gestion du calendrier des employés;
- Les services clients;
- Les services d'assistance informatique;
- La recherche documentaire : il est possible de demander à l'assistant vocal d'afficher des documents à l'écran lors de présentations (bilans financiers, plaquettes de présentation, etc.).

En premier lieu, il est intéressant de constater que la distinction entre reconnaissance vocale et assistance vocale n'est pas très claire... Effectivement, des fonctions de « dictée vocale » ou de « prise de notes » sont liées à la mise en œuvre de capacités de transcription de la voix en texte (du *speech to text* en bon anglais), mais ne sollicitent pas de capacités de compréhension du texte ou de dialogues avancés.

Les autres fonctions sont effectivement liées à la mise en œuvre de véritables fonctions d'assistance vocale. C'est-à-dire qu'une intelligence artificielle doit comprendre l'intention de l'utilisateur à partir de la transcription en texte de sa demande vocale, puis y répondre soit directement, soit à la suite de tours de parole pour raffiner la demande.

On peut néanmoins constater un mélange entre les fonctions externes et les fonctions internes. Les fonctions externes comme le « service client », mettent en œuvre des technologies proches de celles « grand public » citées par les études précédentes. Elles posent des questions classiques sur la gestion des données personnelles, mais ne suscitent aucun doute quant à leur rentabilité. Il paraît évident que le remplacement d'un opérateur humain par un call-bot, pour des tâches de support de premier niveau est un choix économique rentable. Est-ce socialement responsable... c'est une autre question !

Les autres fonctions comme le « calendrier », les « services informatiques », ou la « recherche documentaire » sont effectivement internes. Ces fonctions sont décrites « comme un atout ergonomique » par les auteurs. Cela voudrait dire que l'accès vocal à ces fonctions n'est pas

⁴ <https://www.manutan.fr/blog/conseils/assistants-vocaux-revolution-methodes-travail>

critique, ou n'est pas indispensable. Effectivement, les fonctions citées comme la « prise de rendez-vous », la « réservation de salle » ou la « programmation d'un rendez-vous client » sont déjà accessibles par les outils bureautiques classiques. Les personnes qui les utilisent sont majoritairement dans un bureau, devant un poste informatique. Est-ce vraiment indispensable et plus rapide de définir un rendez-vous client par la voix plutôt qu'en trois clics sur une interface classique ? Mis à part pour des personnes qui présentent des situations de handicap, le retour sur investissement sera-t-il à la hauteur des dépenses de mise en œuvre.

Pour SPIX industry⁵, il est indispensable que la mise en œuvre de l'utilisation de la voix et de l'assistance vocale apporte une réelle plus-value d'une part à son utilisateur, et à l'industriel qui finance tout ça. De la même manière, il est important de considérer dans l'entreprise l'ensemble des personnels qui pourraient bénéficier de solutions vocales ou d'assistance vocale.

La plus-value de l'utilisation de la voix et de l'assistance vocale ne serait-elle pas plutôt à trouver dans les activités secondaires, plutôt quand dans le tertiaire ? C'est la vraie révolution de l'arrivée des assistants vocaux dans l'industrie : investir sur la voix des clos-bleus pour maximiser la compétitivité de l'industrie du futur.

⁵ <https://www.spix-industry.com>

Les limites des assistants vocaux domestiques

Le premier bilan tiré des études précédentes est donc sans appel : les assistants vocaux grand-public sont donc utilisés principalement pour des tâches très simples, et majoritairement sur smartphone avec un retour visuel.

Qu'en est-il de l'adoption et de l'utilisation récurrente des enceintes connectées, et des retours de leurs utilisateurs. La seconde partie de l'étude de l'agence Bloomberg⁶ est beaucoup plus inattendue et donne un éclairage pertinent pour une utilisation industrielle.

Cette deuxième partie de l'étude montre que les principales fonctions d'Amazon-Alexa sont comprises par leurs utilisateurs dans les 3 heures suivant leur installation. Par la suite, 25% d'entre eux délaissent ces dispositifs dans les 2 semaines suivantes... ça laisse rêveur et en même temps ! Les raisons invoquées pour expliquer cette constatation les suivantes :

- **Les fonctions accessibles et utilisées sont jugées peu utiles**

Les fonctions utilisées mises en évidence par les études précédentes ne sont pas révolutionnaires par rapport à l'utilisation tactile classique d'une tablette ou d'un smartphone. La plus-value de ces fonctions reste du domaine soit du jeu, soit du gadget, ou encore du confort. L'étude PwC⁷ montre que des fonctions à beaucoup plus forte plus-value (réservation d'un restaurant ou de billet d'avion par exemple) ne sont pas encore utilisées (16% et 0% respectivement), mais sont largement souhaitées pour l'avenir (32% et 26% respectivement). Les utilisateurs cherchent donc à simplifier des tâches complexes à forte plus-value par l'utilisation de la voix et de l'assistance vocale.

Au regard de l'étude Manutan, on peut légitimement se demander si l'utilisation vocale des fonctions bureautiques classiques ne va pas tomber dans cette catégorie. Une exemption : une dictée vocale efficace pour l'industrie, qui comprend les termes métiers et connaît les noms des salariés.

⁶ <https://www.bloomberg.com/news/articles/2021-12-22/amazon-s-voice-controlled-smart-speaker-alexa-cant-hold-customer-interest-docs>

⁷ <https://www.pwc.com/cisvoiceassistants>

- **Les utilisateurs sont concernés par la confidentialité de leurs données**

Plusieurs études récentes rapportées par les médias^{8,9} démontrent que les capacités des assistants vocaux grand public comme Alexa, Google-Home ou Siri sont améliorées par l'analyse des échanges entre les utilisateurs et leur assistant. Cette analyse passe obligatoirement par l'écoute des commandes vocales énoncées par les utilisateurs pour obtenir le service demandé. Le doute sur la confidentialité des données est donc semé dans l'esprit des utilisateurs.

Dans le contexte industriel, les questions sont les mêmes. Les call-bots de service et d'assistance de premier niveau pour les clients devront donc rapidement clarifier leur position quant au sujet des données manipulées. De même pour les solutions internes : est-ce qu'elles écoutent en permanence, est-ce que la voix est enregistrée, ...

- **Les suggestions spontanées faites par Alexa énervent les utilisateurs**

Enfin, le rapport Bloomberg remonte un constat sans appel : l'utilisateur accepte mal que son assistant vocal lui donne des ordres, ou lui fasse des suggestions d'actions ou d'achat de manière non sollicitée. Premièrement, cela prouve que l'assistant exploite les requêtes précédentes de l'utilisateur ce qui vient renforcer sa méfiance vis-à-vis de ses données personnelles. Enfin, les utilisateurs ont du mal à ce que leur assistant vocal soit à l'initiative de l'échange : l'humain veut garder l'initiative.

Et pourtant, on pourrait imaginer des fonctions vocale associées à des situations critiques, d'alerte incendie, de risque climatique, ou de danger imminent. Ce type de fonction de sollicitation spontanées seraient beaucoup mieux acceptées. Nous verrons que pour que ces fonctions soient efficaces, il convient que l'assistant vocal soit « aware », c'est-à-dire qu'il soit conscient du contexte, de l'environnement dans lequel il est utilisé.

Ainsi, les fonctions disponibles dans les assistants « grand public » sont jugées peu utiles par leurs utilisateurs. Effectivement, pour être « intelligent » et donc pertinent, un assistant vocal doit comprendre le contexte de son utilisation. Dans le domaine du grand public, l'utilisateur (à juste titre) ne souhaite pas partager trop de contexte avec son assistant (*où je suis, ce que je fais, ce que j'aime, mon humeur, ...*). Les fonctions accessibles sont donc nécessairement simples.

La révolution de l'assistance vocale dans l'industrie consiste à l'utiliser pour des tâches critiques (les outils bureautiques ne sont pas utilisables de manière classique), et pour lesquelles un contexte est disponible (un ordre de tâche, un bon de travail, un formulaire à compléter). Il est donc temps d'allumer la voix dans les ateliers de production, de maintenance ou de contrôle qualité de l'industrie.

⁸ <https://www.cnetfrance.fr/>

⁹ <https://www.lefigaro.fr/secteur/high-tech/>

Alors pourquoi les cols-bleus ?

Pour qu'un assistant vocal soit considéré comme « utile », il faut qu'il rende un « service à forte valeur ajoutée » à l'utilisateur et qu'il apporte un « retour sur investissement » (un bénéfice) à celui qui l'a acheté (le payeur).

Comme nous l'avons déjà vu dans un précédent article¹⁰, pour les assistants vocaux grand public le « payeur » est également le « fournisseur » de la technologie. Les assistants vocaux Alexa, Siri ou Google-Home sont effectivement mis à disposition gratuitement des utilisateurs, les contreparties non financières étant largement sous-estimées. Le modèle économique de ces assistants porte donc de manière intrinsèque une partie de la frustration des utilisateurs.

Tel que précisé dans une définition de Wikipedia¹¹, l'intelligence d'un assistant vocal (ou assistant personnel) repose sur la connaissance de l'utilisateur et l'historique des données, autrement dit des bases de connaissances. Ces connaissances dans un contexte grand public sont disponibles en grande quantité, mais obligatoirement imparfaites. Il en résulte donc des services rendus à l'utilisateur qui sont simples et sans réelle valeur ajoutée.

Nous pourrions en conclure que la situation est désespérée... Heureusement non ! Il suffit de trouver des domaines d'application pour des technologies d'assistance vocale qui permettent de remplir les deux conditions : « valeur » pour l'utilisateur et « bénéfice » pour le payeur.

La mise en œuvre de solutions vocales pour les cols-bleus de l'industrie répond aux critères de valeur et de bénéfice, et apporte l'ensemble des ROI¹² attendus.

Depuis plusieurs années, SPIX industry¹³ développe une technologie d'assistance vocale intelligente au service de l'industrie, et plus particulièrement des techniciens et les cols-bleus. Pourquoi cet assistant remplit les deux conditions précédemment énoncées ?

Dans le contexte actuel de développement de l'industrie 5.0¹⁴, les opérateurs manuels sont de plus en plus confrontés à des tâches numériques. La digitalisation des tâches, le déploiement de logiciels complexes de gestion des procédures, des interventions, des rapports... demande aux hommes et aux femmes de terrain de l'industrie d'utiliser des applications numériques sur

¹⁰ <https://www.spix-industry.com/2021/assistants-vocaux-quel-est-prix-du-gratuit/>

¹¹ https://fr.wikipedia.org/wiki/Assistant_personnel_intelligent

¹² <https://www.spix-industry.com/article-roi-innovations-industrie/>

¹³ <https://www.spix-industry.com>

¹⁴ <https://www.spix-industry.com/2021/quest-ce-que-industrie-50-peut-faire-pour-vous>

smartphone, ou sur tablette. Problème : ce n'est pas leur métier, ils ont souvent des gants, et n'aiment pas ce type de tâche sur lesquelles ils « perdent » beaucoup de temps.

L'utilisation d'un assistant vocal en interface des logiciels qu'ils utilisent déjà sur tablette ou smartphone, leur simplifie les tâches numériques. Ils réalisent les actions digitales qui leur sont demandées, mais de manière plus simple et en gardant les mains et les yeux libres. Ils conservent le support visuel de leur logiciel habituel, ce qui leur permet de monter en maturité dans l'utilisation de la voix. Un jour, ils pourront utiliser leur logiciel sans regarder l'écran !

Dans ce cas, la valeur ajoutée et le service rendu à l'utilisateur sont élevés, car l'assistant vocal dispose de l'ensemble des connaissances nécessaires à son fonctionnement. En effet, les procédures, ordre de tâche, instructions de travail ou formulaires de rapports, complétés par des bases de vocabulaire et d'ontologies propres à l'industrie, constituent l'ensemble des connaissances utilisées pour définir le contexte dans lequel l'utilisateur fait appel à son assistant vocal.

Les utilisateurs ont confiance dans cet assistant vocal intelligent Spix, car les règles d'utilisation de leur voix et des données de l'entreprise sont claires. En effet, les SKILLS de Spix¹⁵ permettent d'assembler un assistant vocal intelligent embarqué, sans connexion à un cloud externe et qui ne nécessite pas d'accès à des données d'entreprise en temps réel. Il fonctionne en interface des logiciels métiers des opérateurs, ces logiciels restent en charge de la gestion des données métiers qui leur sont propres. Du côté opérateur, leur voix n'est pas enregistrée, ils peuvent parler à leurs collègues sans interférences avec leur assistant, et le contexte d'utilisation est limité à leur contexte de travail. L'opérateur reste toujours à l'initiative de l'interaction avec son assistant, sauf en cas d'alerte de sécurité. Mais dans ce cas, la plus-value de l'intervention de l'assistant est évidente.

Du point de vue de l'industrie, l'assistant vocal Spix de SPIX industry est un logiciel commercial. La technologie développée par l'entreprise respecte les règles de confidentialité des données des industriels, s'interface avec leurs logiciels métiers par des librairies¹⁶ et répond aux contraintes de validation pour une utilisation opérationnelle en milieu industriel. Ainsi, l'industrie fait l'acquisition de cet assistant vocal pour permettre à ses techniciens et opérateurs de terrain de mieux utiliser leurs applications numériques.

Dans ce cas, l'industriel est le payeur. Son retour sur investissement¹⁷ se trouve dans la réduction de la non-qualité de production, la réduction des temps à non-valeur ajoutée de ses opérateurs, l'augmentation de la sécurité des interventions humaines. Enfin, l'industriel trouve son bénéfice dans l'augmentation des données qu'il est capable de collecter en temps réel sur les opérations de terrain effectuées par ses techniciens.

¹⁵ <https://www.spix-industry.com/service-vocal-intelligent-digitalisation-industrie/>

¹⁶ <https://www.spix-industry.com/service-vocal-intelligent-digitalisation-industrie/>

¹⁷ <https://www.spix-industry.com/article-roi-innovations-industrie/>

Conclusion

L'adoption de l'utilisation de la voix, pour nos besoins privés comme pour une utilisation industrielle, passera par l'émergence de services « *à forte valeur ajoutée* » pour les utilisateurs. Dans le grand public, c'est compliqué, car tout repose sur les données que nous sommes tous disposés à partager.

Pour l'industrie, il ne faut pas se tromper de cible. Piloter un logiciel de bureau par la voix alors que la souris et le clavier sont disponibles et utilisable, n'apporte pas une réelle plus-value. Comme nous l'avons vu, seule l'introduction de serveurs internes et confidentiels pour une transcription vocale capable de traiter des termes métiers apporte un réel service à valeur ajoutée aux utilisateurs bureautiques.

La révolution de l'utilisation de la voix passera donc par les techniciens et les cols-bleus qui ont les mains occupées à la réalisation de leurs tâches et à qui l'on demande de plus en plus d'interactions numériques.

Chez SPIX industry, nous sommes convaincus par le développement de solutions d'assistance vocale intelligente sur des verticaux métiers. Sur des métiers de production, d'inspection, de contrôle qualité ou de maintenance, il est possible d'assembler les composants nécessaires pour assurer l'utilité, l'utilisabilité et l'acceptabilité des services vocaux rendus à l'utilisateur, tout en satisfaisant le retour sur l'investissement de l'industriel.

A propos de SPIX industry

L'entreprise a été créée pour développer l'utilisation de la VOIX dans l'industrie sous toutes ses formes : pilotage vocal de logiciels existants, assistance conversationnelle avancée pour guider les opérateurs sur leurs procédures, remplissage par la voix de formulaire « yeux et mains libres », transcription temps réel ou différé de notes audio, co-développement d'une application métier intégrant un Assistant Vocal Intelligent.

Entreprise résolument innovante, totalement engagée dans la mise en œuvre opérationnelle de technologies issues de l'Intelligence Artificielle au bénéfice du développement de l'industrie 4.0 et de l'opérateur augmenté, SPIX industry est leader sur son marché en Europe.

Depuis 2013, SPIX industry conduit un effort de R&D important pour développer une technologie 100% française d'Assistance Vocale Intelligente « Spix » et bâtir des solutions opérationnelles dans un environnement industriel exigeant.

Site web : spix-industry.com

Linkedin : linkedin.com/company/spix-industry/

Contact presse

André JOLY – Directeur Général

Tél. : +33 6 25 17 27 94

Mail : andre.joly@spix-industry.com

Entité légale

Simsoft3D SAS - 1244 rue l'Occitane - 31670 Labège (France)

« SPIX » et « SPIX industry » sont des marques déposées de Simsoft3D SAS.