

Table des matières

Table des matières	1
1 Assistants virtuels intelligents	3
1.1 Introduction	3
1.2 Un robot à tout faire	5
1.3 Architecture générale d'un AVI	5
1.4 Caractéristiques principales d'un AVI	6
1.4.1 Sensible au contexte	6
1.4.2 Évolutif	7
1.4.3 Multimodal	7
1.4.4 Anthropomorphe	8
1.4.5 Multi plateforme et Flexible	8
1.5 Domaines d'applications des AVIs	8
1.5.1 Vie quotidienne	8
1.5.2 Assistance professionnel	9
1.5.3 E-enseignement	9
1.6 Types d'assistants virtuel	9
1.6.1 ChatBot(Agent de conversation)	9
1.6.2 Virtual Administrative Assistant	9
1.6.3 Social Media Marketing Virtual Assistant	9
1.6.4 Virtual Assistant Writers	9
1.6.5 Virtual Research Assistant	9
1.6.6 eCommerce Virtual Assistant	9
1.6.7 Data Entry Virtual Assistant	9
1.7 Exemples d'AVI	10
1.8 Points fort des AVI's	16
1.9 Conclusion	16
Bibliographie	18

INTRODUCTION GÉNÉRALE

- Ici on parlera des motivations qui ont aboutis à ce projet, des objectifs de ce dernier ainsi que ses perspectives

1.1 Introduction

Depuis la commercialisation du premier ordinateur grand public(**Xerox PARC Alto**) en 1973, le monde découvris pour la première fois ce qui allait devenir l'apparence basique de chaque ordinateur moderne, en effet ils furent les premiers à proposer une interface graphique dotée de fenêtres, d'icônes et d'uns souris pour se déplacer et d'un clavier pour écrire du texte. Bien que basique, cette idée lança alors plusieurs autres grandes marque sur le même chemin (IBM,Apple,Compaq ...), Par la suite beaucoup ont essayé d'améliorer la façon dont l'homme utilisait sa machine : souris plus précise, écran doté d'une plus grande résolution, clavier plus enrichis, voir même l'introduction des écrans tactiles dans certains systèmes embarqués.

Cependant, certains voyaient encore cette façon d'utiliser la machine comme **trop primitive**, et peu intuitive, en effet laisser un enfant devant un ordinateur et il prendrait un bon moment pour apprendre à ne serait ce qu'éditer un simple fichier. pour citer Donald A. Norman

“We must design for the way people behave, not for how we would wish them to behave.”[1]

Traduit par :

“Nous devons concevoir selon le comportement des utilisateurs, et non pas selon la façon dont nous voudrions qu'ils se comportent.”

L'humanité à fait beaucoup de chemin depuis les années 70, utiliser un ordinateur de nos jours avec les moyens **classiques** (souris, clavier, écran ...) est devenu une tâche triviale, voir même une seconde nature, ça reste cependant dû au fait que de plus en plus de jeunes enfants sont exposés depuis leur plus jeune âge au monde technologique qui les entoure, le processus d'apprentissage reste cependant présent, l'effort d'utiliser les outils communs reste lui aussi présent.

Qu'on le veuille ou pas, la plus naturelle et plus ancienne façon de communiquer pour l'homme à toujours été la parole, l'invention de langues toutes aussi riches et complexes les unes que les autres à permis à l'humanité de briser plusieurs barrières sociales. l'avancement le plus naturel pour cette façon de communiquer serait donc de l'étendre non plus seulement qu'aux humains, mais aussi aux machines que l'homme ç su construire et améliorer au fil des années.

Motivé par cette manière que l'on a de communiquer entre nous, et épaulé par les récentes technologies tel que l'apprentissage automatique?? et le traitement automatique du langage naturel?? et de l'intelligence artificielle en tant que tout un domaine, les plus brillants des chercheurs ont entamé leurs recherches dans un tout nouveau domaine.

Les Assistants Virtuels Intelligents sont donc le produit de plusieurs années de recherche, visant tout d'abord à faciliter certaines tâches pour l'utilisateur, les premiers AVI était conçu comme des agents de conversation ou **Chatbots**, limités dans leurs actions et dépendant toujours d'un moyen de communication textuel, ce ne fut pas la forme désirée de l'AVI par excellence, avec l'avancement des recherches sur la reconnaissance automatique de la parole et l'émergence de l'apprentissage automatique (les tout premiers assistant virtuels utilisant la RAP étant spécialisés dans certains domaines comme des système d'aides à la décision médicaux), il a ensuite été plus aisé de briser la barrière et de réaliser ce qui était encore une esquisse d'un AVI personnalisé avec l'émergence de l'apprentissage profond(??) et la popularisation des Smartphones¹, de nouveaux AVIs comme **Apple Siri??** en 2011 et **Google assistant** et **Amazon Alexa** ont fait leur apparitions, offrant de plus en plus de services personnalisés et spécifiques à chaque utilisateurs. Dans la suite de ce chapitre nous essayerons de mieux détailler ce qu'est un AVI, ce qui est demandé d'un tel système, des différents types d'AVIs et des domaines d'application de ces derniers, pour enfin conclure sur les limitations actuelles et les motivations de ce projet.

1. Text about smartphones here

1.2 Un robot à tout faire

Informellement, un AVI est type d'**agent**² logiciel qui peut effectuer certaines tâches et proposer des services dédiés aux utilisateurs qui vont d'une simple tâche (Ouvrir une fenêtre, lancer une application ...) à la réalisation de requêtes un peu plus complexes (). Pour répondre efficacement à toutes sortes de requêtes, un AVI se doit donc de garder trace du contexte courant de sa conversation avec l'utilisateur, il doit donc disposer d'un système capable d'enregistrer les informations pertinentes et de savoir les réutiliser, mais aussi de pouvoir déduire les quelles de ces informations sont manquantes, on parle ici de Context-Awareness ou Sensibilité au contexte, comme vu dans [2].

toujours d'après [2] et [3] *Day* et *Abwod* définissent un contexte comme suit :

“A context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves”

Qui peut être traduit par :

“Un contexte est une information qui peut être utiliser pour caractériser l'état d'une entité. Un entité peut être une personne une place ou un objet considéré comme pertinent vis à vis de l'interaction entre l'utilisateur et l'application, ainsi que ces deux derniers en eux mêmes”

Il en découle que pour parvenir à développer un système qui puisse répondre aux besoins individuels et spécifiques de chaque personne, parvenir à modéliser et prendre en compte le contexte semble être une solution prometteuse.

1.3 Architecture générale d'un AVI

Il n'existe pas vraiment d'architecture type pour un AVIs, chaque équipes de chercheurs à tenté sa propre implémentation, il existe en revanche une représentation haut-niveau des composants d'un tel système, nous proposons à titre d'exemple la pseudo-architecture interne de l'assistant de Google :

2. un agent est ...

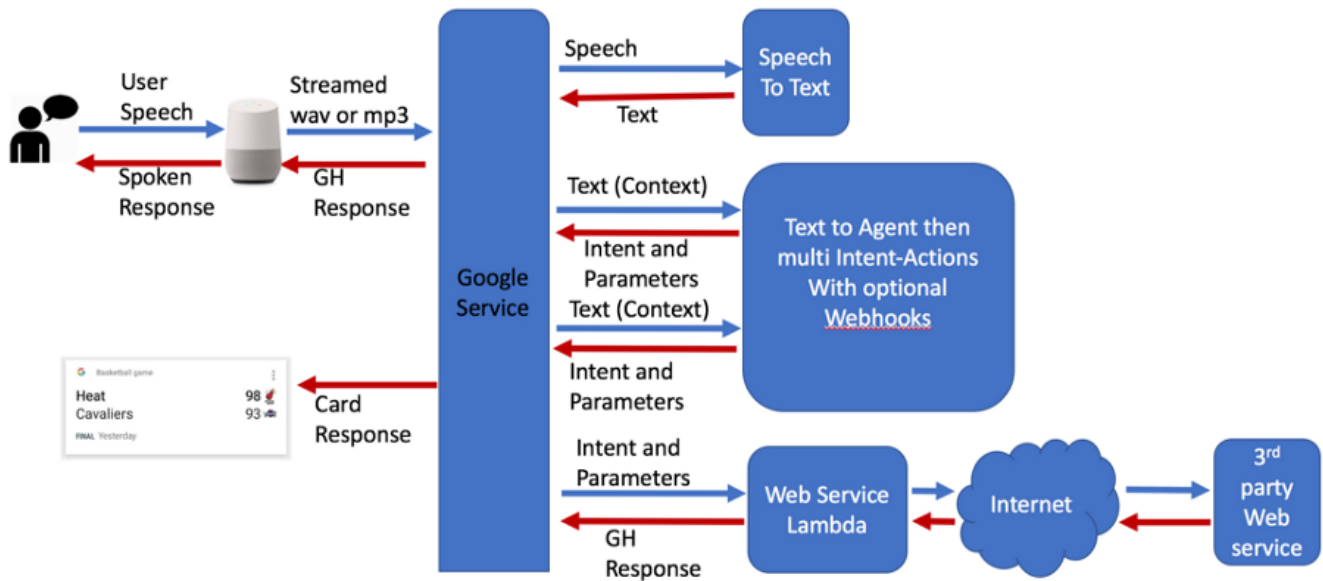
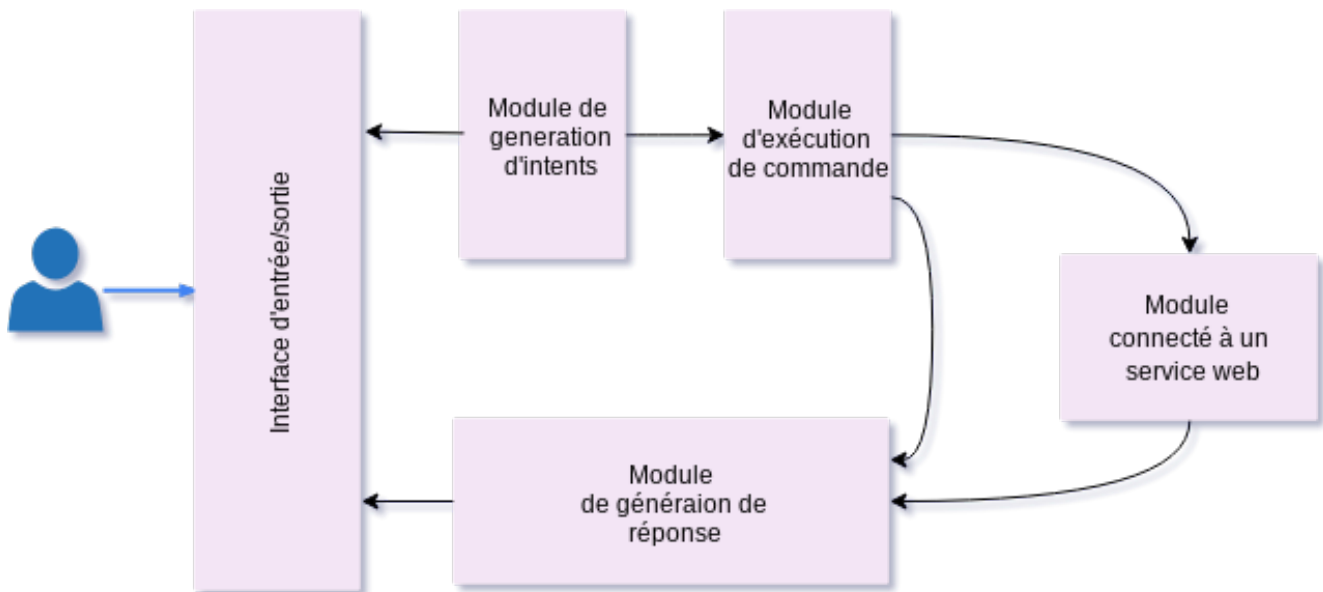


FIGURE 1.1 – Architecture haut niveau de Google Assistant [4]

On peut ainsi dégager le schéma simplifié abstrait suivant :



1.4 Caractéristiques principales d'un AVI

1.4.1 Sensible au contexte

Comme précédemment vu dans la définition du contexte (voir 1.2), ce dernier peut être interprété comme tout aspect d'une entité (Position d'un objet, couleur d'un objet, température d'une chambre ..etc), Un assistant dit **intelligent** doit donc être capable de capturer le concept du contexte, d'utiliser et de traiter toute informations catégorisée comme contextuelle. Pour être plus

précis, l'AVI doit être sensible à *l'évolution* du contexte courant, par le biais de capteur optiques de microphones .. etc, tout ce qui pourrait amener l'utilisateur à faire évoluer la demande qu'il a émis, l'assistant devra donc proposer un système de mise à jours du contexte pour éliminer les informations inutiles et garder celles qui pourraient aider à répondre à la requête de l'utilisateur.

1.4.2 Évolutif

Il a été vu dans 1.2 qu'un AVI peut être vu comme un type d'agent. Pour rappel d'après *Russel* et *Norvig* un agent (qui est une entité autonome pouvant interagir avec son environnement afin d'accomplir certaines tâches) peut être de plusieurs types :

- Agent simple réflexe : agent exécutant ses actions à base de règle conditionnelles simple (c.à.d **Si Condition alors exécuter actions**), ils sont ainsi très simplistes et limité dans la portée de leur actions.
- Agent basé modèle : semblables aux agents simple réflexe, il sont toute fois doté d'un modèle interne complexe sensé représenter le monde extérieur auquel l'agent à accès, cependant ils appliquent les actions de la même manière que le précédent type d'agents.
- Agent à but : une amélioration des agents simple qui sont dotés d'un ensemble d'états buts à atteindre d'une façon ou d'une autre
- Agent à utilité : des agents à buts qui tentent d'aboutir à leurs buts d'une manière optimisée (intelligente) utilisant une fonction de mesure adéquate pour le choix des différents états à atteindre.
- Agent apprenant : agent à utilité enrichi par un module d'apprentissage qui sert de juge pour répondre aux **"critiques"** des actions qu'il entreprenant, le terme **agent évolutif** est aussi employé.

Pour ce qu'il en est des AVIs, les plus récents systèmes (ex : **Amazon Alexa** qui améliore son module de reconnaissance de la parole après chaque réponse non déclinée par l'utilisateur), peuvent être considérés comme des agent apprenants, répondant de ce fait à la contrainte évolutive imposée, cependant le domaine de l'auto-évolution des systèmes intelligents est encore un domaine nouveau qui se voit épaulé par les récentes avancées dans le domaine de l'apprentissage automatique [2]

1.4.3 Multimodal

Afin d'assurer une aisance d'utilisation, les AVI's sont fréquemment amenés à récupérer les requêtes(ou données) en entrée de la manière la plus naturelle possible(par exemple par le biais de la paroles). Cependant, pour garantir une expérience d'utilisation adéquate, l'assistant sera souvent confronté à récupérer ses requêtes de différentes manières, que ce soit à travers une interface graphique(écran tactile) ou à travers d'un texte brut tapé au clavier voir même à travers des expressions faciales ou des états cognitives/émotionnels ([5]), pour ensuite produire une réponse qui, elle aussi pourrait éventuellement être de la forme textuelle ou sonore ou les deux ..., cette capacité à recevoir en entrée et/ou produire une sortie de plusieurs façon différentes est la multi-modalité d'après [6], et cette caractéristique qui permet de masquer à l'utilisateur toute la complexité d'acquisition de ses requêtes.

1.4.4 Anthropomorphe

Plusieurs auteurs tendent à attribuer une grande importance à l'anthropomorphisme des AVIs ([7]), qui est

“Un mécanisme qui pousse les êtres humains à induire qu’une entité non-humanoïde possède des caractéristiques et comportements propres à l’homme”[8]

Ce comportement humanoïde pousserait donc l'utilisateur à se sentir plus à l'aise avec l'assistant, le conduisant ainsi à adopter une façon de communiquer plus humaine et moins structuré qu'avec les autres machines, c'est ainsi une caractéristique majeure d'un AVIs se disant **personnalisé**.

1.4.5 Multi plateforme et Flexible

Malgré leur récentes prouesses, certains AVIs sont encore restreints à un écosystème fortement dépendant du fabricant, Cowan et al. mentionnent dans [9] que Apple Siri est limité à l'environnement constitué des produits de la firme à la pomme, n'ouvrant par défaut que les applications de cette dernière quand une requête lui est transmise.

C'est un comportement que les assistants devrait éviter, car une indépendance des plateformes utilisées est, certes, très complexe à instaurer, mais offre plus de possibilités aux utilisateurs et aux développeurs, pouvant ainsi exploiter la puissance de certaines plateformes (Smartphones, TV connectées ...). Avec l'émergence de l'IoT (Internet of Things) et des maisons intelligentes par exemple, c'est un tout nouveau terrain de jeu qui est présenté pour les AVIs, offrant plus d'opportunités pour les utilisateurs.

1.5 Domaines d'applications des AVIs

Après avoir vu les différents aspects que les AVIs doivent traiter, nous nous intéresserons maintenant aux types de services et applications que ces derniers pourraient fournir pour démontrer qu'ils peuvent bel et bien faciliter certaines tâches à l'homme.

1.5.1 Vie quotidienne

À la base, les AVIs étaient destinés à un usage très personnel comme la gestion des achats dans les supermarchés, ou des guides touristiques de plusieurs destinations de voyage, cette spécificité a commencé à s'estomper petit à petit avec l'émergence de nouveaux systèmes dédiés à des applications plus générales, comme les maisons intelligentes ou les assistants de planifications de tâches, ce qui a permis de mettre encore plus l'accent sur cet aspect de convivialité que les tout premiers AVIs en tenter de perfectionner, ainsi tout ces assistants spécialisés dans des domaines restreints (Tourisme, shopping, détente ...) sont regroupés dans un seul système plus polyvalent capable de répondre à des besoins quotidiens divers et variés, allant même à fournir une assistance aux personnes âgées pour leur faciliter les tâches rudimentaires devenues trop fatigantes

1.5.2 Assistance professionnel

1.5.3 E-enseignement

1.6 Types d'assistants virtuel

1.6.1 ChatBot(Agent de conversation)

1.6.2 Virtual Administrative Assistant

1.6.3 Social Media Marketing Virtual Assistant

1.6.4 Virtual Assistant Writers

1.6.5 Virtual Research Assistant

1.6.6 eCommerce Virtual Assistant

1.6.7 Data Entry Virtual Assistant

'interesring links : ' [first](<https://www.acelerartech.com/va-guide/types-of-virtual-assistant-services/>) [second](<http://outsourceworkers.com.au/different-kinds-of-virtual-assistants/>)

1.7 Exemples d'AVI

Pour illustrer la puissance des AVIs les plus récents, nous avons décidé de mettre en évidence les quatre produits qui dominent le marché courant :

Google assistant

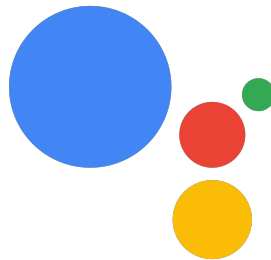


FIGURE 1.2 – Logo de Google assistant

Lancé en 2016 comme un chatbot(1.6.1)intégré dans l'application Google Allo, G.A s'est vu ensuite être directement intégré sur les système d'exploitation Android(que ce soit sur smartphones ou tablettes, et plus récemment sur Google Home³), G.A est un assistant à tout faire concocter par les ingénieurs de Google dans le but de faciliter la recherche sur internet, la planification des tâches, l'ajustement des réglages de l'appareil ..., sont point fort est sa capacité à engager une conversation bi-directionnel avec l'utilisateur, assurant ainsi une interaction personnalisé variant d'un utilisateur à un autre, cela lui permet par exemple de proposer certains résultats de recherche selon les les précédentes interactions avec l'utilisateur, ou de lui proposer une activité si ce dernier lui reproche de s'ennuyer.

3. Appareil servant à contrôler les composant d'une smart-house ainsi que l'utilisation des différents services de google

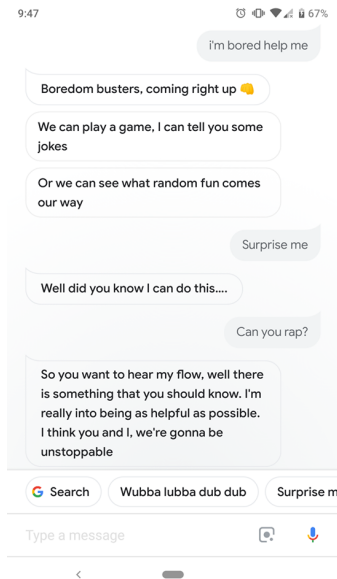


FIGURE 1.3 – Conversation aléatoire 1

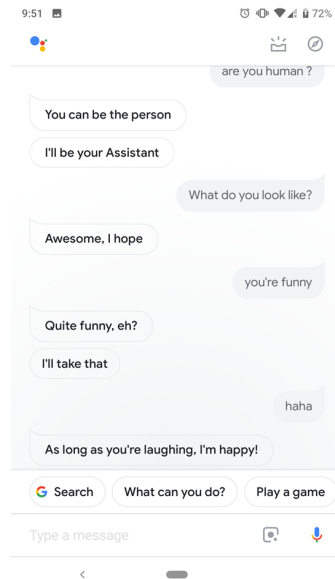


FIGURE 1.4 – Conversation aléatoire 2

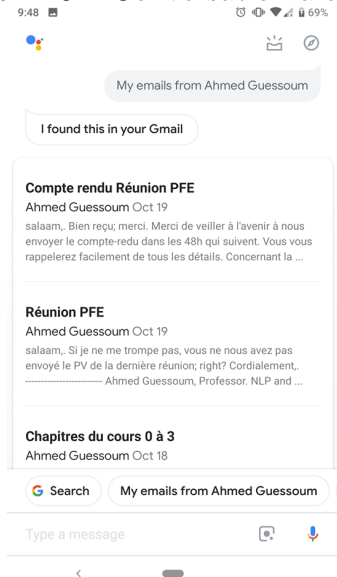


FIGURE 1.5 – Requête simple 1

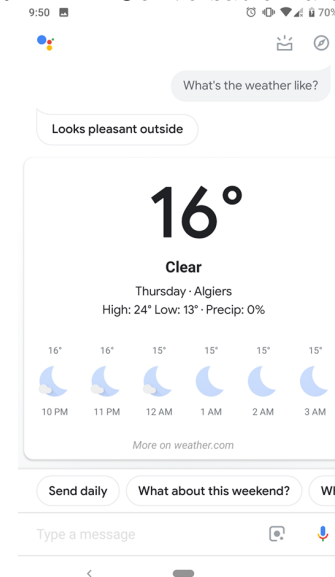


FIGURE 1.6 – Requête simple 2

Google duplex



FIGURE 1.7 – Google duplex en réservant une place dans un salon de coiffure

une des nouveautés impressionnante de G.A est la fonctionnalité Google Duplex, toujours en phase de développement, ce module est capable de passer des appels a de vrais personnes et d'avoir une conversation avec elles afin de réaliser une tâche demandé par l'utilisateur comme par exemple réserver une chambre d'hôtel, une table au restaurant ...

Apple Siri



FIGURE 1.8 – Logo d'Apple Siri

Siri est l'assistant virtuel développé par Apple, contrairement au AVIs durant sa sortie, Siri proposait une nouvelle façon de communiquer avec l'utilisateur, proposant une interface de requêtes vocale, et une façon de communiquer très humanoïde (satisfaisant ainsi le critère d'anthropomorphisme voir ??). Siri est capable de répondre à des questions précises, de proposer des recommandations, ou bien déléguer la requête à des services web. Il à l'avantage(et l'inconvénient) d'être disponible sur la multitude des appareils qui composent l'écosystème d'Apple (MacBook,iPhone,iWatch ...).



FIGURE 1.9 – Intégration aux applications 1 [10]

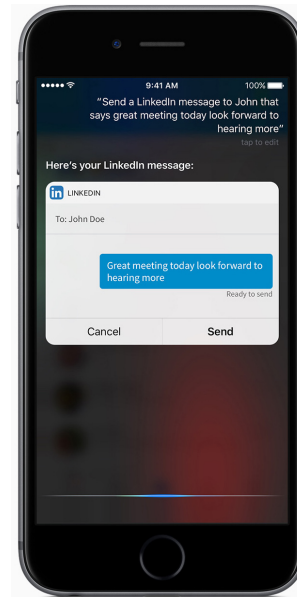


FIGURE 1.10 – Intégration aux applications 2 [10]

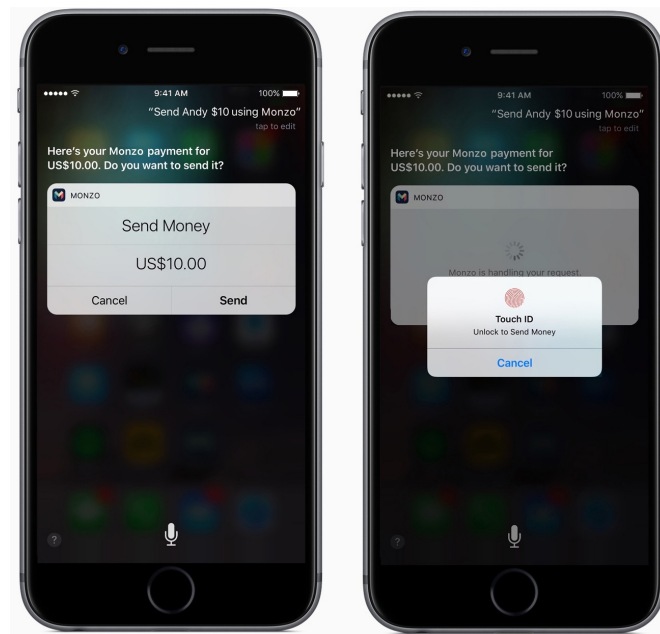


FIGURE 1.11 – Service paiement 1 [10]

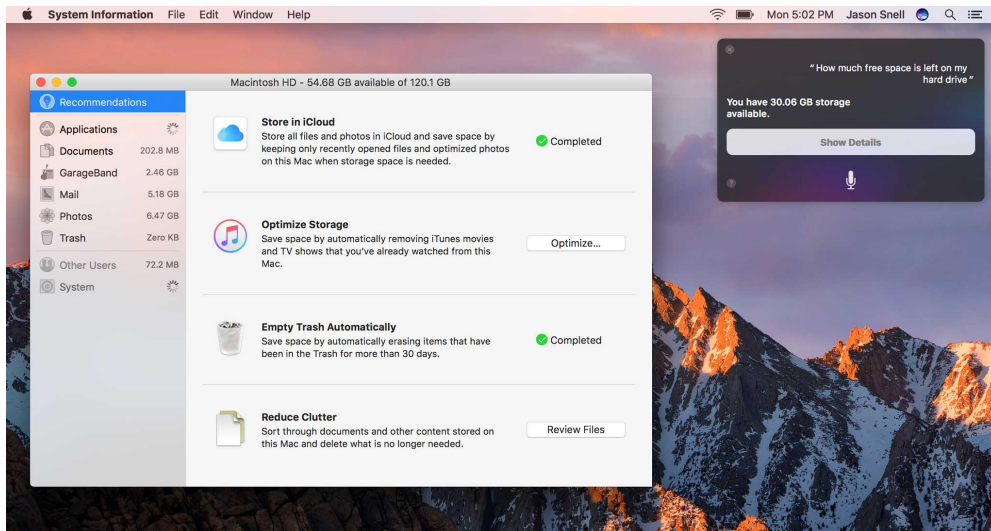


FIGURE 1.12 – Siri sur un laptop [11]

Amazon Alexa



FIGURE 1.13 – Logo d'Amazon Alexa

Amazon Alexa est un assistant exclusivement intégré au dispositif Amazon Echo (un haut-parleur portatif), à l'instar de Siri, il est aussi capable de communiquer avec l'utilisateur par le biais de la parole, pouvant ainsi exécuter diverses commandes comme joueur de la musique, réciter des livres audios, annoncer des news en temps réel (Résultats sportifs, tendances politiques ...), son atout majeur est sa capacité à s'intégrer à plusieurs appareils connectés (Contrôleur de thermostat ou de lumière ambiantes dans une Smart-House) ainsi que la possibilité d'ajout de SKILLS (ou compétences) de la part des développeurs tiers pour enrichir la panoplie de services que peut offrir Alexa.



FIGURE 1.14 – Haut parleur Echo embarquant l’AVI Alexa

Microsoft Cortana



FIGURE 1.15 – Logo de Microsoft Cortana

Cortana est la tentative de la part de Microsoft d’intégrer un assistant dans son système d’exploitation Windows 10 et WindowsPhone, il propose divers services de base tel que planifier des tâches, exécuter des commandes via la parole, et analyser des résultats de recherche sur le moteur de recherche de Microsoft, à savoir Bing, pour répondre à des questions.

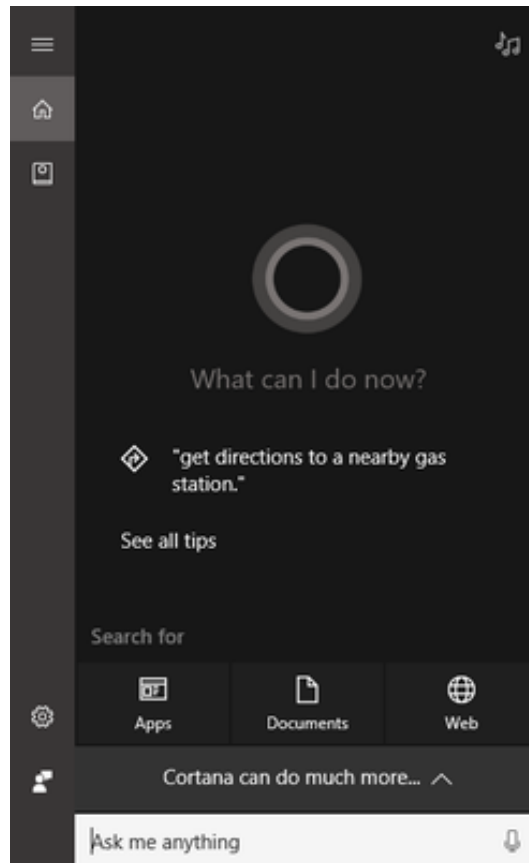


FIGURE 1.16 – Cortana répondant à une requête utilisant Bing

1.8 Points fort des AVI's

A travers les sections précédentes ,nous avons appris à découvrir les différents aspect d'un assistants virtuel intelligent, nous avons pu donc apprécié la puissance d'un tel système s'il venait à être perfectionner d'avantage.

En effet, en examinant les domaines d'applications, il est facile de déduire que le recours à un AVI peut grandement faciliter certaines tâches, que ce soit pour celles qui sont les plus triviales et donc peuvent retarder d'autres tâches plus importantes, ou bien celles qui doivent faire appel à la précision ou la grande capacité de calcul des machines, assurant ainsi des résultats précis et produits rapidement.

1.9 Conclusion

À la fin de ce chapitre nous pouvons donc mettre en valeur la place primordiale que pourrait avoir les AVIs s'ils arrivaient à maturité(c.à.d à briser la barrière qui sépare les humains de la machine et à faire partie de la vie quotidienne des utilisateurs).

Nous avons d'abord discuté de l'aspect historique des AVIs, ainsi que les caractéristiques essentielles d'un AVI idéal, nous avons ensuite cité quelques domaines d'applications qui leur sont

propres, pour ensuite citer quelques exemples d'assistants existants, pour enfin clôturer avec les principaux avantages de tels assistants.

Dans le prochain chapitre nous allons principalement aborder les aspects techniques des différents composants de l'AVI que nous désirerions réaliser.

- [1] D. A. Norman, *The Design of Everyday Things*. Basic Books, 2002.
- [2] R. Knote, A. Janson, L. Eigenbrod, and M. Söllner, “The what and how of smart personal assistants : Principles and application domains for is research,” in *Multikonferenz Wirtschaftsinformatik (MKWI)*, 2018.
- [3] *Handheld and Ubiquitous Computing : First International Symposium, HUC’99, Karlsruhe, Germany, September 27-29, 1999, Proceedings (Lecture Notes in Computer Science)*. Springer, 1999.
- [4] “Understanding the differences between alexa, api.ai, wit.ai, and luis/cortana.” <https://medium.com/@abraham.kang/understanding-the-differences-between-alexa-api-ai-wit-ai-a> (Accessed on 10/29/2018).
- [5] T. Dingler, “Cognition-aware systems as mobile personal assistants,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16*, ACM Press, 2016.
- [6] E. Luger and A. Sellen, “”like having a really bad PA”,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, ACM Press, 2016.
- [7] R. Trappl, ed., *Your Virtual Butler*. Springer Berlin Heidelberg, 2013.
- [8] A. Purington, J. G. Taft, S. Sannon, N. N. Bazarova, and S. H. Taylor, “”alexa is my new BFF”,” in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*, ACM Press, 2017.
- [9] B. R. Cowan, N. Pantidi, D. Coyle, K. Morrissey, P. Clarke, S. Al-Shehri, D. Earley, and N. Bandeira, “”what can i help you with?”,” in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '17*, ACM Press, 2017.
- [10] “Apple shares examples of siri’s third-party app integration on ios 10.” <https://www.idownloadblog.com/2016/09/01/apple-siri-ios-10-app-integration/>. (Accessed on 10/29/2018).
- [11] “macos sierra review : Hey siri, where did my files go ? - six colors.” <https://sixcolors.com/post/2016/09/sierra-review/>, 2016. (Accessed on 10/29/2018).