Haute Ecole

Groupe ICHEC - ISC St-Louis - ISFSC

Enseignement supérieur de type long de niveau universitaire

Artificial Intelligence and marketing: An ongoing revolution

Artificial Intelligence is helping companies to project marketing in a new era with its technologies applied to language manipulation

Mémoire présenté par
Luca BISACCHI

pour l'obtention du diplôme de
Master en Gestion de l'Entreprise- MIBM-120

Academic year 2018-2020

Promoter:
Monsieur **Sed SAAD**

Boulevard Brand Whitlock 2 - 1150 Brussels

SUMMARY

We live in a constantly changing world where innovators take it all. Companies are obliged to keep on innovating to find new ways to differentiate themselves from their competitors and opportunities for growth.

Artificial Intelligence, a discipline in full development and considered as the rising star of the IT horizon, will bring a real revolution in the way companies conduct their business and will be a key ally in all their areas of activity. In the field of marketing in particular the application of artificial intelligence technologies is helping companies to rewrite their marketing strategies to completely change the customer experience.
One of the fields of artificial intelligence where the most remarkable progress is being made and which still has great scope for development is language technology. Taking advantage of innovative features such as deep learning, voice recognition and language recognition, the most innovative companies are introducing new marketing techniques and tools that offer added value to customers, diversifying their purchasing routines, making them more interactive and stimulating and ultimately more profitable for companies. Companies that invest in these technologies will be able to gain a competitive advantage and take the lead over the competition, and companies that fall behind may never catch up again.

One of these innovations based on these features, which is becoming increasingly popular, is the voice marketing assistant, the trend of the moment. Voice assistants are now found in every smart device, mobile or fixed, and with the spread of 5G and the Internet of things they will be even more ubiquitous.
Numerous surveys show that the customer believes that the use of voice and language will improve the effectiveness and quality of research to prepare the purchase, both in terms of speed and accuracy of the result, making the consumer experience more stimulating and rewarding.

However, some doubts remain about the unconditional acceptance of these technologies. Many people believe that the personal voice assistant is not yet a sufficiently reliable support and more work still needs to be done to make it truly reliable and improve its image. In any case, it is clear that these technologies are radically changing the marketing function and companies need to adopt them fully to build customer loyalty and develop new markets.

This thesis aims first of all to present what has already been achieved in the field of voice marketing. Then we will examine how international market leaders who are redesigning their business model to integrate this and other artificial intelligence techniques are positioning themselves against the competition and how a business model that actively integrates these techniques can offer a competitive advantage over companies that will not seize these new opportunities in time.

Artificial intelligence and marketing: An ongoing revolution

Artificial Intelligence is helping companies to project marketing in a new era with its technologies applied to language manipulation

ACKNOWLEDGMENTS

The concepts addressed in this thesis were new to me and it has taken some efforts to assimilate and manage them with the necessary mastery in a relatively short time.

My sincerest gratitude goes first of all to Professor Sed Saad who in his lectures made me discover and appreciate the concepts of Artificial Intelligence that I developed in this thesis. Without his valuable advice and help throughout the development of this thesis, I would not have been able to produce the quality work that I promised myself to present.

Equally valuable and irreplaceable was the help and encouragement given to me by my thesis supervisor, Dr. Sara Assi at Ogilvy - Social Lab, throughout my internship and afterwards. Also, at Ogilvy I was lucky to be able to count on the help and encouragement of Dr. Isabelle Pin. My sincere thanks go to both of them.

My last words go to my family, who have always been by my side during the most difficult moments of my academic career, with their advice, scolding when necessary, encouragement and above all a lot of love. Thank you from the bottom of my heart.

ANTI-PLAGIARISM COMMITMENT

"I, BISACCHI, Luca, study year 2019-2020, hereby declare that the attached Thesis is free from plagiarism and respects in all aspects the rules of studies in matters of loans, quotations and exploitation of various sources signed during my registration at ICHEC, as well as the instructions concerning the referencing in the text respecting the APA standard, the bibliography respecting the APA standard, etc. made available to me on Moodle.

On my honor, I certify that I have read the aforementioned documents and I confirm that the Thesis presented is original and free from any loan from a third party who is not correctly cited. "

As part of this online filing, the signature consists of the introduction of the thesis via the ICHEC-Student platform.

TABLE OF CONTENTS

1. Introduction

1.1. Subject matter

Not everyone can afford the luxury of owning a butler who takes orders, executes them to the letter and tries every solution in order to satisfy our request. A butler who knows our preferences, our habits, our forgetfulness. If we try to pick up our phone and scroll through our address book, we will most likely find that the percentage of these people is minimal. However, this luxury is tempting to many and the giants in the technology market have long understood this. In fact, the luxury I am referring to earlier has been partially satisfied by what are referred to as "personal assistants".

While these assistants have actually been present in our lives for some time, see iPhone with its Siri personal assistant, able to provide daily information and perform small tasks, or Cortana, the personal assistant provided by Microsoft with more or less the same functionality, nowadays the focus has shifted to personal assistants installed on independent devices called voice enabled devices.

These smart assistants, if properly integrated in our habitat, can help to simplify our lives considerably, replacing many mechanisms that normally require time, physical and mental effort, and most often lead to stress and other unpleasant consequences. They can perform a wide range of actions. Initially these were limited to looking for simple information such as the date of birth of a famous singer, the weather forecast or simply setting an alarm clock, but now they can perform more complex commands such as buying some shares in a listed company, suggesting information based on your previous research, controlling the entire lighting and heating system of a house. These are just some of the unlimited capabilities that these systems can already offer.

The novelty is easy to see. You no longer need to pick up your mobile phone and type your request in the search bar. With a simple voice command that "wakes up" your personal assistant, you have access to a sea of features. An experiment produced by Stanford University has shown that people are significantly faster when they have to dictate a command using their voice than typing on the screen of a phone or any other device with a screen. This study showed that normally, under optimal conditions (without disturbing elements that could affect the result of the experiment) the voice is 2.9 times faster than our fingers, 153 vs 52 pps, and that, even if we unconsciously think that our fingers have a lower margin of error

than the voice input, in reality our voice is subject to greater accuracy, 3.93% vs 4.72% (Ruan S., Wobbrock J., Liou K., Ng A., Landay J., 2017).

These functionalities are the technological answer to issues raised by the extraordinary advances in some domains. I am referring to artificial intelligence and the internet of things. While the internet of things has allowed every single device with an internet connection to become a small piece of an immense puzzle, artificial intelligence is transforming the traditional processes with which we were used to do things, creating a new technological revolution. Extensive evidence demonstrates that artificial intelligence is the best tool to support this innovation. In fact, in order to achieve this result, the decoding process can be summarized as the recognition of a sequence of words that correspond to a certain sound and meaning in relation to the input received (Huang X., 2014).

Here are some numerical data that confirm that we are not dealing with a simple speculative bubble. Microsoft produced a report between 2018 and 2019 in which a questionnaire was used to collect data on the spread of smart assistants in the homes of Americans. The results show that in the space of 12 months the number of respondents with such accessories almost doubled from 23% to 45%. In addition, it would seem that there is also a certain rush to complete the purchase. Of the 30% of respondents who responded that they will buy a smart assistant, 78% think they will make the purchase within six months (Olson, C., Kemery, K., 2019). This "urgency" means that large manufacturers such as Apple, Google, Amazon and many others must work hard to ensure the greatest number of features for the end user.

1.2. Motivation and objective of the work

The functionalities related to language technologies will bring a competitive advantage in the market to companies that will be able to actively integrate them into their business models ensuring their long-term success. While there are industries that are already benefiting from the introduction of these features, primarily the financial, consumer goods sales and medical sectors, as we will see later, with this thesis we aim to study how leading companies adopt or plan to adopt these technologies and what benefits they can expect to have. We will then propose a concrete project to integrate voice recognition into the learning industry.

The thesis will begin with a presentation of the artificial intelligence fields necessary for the proper functioning of a smart speaker. Then I will compare some of the industries where voice recognition has brought the best benefits at the moment, namely finance, consumer goods sales and medicine, with a further focus on the business models adopted by the various leaders in the two sectors. After examining in a general way how market leaders can benefit from the integration of language technologies into their business models, the focus will shift to a project developed within the learning business. In fact, voice recognition has already been introduced into the educational system and has reached a high number of consents (introducing voice recognition into higher education). Differently from what has been done previously, however, I want to theorize how the use of recognition can exponentially increase the language skills of an individual. The micro category under examination and that of foreign languages, a field in which it has been demonstrated that practice brings the greatest benefits. The final goal is to define a business model in which speech recognition and machine learning can bring a tangible competitive advantage for the company that will integrate them effectively into its distribution strategy.

1.3. COVID-19 Impact

The assumptions that I will develop will seek confirmation in the data that I will collect through research in forums dedicated to the topic and user questionnaires. Unfortunately, at the time of writing we are facing a rather particular experience and the search for more specific and probably more crucial data for the realization has undoubtedly been penalized. The Covid-19 is having a major impact on everyday life and unfortunately this has not made it possible to carry out a part of the research that I had considered as key to the project. I refer to specific conversations with specialists in the field that would undoubtedly have formed a solid basis for my thesis. However, I am sure that the information I have gathered is sufficient for a concrete analysis and the definition of the business model, the main objective of this document.

2. Literature Review

2.1. Technology

2.1.1. Artificial Intelligence and Language Technology

One of the fields in which AI has made the most remarkable progress is the study and treatment of language. However, if in the collective imagination one has the impression that these technologies are already very advanced (we remember already 40 years ago the surreal dialogues in the film "2001: a space odyssey" between the astronaut and the on-board computer), in reality the progress has been much slower and partial. Although much progress has been made in the construction of machines with which to talk, and which can respond in an intelligible and sensible way, intelligent communication between machines and human beings remains something far away.

However, while the development of computers that can actually converse intelligently with humans and handle all the facets of spontaneous conversation is still far away, in the field of computerized language processing the progress is remarkable. As things stand today, a computer is not yet able to replace a human being if not in routine tasks but in recent years the progress has been so remarkable that the results obtained are closer and closer to what was hoped to achieve when 60 years ago the academics started to deal with the problem (Lake, B., Ullman, T., Tenenbaum, J., & Gershman, S. 2017).

One of the fields of the study of language in which progress is remarkable is the treatment of spoken language, which is a problem of considerable complexity.

Many systems have been developed that can execute simple spoken or typed commands and answer basic questions, but these systems cannot hold a conversation and have no real understanding of the words they use. But things evolve rapidly. At companies like Google, Facebook, and Amazon, as well as at leading academic AI labs, researchers are finally trying to solve that seemingly intractable problem, using AI's most advanced tools like deep learning and neural networks. The fact that we can effectively communicate with a machine as with a human being is fundamental to determine whether these devices will become entities with which we interact on a daily basis or whether they will remain black boxes that can only perform a limited and predetermined series of operations. Josh Tenenbaum, professor of cognitive science and computational science at MIT, clearly states that an artificial intelligence system whose behavior can replicate the

human being , must have language at its center (Lake, B., Ullman, T., Tenenbaum, J., & Gershman, S. 2017).

Will Knight, senior editor who writes about AI and robotics at MIT Technology Review, was very interested in the topic of AI applied to the treatment of language.

He followed the work of many pioneers of these technologies, starting with Professor Winograd who already in 1968 had tried to build a rudimentary machine that could converse with the user by writing sentences on the monitor. It was quite disheartening the opinion of the famous and influential linguist Noam Chomsky, professor at MIT, who considered extremely difficult to create machines that could treat language in an "intelligent" way, since the very knowledge of the mechanisms of human language is poorly understood (Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J, 2016).

However, attempts have been made, for example with the ELIZA program, the first known "chatbot", built by Professor Joseph Weizenbaum of MIT, or Winograd himself, with the SHRDLU program, able to handle a conversation on a limited set of objects, with alternating results that however have spurred more and more the study of the problem. Nevertheless these experiments were quite quickly confronted with the technological limits of the time: to broaden the field of conversation and teach the system the rules for dealing with more complex problems, required computing resources not existing at the time.

It was necessary to change strategy. This was done by some researchers who, inspired by neuroscience, were experimenting with artificial neural networks, sort of layers of mathematically simulated neurons that could be trained to react in response to certain inputs(Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J, 2016). However, these systems were very slow and for this reason the approach was not considered suitable to deal with logic and reasoning. However, the approach was conceptually right. Neural networks could learn to do things that could not be encoded with rules, such as the recognition of handwritten characters, which led to practical and useful applications. Supporters argued that neural networks would eventually allow machines to handle human language.

In recent years, neural networks have become much more complex and powerful. This has been allowed by important developments in the underlying mathematical theories and the availability of faster and faster hardware and more data. In 2009, researchers at the University of Toronto produced a multi-level deep learning network capable of recognizing speech with record accuracy and later developed a machine vision system that used a surprisingly accurate deep learning algorithm (Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J, 2016).

A deep learning neural network is able to recognize objects in images with very effective mathematical techniques, which even allow the system to learn. Such systems can now recognize objects, animals or faces with a similar accuracy to that of human beings. There is an obvious problem in applying deep learning to language. Indeed, words are arbitrary symbols, and as such they are fundamentally different from images. Two words can have a similar meaning, but contain completely different letters. Just as the same word can mean different things in different contexts. Here, too, many scholars have developed mathematical techniques that have allowed words to be encoded into symbols that could be treated correctly by neural networks.

One of the most committed researchers in the development of these technologies is Dr. Quoc Le who works for Google in Mountain View, California. Quoc Le's ambition is to build a machine that can hold a proper conversation with a human being and even simulate autonomous thoughts. Google is already in the process of teaching language to computers. Among other things, systems have been developed that understand syntax, recognizing nouns, verbs and other elements of the text. It is clear that for Google, progress in understanding and processing language is a strategic development factor. The basic Google search algorithm is used to simply trace keywords and links between web pages. Developments go in the direction of reading the text on the pages, trying to understand its meaning and provide better results. Dr. Le is going further: for example, by adapting the system used for the translation and captioning of images, a system has been developed that reads e-mails and can process responses to messages that require them. Progress is also being made on systems that can provide sensible answers to open-ended questions on increasingly broad topics. Sometimes the answers are so surprising that one really has the impression of dialoguing with a human being. Of course, one should have no illusions. Systems, however developed, have nothing that comes close to what we would call common sense. They are more and more sophisticated and capable of a large scale processing of ever larger amounts of data, but they do not imply a real "knowledge" of the subject being dealt with (Goodman N. D., Stuhlmüller A., ND).

Other fields that Dr Li is developing are machine learning and machine vision. With his team he has invested enormous efforts to build databases of millions of object images and associate them with a large number of notes and descriptors, with the aim of enabling deep learning systems to better understand the physical world. Dr. Le affirms that the linguistic part of the brain also receives a lot of information from the visual system and that therefore the integration of these systems is an essential block to allow the realization of true AI. However, the only deep learning of a large amount of data will not be enough: it will be necessary to think about things like emotional and social intelligence, because human beings

are not good at processing large amounts of data but are very skilled in abstraction and creativity. Even the largest and most sophisticated neural networks lack important components such as the ability to learn very quickly from a relatively small amount of data and the ability to model the world in 3D very efficiently. It is evident that the language is based on other even more basic skills, present in a newborn baby even before it learns to speak: perceiving the world visually and acting on the rest of the human body's motor systems. It is therefore clear that a real mastery of language by artificial intelligence systems will have to pass through a replication of learning and the construction of mental models with a contribution from psychology (Goodman N. D., Stuhlmüller A., ND).

On the subject of psychology, we need to talk about the work done by Dr Noah Goodman in the Department of Psychology at Stanford. He says that the technique of language is special because it is based on a great knowledge of the language, but also on an enormous amount of common sense knowledge about the world. The two concepts are inseparable. Among the various achievements of Goodman there is for example a programming language, called Webppl, which can help to give computers a sort of probabilistic common sense and can therefore make a conversation with a machine more sensible. He even managed to make the machine understand puns or interpret hyperbole. For example, if we say that we had to wait "forever" for a table in a restaurant, the machine will discard the literal meaning and understand that instead we stood waiting for a long time and this made us nervous. It is clear that the system is far from intelligent, but it shows how new probabilistic approaches introducing elements of psychology could help AI systems to speak more realistically.

Yet, despite the difficulty and complexity of the problem, the successes achieved by researchers using deep learning techniques in other fields of AI, suggest that great progress will also be made in the treatment of language. Everyone agrees that if AI is to become a widely used tool to help people to perform increasingly complex tasks, language must be a fundamental part of these systems. For example, Toyota, which invests heavily in automatic guidance technology, is working with Professor Gerald Sussman of MIT, an expert in artificial intelligence and programming languages, to develop automatic guidance systems with speech capabilities, which would be able to talk to the passenger and explain what they are doing, for example. Ideally, the system should provide not only answers, but also explanations (Goodman N. D., Stuhlmüller A., ND).

The large technology companies, led by Google and Amazon, are considering how to commercialize the technology, as evidently the ultimate goal of these companies is to increase profits. For example, intelligent assistants and health care tools are being developed. It is clear that it is important for these systems to be

able to talk to the AI behind them. Without knowing why a decision is being made, a relationship of trust between man and machine cannot be created. In fact, as AI systems become more and more sophisticated and complex, one cannot imagine a real collaboration without using language, without being able to ask them: "Why?

If these leading companies in their sectors invest so heavily in these technologies and employ the brightest minds they have in these projects, it is because they are convinced that in the near future these technologies will play an essential role in the strategic development of their business.

Complementary to spoken language processing technology is machine translation technology, where progress is certainly more substantial. For texts that are not particularly specialistic or complicated, the result of machine translation is now so good that it only requires a superficial revision and purely improvement modifications by the editor (Goodman N. D., Stuhlmüller A., ND).

With the advancement of hardware performance and the exponential increase in the computing power of today's computers, which is still in progress, we have turned to the adoption of mathematical and statistical techniques that process and model enormous amounts of data, unimaginable a few years ago. These software process enormous amounts of text, produce models, learn from texts already processed by computers or human beings and then try to apply all this information to unknown texts.

Ronald Schmelzer, principal analyst, partner and founder of the consulting firm Cognilytica, focused on AI technologies, gives us a brief presentation of the state of the art of AI applied to machine translation.

The preamble is that machine translation has received a significant boost from the application of state-of-the-art AI technologies such as deep learning but continues to face the complexities and nuances of the multitude of languages and dialects.

While the concepts behind machine translation are simple, the process of fully automatic translation is extremely complex. Human languages have a wide variety of nuances, dialects, colloquial expressions, similar words, false friends, synonyms, words with multiple meanings. The proof of this complexity is that the figure of the professional translator is still far from being replaced by machines and that when the translation of texts with legal value or highly specialized content is required, the use of a professional translator is indispensable (Goodman N. D., Stuhlmüller A., ND). It is therefore understandable why making a machine capable of processing and understanding all these complexities to provide an accurate

translation is a challenge that has put and continues to put the brightest minds of AI applied to language to the test. To move forward, it is necessary to use all known AI technologies together, such as deep learning, large database analysis and language analysis.

We can, in principle, identify three different types of machine translation technologies: rule-based machine translation (RBMT), statistical machine translation and hybrid systems that combine the two. As the name suggests, in a rule-based machine translation system, you program the system with as complete a set of rules as possible to determine how the text should be translated. The vast majority of machine translation systems developed and made available to the public use rule-based machine translation, as this technique was the first to be developed and achieve acceptable results. It is particularly suitable for technical texts because almost always what is required is a literal translation and requires only minimal revision by the human being. It also responds to a very frequent daily practical need, which is to immediately understand the essential content of a text. The progress of this technique has gone hand in hand with the evolution of technology: increasingly powerful machines and almost unlimited memory capacity(Goodman N. D., Stuhlmüller A., ND).

However, the limits of technology are quickly reached, as dictionaries are confronted with the problem of words that are mistranslated from one language to another or even require paraphrasing.

The solution to these problems is a new method of translation, known as Statistical Machine Translation. Instead of being based on the use of dictionaries and rules, the technology works with the processing of large quantities of texts with the original and translation on the side. With the evolution of AI, rule-based systems are no longer the best performing systems. The new systems increasingly use statistical techniques combined with artificial intelligence techniques such as deep learning. For example, in the task of breaking down sentences into their grammatical components, software learns from large bodies of text that have already been analyzed and translated by humans and then uses what it has learned to make its best guess about a text never seen before. Thanks to the increasing power of processors, falling prices for data storage and, above all, the explosion of available data, this approach is giving the most remarkable results. Hybrid machine translation combines statistical and rule-based approaches. Dictionaries serve as the basis for translation. The computer then learns from already translated texts the nuances that need to be rendered in the final text. Because machine learning-based translation works with the analysis of very large amounts of data and real text, it is clear that large IT giants providing data management services such as cloud computing and cloud storage are at the

forefront of the development and improvement of these technologies. Among others, Amazon, Google, Microsoft, Facebook are constantly improving their machine translation capabilities, taking advantage of the huge amount of conversations and texts that are being delivered to their platforms in virtually every language on the planet. The probability that a computer user has used Google Translate© is probably close to 100%, although the system still suffers from significant accuracy issues. Facebook presented a machine translation system that does self-learning without supervision, which seems to be much more reliable. Amazon also provides machine translation on its Amazon Web Services (AWS) platform (Goodman N. D., Stuhlmüller A., ND).

In addition to IT giants, there are at least 45 machine translation companies operating worldwide. Some suppliers focus on niche areas of professional document translation in specific specialist areas. Other companies integrate the human component into the machine translation cycle to achieve the accuracy thresholds often required by the contractual terms and conditions of service provision. Indeed, these companies providing machine translation services for businesses and for professional purposes are increasingly successful as they save organizations time and money, to the extent that even large international public organizations, such as the European Union, make use of these services in specific contexts where the quality of these translations may be acceptable. In fact, even in formulas with human intervention, machine translation can achieve productivity up to three times higher than traditional translation (Goodman N. D., Stuhlmüller A., ND).

Having exposed all the advantages and evolutions of machine translation, we cannot, however, fail to mention its limitations and difficulties in advancing. For example, the difficulty of managing the nuances of a daily speech, when it is enriched by colloquial expressions or not very widespread, are evident. Also evident is the difficulty of making a translation into a specialist language in the medical or legal fields. These are documents in which an incorrect translation has important consequences for the person. For example, medical terminology is often difficult to translate and many concepts in the psychiatric field do not have an exact translation between languages.

On a purely anecdotal basis, we can also talk about the field of literary writing and works of fiction, where machine translation is inconceivable. To write a work the author uses all his narrative and imaginary talent, using images and expressions that only another literary author can make in translation, sometimes using expressions and metaphors completely different from the original. One can never expect from a computer to express ideas such as humor or sarcasm. You can never expect an accurate translation of novels in the short term. Even culture-

specific slang and dialects are very difficult to handle for machine learning, because on the one hand the exact translation does not exist and on the other hand the translation can be highly context-dependent in an unpredictable way. Sometimes the translation of slang is simply impossible (Goodman N. D., Stuhlmüller A., ND).

Machine translation systems have made and continue to make tremendous progress and the amount of work required by a human reviewer is constantly decreasing. These improvements meet the needs of an increasing number of businesses, including commercial and marketing services that will be developed later. The fact remains that human intervention in the translation process will remain necessary for a long time to come for any translation needs with a legal or personal impact. However, the contribution of human feedback to machine translation, which the new artificial intelligence techniques allow for increasingly active integration, will allow for a constant improvement in translation and the gradual reduction of the gap between human and machine translation quality (Goodman N. D., Stuhlmüller A., ND).

2.1.2. Voice enabled devices – Smart assistants and smarts speakers

The first speech recognition systems were developed in 1950. Even if we talk about speech recognition, in reality, these systems focused on the recognition of numbers rather than words. Ten years later it was IBM that introduced the first system capable of recognizing actual words. It could recognize 16 different words and was called the Shoebox. (Sonix, 2020).

Nowadays, voice assistants have an accuracy of 95%, as shown in Figure 3, bordering on perfection. Although in those days it was expected to reach perfection much earlier, as explained by the professor of artificial intelligence at the University of Berkeley Stuart Russell in his book "Human compatible, the problem of control", the path proved to be more tortuous than expected (Russell, S., 2019). However, today smart speakers are able to recognize complex voice commands even in hostile environments, i.e. when there are background noises, called 'disturbing elements' (Russell, S., 2019).

If you analyze the meaning of smart assistant from a linguistic point of view, you can see that it consists of two terms: Smart and Assistant. According to the definition provided by the Cambridge dictionary, Smart can be traced back to a plurality of meanings: fashionable, wealthy, intelligent, quick though in action, something that uses computers to make accurate performance (Cambridge, 2020). On the other hand, on the other hand, we have assistants, simply "a helper" (Cambridge, 2020). So, if we had to give a definition of smart assistant, we could define it as a fashionable and quick in action helper.

The definition of smart assistant can be extended to all devices that have an AI system integrated in them: with such a definition we could talk about a car, a watch, a television, a telephone and many other devices. However, as Voicebot specifies in its report "Voice Assistant Consumer Adoption Report 2018" developed in 2018, if we really want to capture the essence of the concept of voice in relation to smart speakers, "we have to look at voice as a category distinct from programmatic visual interfaces on smartphones, smart watches and televisions" (Kinsella, B., Mutchler, A., 2018).

As we have previously introduced, the market for smart assistants integrated in a special device, smart speakers, is anything but saturated. Statista, a renowned source of accurate forecasts, has estimated that by 2025 the world market for smart speakers will exceed $35.5 billion (Statista, 2020). The

contenders for a slice of the market are certainly not lacking, although in reality it already seems that there are some clues as to who will be able to win.

VoiceBot, a leader in voice recognition information, also attributes the role of market leader to Amazon; in fact, with various affordable models more or less accessible to all, Amazon was able to take possession of about 72% of the global market for smart speakers in 2018, with a slight decrease in 2019 to 61%. Second place was given to Google with its Google home, which grew from 18.4% in 2018 to 24% in 2019, laying the groundwork for a potential race for two. (Kinsella, B., Mutchler, A., 2019).

Hereafter we will analyze both devices with the aim of defining their merits, defects and fields of use.

2.1.2.1. Amazon Echo Alexa

Through the analysis of the article written by Priest and Price, we can undoubtedly say that Amazon Echo, with its alexa AI system, was the first ever smart speaker (Priest D., Price M., 2020), with sales of over $5 million in just two years after the initial launch (Lei X., Tu G., Liu A., Ali K., Li C., Xie T., 2019), which took place at the end of 2014 (Vigliarolo B., 2018).

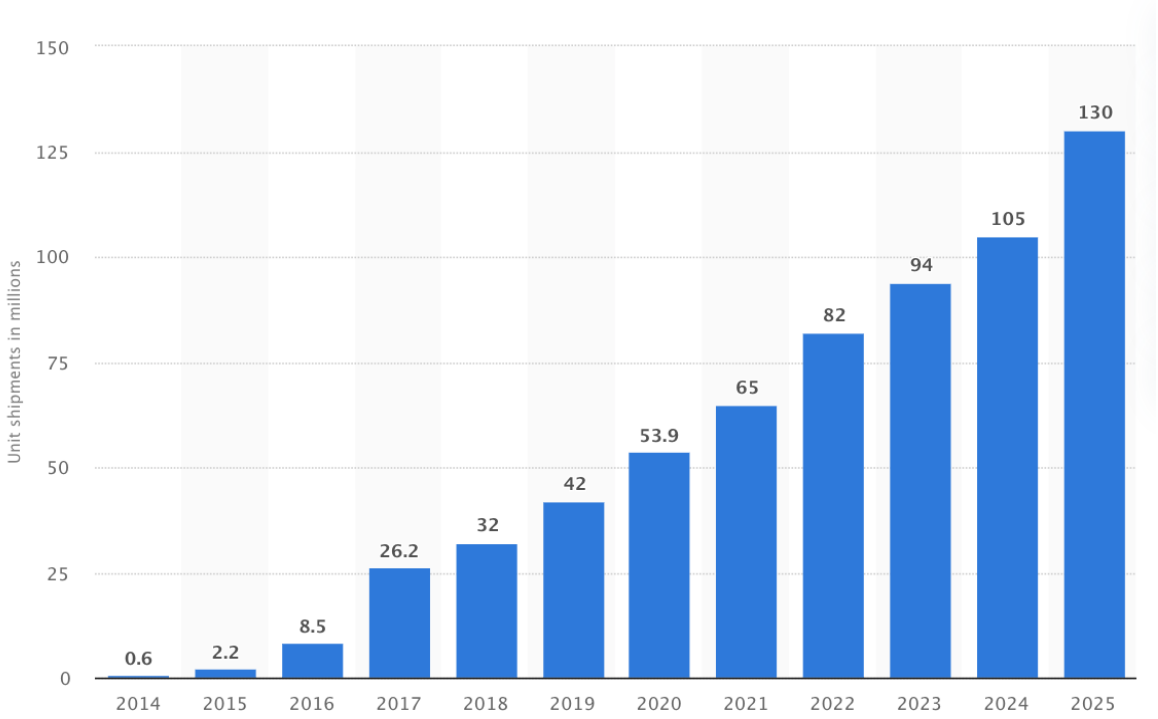

Figure 1: Prevision of Amazon Echo sold by 2025, by Statista

Alexa can perform a large number of actions, such as ordering products directly from its native marketplace, reading emails, creating events and saving notes on the calendar (Vigliarolo B., 2019). According to Wired, in 2017 Amazon Alexa would have been optimized to perform more than 10,000 required actions by voice command, (Barrett, B., 2017). This gives it a clear advantage over its competitors, who have a very limited arsenal of functionality in comparison.

2.1.2.2. Google Home

Google Home is the smart speaker developed by Google that aims to challenge Amazon Echo Alexa in the race for innovation in the field of voice combined with electronic devices (Kinsella, B., Mutchler, A., 2019). The device was officially introduced on the market in 2016 (Kinsella, B., 2020). Thanks to Amazon's pioneering work to create a new market, in the two years following the release of Google Home, Google announced that the sale of its smart device had contributed to sales of about $3.4 billion (Kinsella, B., 2020).

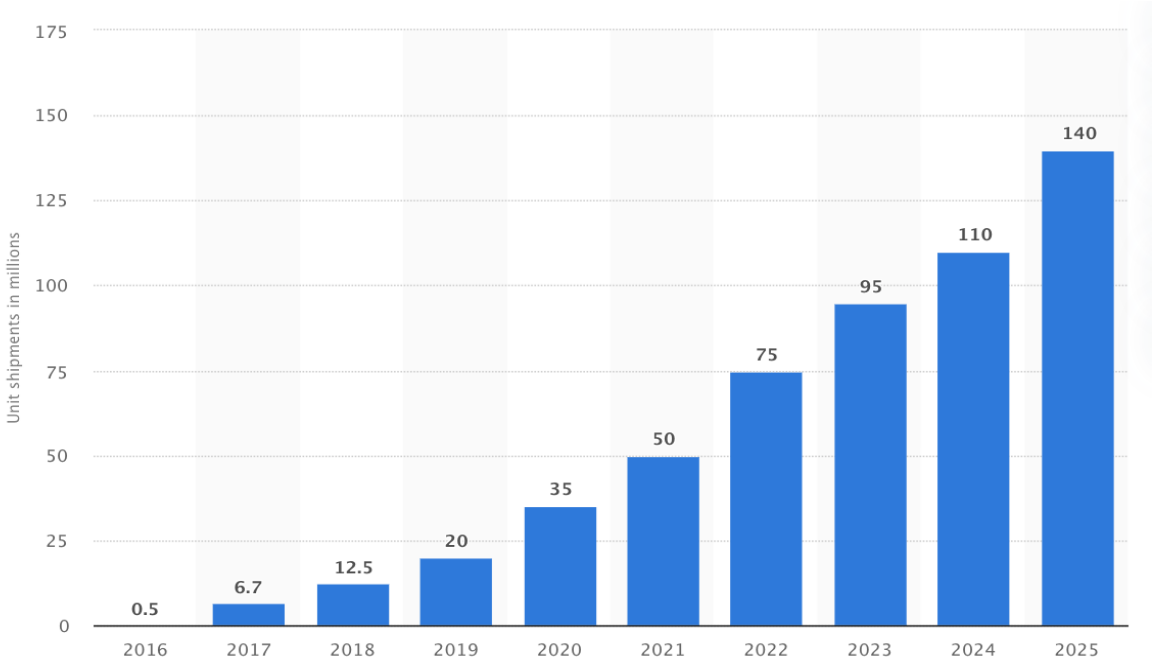

Figure 2: Prevision of Google Home sold by 2025, by Statista

As for the actions available in Google Home's portfolio, just to have a certain degree of consistency in the writing, Google Home is also able to execute the most classic voice-activated requests: record an audio note, check the weather, search for information on the Internet (Vigliarolo B., 2019).

2.1.2.3. A Comparison Between the two leaders of the market

As we have already introduced, the two smart speakers distributed by Amazon and Google are considered the two market leaders. There are several factors that make the performance of Google home and Amazon Echo Alexa superior to those of their opponents.

First, both devices can be connected to any other 'smart' device via a simple internet or Bluetooth connection (Alyssa, 2020). (You have certainly heard of the Internet of Things). It is worth reminding that under the generic expression “the Internet of Things (IoT)” falls a bundle of technologies that allows any device that has an internet connection to be integrated into a larger network of systems creating synergies that will totally change our lives and the way we do things (Morgan, J., 2014).

In addition, the two devices distinguished themselves from other potential competitors by the high accuracy of data reception. Thanks to the accuracy achieved by AI in the field of voice recognition (Business Insider, 2017) (Figure 3), the two smart speakers have earned the trust of their consumers. In fact, in a report presented by Microsoft between 2018 and 2019, the number of people who said they did not intend to buy a smart speaker fell by about 40%, from 47% in 2018 to 29% in 2019 (Olson, C., Kemery, K., 2019).

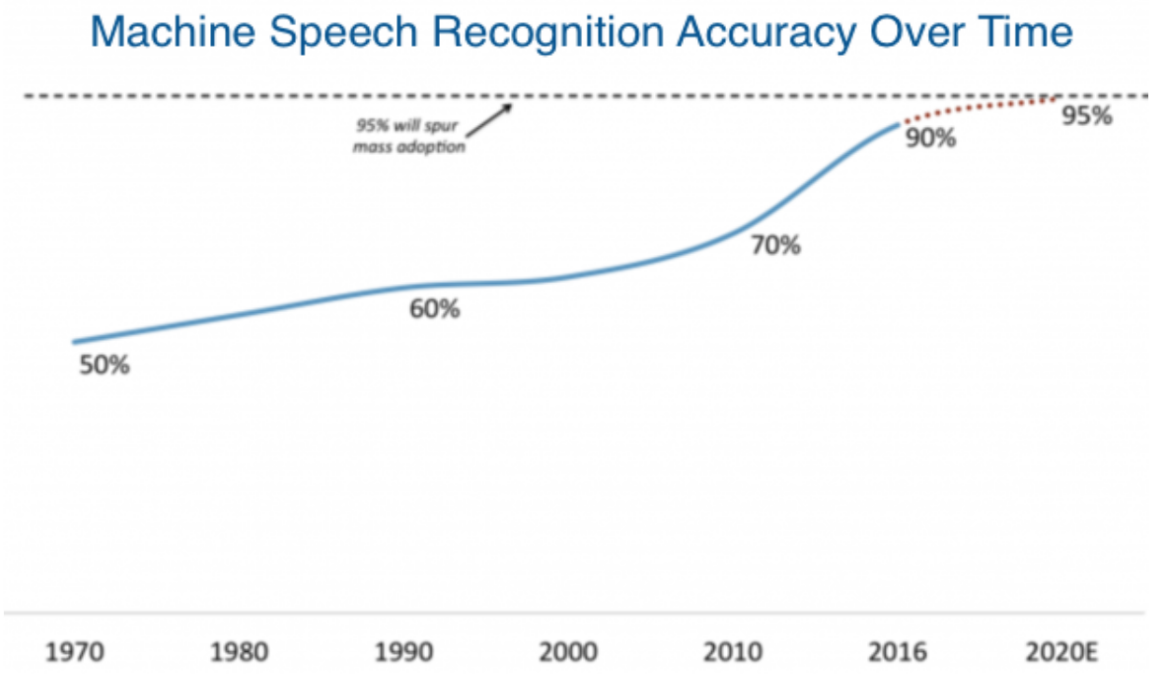

Figure 3: Machine speech recognition accuracy over time, by Business insider

In fact, as reported by Voice bot in its report on the adoption of smart speakers by consumers (Kinsella, B., Mutchler, A., 2019), consumer confidence has increased compared to the introduction of voice command requests, even by those who do not have a smart speaker.

Let's now try to make a comparison on some of the components of the two devices; From a set up point of view, both devices must be connected to their own accounts, either Google or Amazon (Hestia, 2019). This ensures that the devices can draw from a virtually unlimited database of information. Unlike Amazon Echo Alexa, however, Google home can also serve as a host for other smart assistants such as Cortana and Siri (Vigliarolo B., 2019). You need to download additional plug ins and you're done. As far as functionality is concerned, Amazon Echo has more than 10,000 solutions available (Barrett, B., 2017), all of which can be downloaded separately and integrated in a few moves; Google Home, on the other hand, has much more limited functions even if all the necessary ones are present. In terms of voice recognition, the most important feature, a study produced by a group of researchers has shown empirically that there are no major differences in their performance (Berdasco A., López G., Diaz I., Quesada L., Guerrero L., 2019).

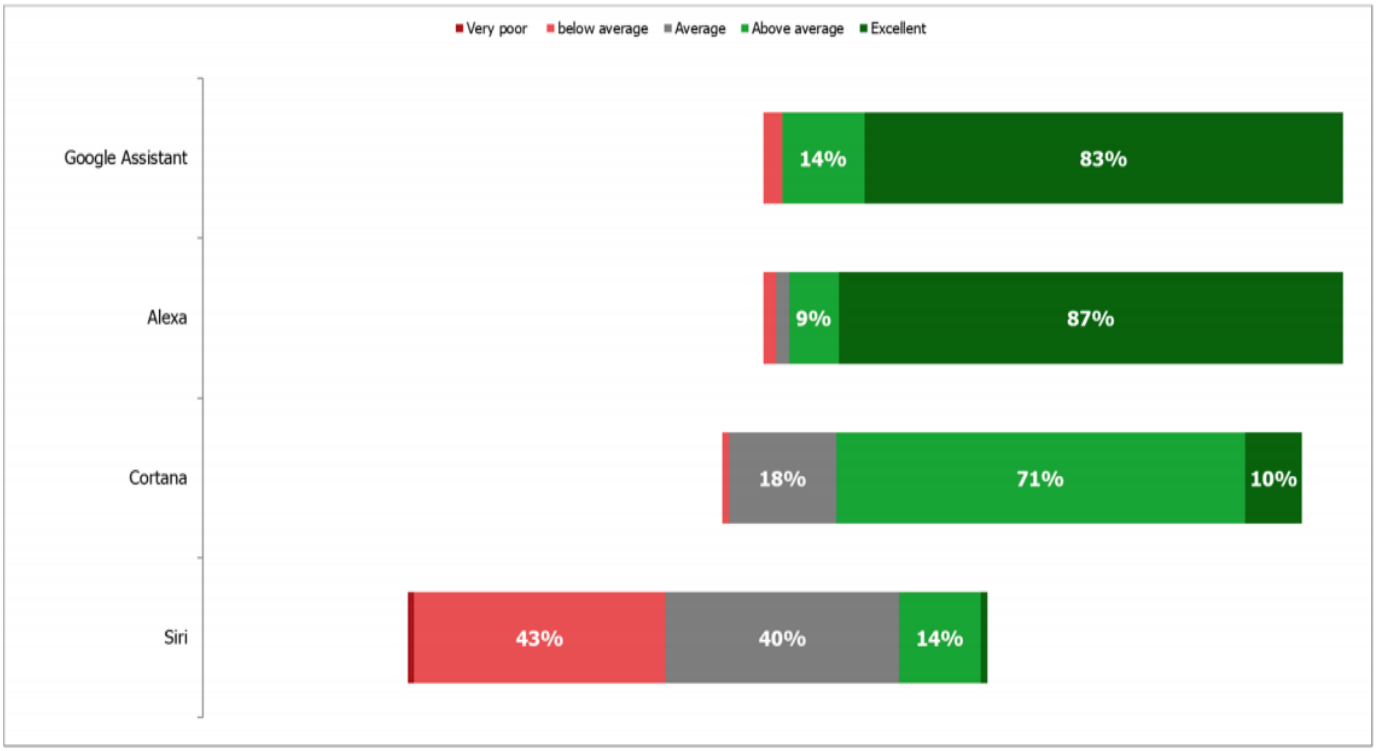

Figure 4: Evaluation through test of the speech recognition accuracy per smart assistant, by Berdasco A., López G., Diaz I., Quesada L., Guerrero L., 2019

As regards the competitors, which were also examined in the same study, the same cannot be said. The results show that Siri, despite being the longest available smart assistant in the voice enabled devices market, is the one with the worst results (Berdasco A., López G., Diaz I., Quesada L., Guerrero L., 2019).

2.2. Language, voice technology and voice marketing applied to different types of business

The use of voice marketing as a means of communication with its audience has already proven effective in several industries (Kayla, 2020). As explained in an article written by VoiceBot, a leading company in the field of innovation related to voice recognition, the industries that have benefited most from the introduction of this support in their business model are of all kinds: from the finance industry to the medical industry, from the entertainment industry to the salt and retailing industry; but also the catering industry, the automation industry (Voicebot, 2019). In practice, if applied following the right strategies, voice marketing can exponentially increase the value that a company can distribute (Digital Marketing Institute, 2020). In figure n, a graph on the investments in Artificial Intelligence (and therefore, also in voice recognition) supplied by KPMG, we can see which are the industries that have invested the most in technological innovation: in the first place, in fact, we find the Financial System, which has grown by 830% from 2015 to today; in second place, the retail sector, which has passed from an investment of 0.8 billion in 2015 to 6.6 billion in 2020; instead, Healthcare is in fourth place, with a total investment in 2020 of 3.8 billion (KPMG, 2019).

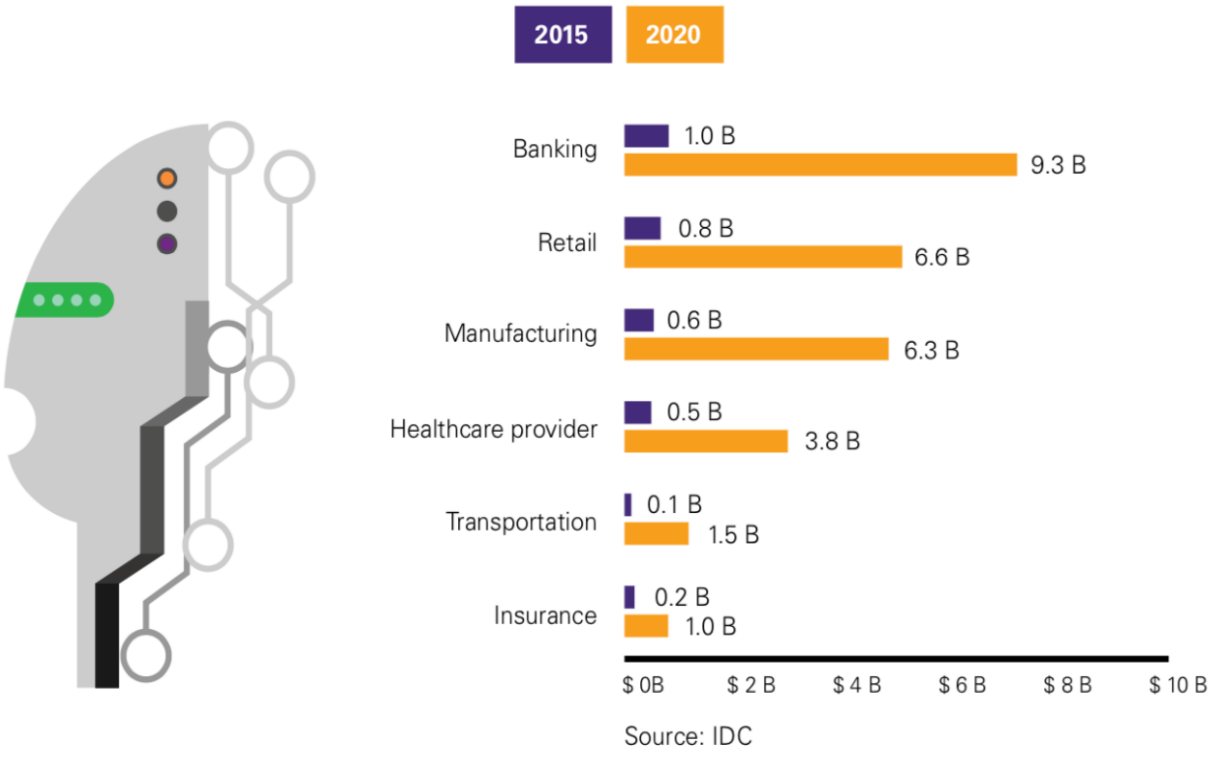

Figure 5: investments for integration of voice marketing in different business area by 2025, by IDC

So let's start with a definition of voice marketing. Trinity Audio defines voice marketing as "a set of strategies and tactics that are used to reach specific targets through the use of different voice enabled devices powered by voice assistants" (Jaworsky, R., 2020).

In practice, through this process there is the possibility to modify and adapt its contents according to the audience we are facing. Since these smart devices continuously collect data about their users, they are able to provide them with the content adapted in detail, thus altering consumer preferences and leading them to choose one product rather than another (Holzwarth, M., Janiszewski, C., Neumann, M.M., 2006). However, customers are not the only stakeholders involved in the process. Indeed, in order for the strategy developed to be effective, the manager leading the project must also act accordingly (Giebelhausen, M., Robinson, S.G., Sirianni, N.J., Brady, M.K., 2014).

Alex Mari, researcher at the University of Zurich, says that these agents (Enabled speaker devices) will exponentially increase their influence on consumer behavior, because through their continuous use, they will gather more and more detailed information about our personality and will be able to provide us with the right solution at the right time (Mari A., 2020). It (Mari) explains how these devices will become a key element in the consumer's journey as direct intermediaries between the market and the consumer. In fact, smart devices will be the new "touch point for brands", which will have to adopt new strategies and solutions in order to be able to "talk" with their target. Only by building a strong presence on this new channel will companies be able to reach their consumers everywhere. For example, the famous American coffee chain starbucks has developed a voice assistant on its mobile phone application that allows you to order your favorite drink through the support of Amazon Alexa (Mutchler, A., 2019). Pizza Domino has also adopted a similar process. The user can order their pizza via the mobile app without having to wait for the long queues that would normally occur with a single orderer. However, unlike starbucks, the app is available not only for Amazon Alexa, but also on the rest of the smart assistants like google home and Siri (Kinsella, B., 2017).

2.2.1. Financial and Banking industry

One of the areas in which most progress has been made through the use of voice marketing is finance. Many banks and credit institutions are struggling to make the best use of the functions available in smart assistants such as Amazon Alexa and Google Assistant (Streeter B., 2020) As the number of calls to banks has increased significantly in recent years, managers and insiders have realized the need to better integrate this technological support (Streeter B., 2020).

According to Dipanjan Chatterjee, a professional analyst, companies that manage to efficiently integrate an automated voice support solution will see their net score increase by 25% on the scale under consideration.

Chatterjee explains that there are 3 simple reasons why using voice assistants will improve a company's performance:

- It's **natural**. People normally prefer to talk rather than write or press keys.

- It's **simple**. Because we always have hundreds of things to deal with, being able to keep our hands free and use only our voice would decrease the time required to perform a task.

- It's **emotional**. People are able to express concepts more simply when they speak than when they write, especially when dealing with topics they don't know thoroughly (Chatterjee D., 2019).

According to a report published by Business Insider considering the introduction of artificial intelligence into the banking system, it is estimated that banks could save approximately $447 billion by 2023. (Business Insider, 2019). If we consider that $199 billion, about 45% of the total, would derive from the introduction of automatic systems of banking conversation, and therefore voice assistance, we can guess the potential linked to the matter. Analyzing the graph in figure 4, the financial system would benefit in two ways:

- **AI biometrics technology**: the system would be able to recognize the voice of the person having the conversation with the machine. Thanks to this function, billions of dollars' worth of fraud could be avoided.

- **Personalized insights**: The system would be able to understand exactly what the client is looking for and could offer the right support in addition to the right advice. This would result in an improvement in customer care.

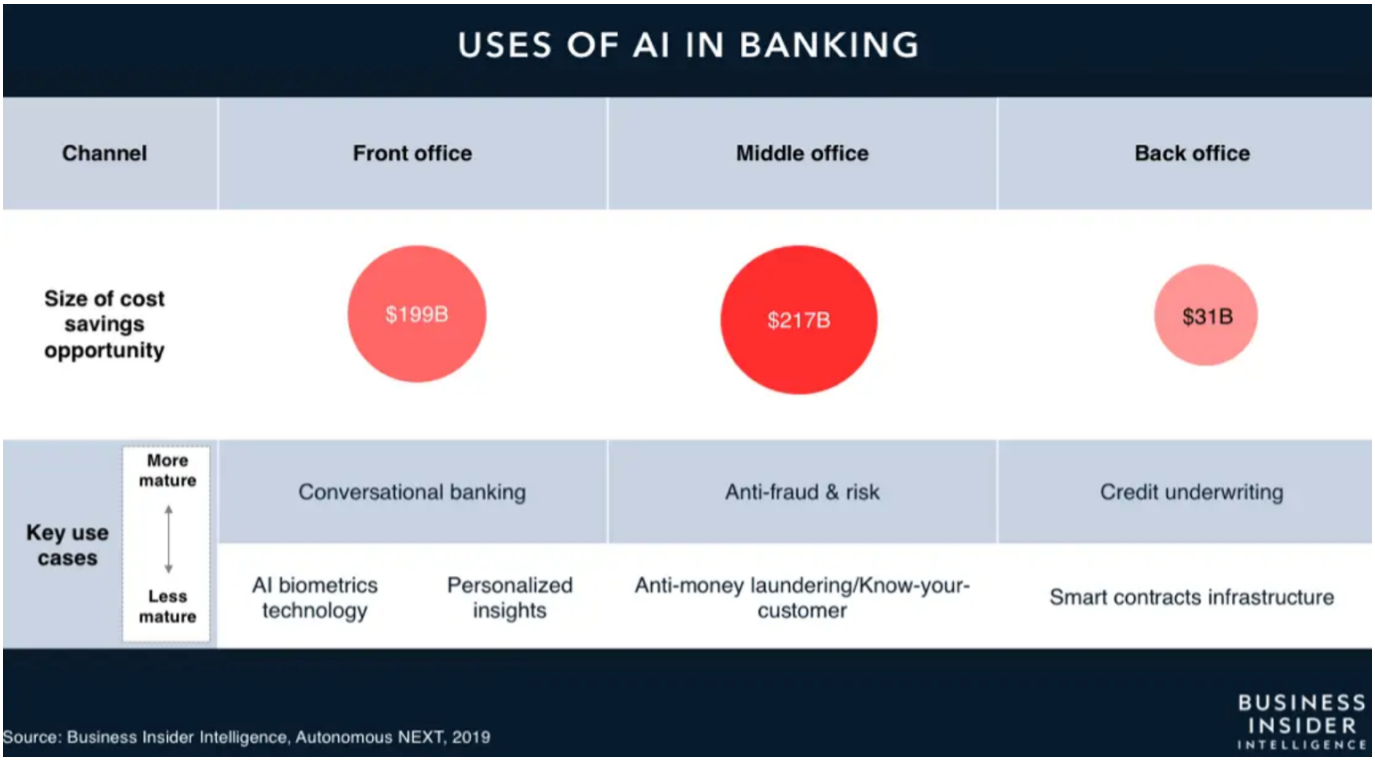

Figure 6: Different uses of AI in banking, by Business Insider

While the first functionality would ensure direct savings due only to security, the second functionality would help improve customer management and retention, a key aspect according to Andres Abumohor, Co-Founder of OmniBank, a bank offering financial support to small and medium sized businesses in Latin America. In fact, in an article published by ReadWrite in early 2020, Abumohor explains how customer service is also a key service in the banking industry. Since this technology could more easily understand the customer's demand than a human being, it would be able to offer a higher and more concrete support avoiding the frustration of the customer himself. Due to increased competition in the sector, he

continues, banks "cannot afford to annoy the customer with inadequate information". (Abumohor, A., 2020).

In addition, Abumohor also makes a consideration regarding safety, confirming what we said earlier by resuming the report produced by business insiders. The founder of the small banking institution explains that by using a voice recognition system, the clients of the different banks would feel more secure in relation to the type of data they normally give to the banks. This sensitive data is the preferred target of hackers and the theft would certainly compromise the relationship between the two parties (Abumohor, A., 2020).

2.2.1.1. Business model comparison

As we have introduced in the previous paragraph, if financial institutions want to benefit from the inclusion of voice marketing and therefore artificial intelligence, they must develop a business model that combines machines and people within the same system (Deloitte, 2020).

Hereafter we will try to deepen the concept by bringing examples of banks that have succeeded in doing so, producing value for themselves and for their customers.

2.2.1.2. Bank of America

The American bank presented its smart assistant at the end of 2016 (Taylor A., 2016). Erica, so named to recall the name of the Bank, was then officially launched in spring 2018 (Bank of America, 2018), with the aim of making the System with which operators and clients dialogue and exchange information leaner. In fact, the smart assistant, thanks to its system based on artificial intelligence, can make analysis and predictions in order to help clients to execute their payments, check their bank accounts, manage their savings and pay taxes (Taylor A., 2016). The main advantage over the classic banking consultant is that Erica is able to collect detailed information as the user communicates with her. Through these, she is able to provide solutions developed specifically for the user, suggesting solutions that could not come from a human being (Future Digital Finance, 2020).

The success of the assistant is traceable in time. On June 12, 2018, the American Bank announced that Erica was used by more than one million users. As the only one of its kind, in the sense that it was the first smart assistant introduced in the world of finance (Bank of America, 2018), Erica was a huge success right from the start. She managed to simplify processes that had been taboo for most users until then. Michelle Moore, head of digital banking at the Bank of America said that customers were really satisfied with the result.

Only one month later, on July 13, 2018, the American bank published an article on its home page stating that Erica had already doubled the number of users who had interacted with it (Bank of America, 2018). This allowed the bank to increase its transactions by 118% in Q2 compared to the previous year (Bank of America, 2018).

Less than a year after this incredible achievement, the Bank of America is back to talk about itself for another success achieved thanks to its digital assistant Erica. In fact, on May 28, 2019 there were more than 7 million users who had made transactions using the digital assistant (Bank of America, 2019). From the point of view of the possible requests that a user can make, the data reported by the American bank show an increase of 100% compared to when the service was launched, going from 200,000 possible requests to over 400,000 (Bank of America, 2019).

Let's look at the customer transaction data, where we can find various forms of deposits, loans and investments. Clients have access to 4,300 centers in 34 different countries, in addition to the online service where they can interact with the smart assistant Erica (Bank of America, 2018). If we compare the results achieved in the three years 2017 - 2018 - 2019, we can see that the results are continuously increasing.

Consumer Banking

(Dollars in millions)	Deposits 2018	Deposits 2017	Consumer Lending 2018	Consumer Lending 2017	Total Consumer Banking 2018	Total Consumer Banking 2017	% Change
Net interest income	$ 16,024	$ 13,353	$ 11,099	$ 10,954	$ 27,123	$ 24,307	12%
Noninterest income:							
Card income	8	8	5,281	5,062	5,289	5,070	4
Service charges	4,298	4,265	2	1	4,300	4,266	1
All other income	430	391	381	487	811	878	(8)
Total noninterest income	4,736	4,664	5,664	5,550	10,400	10,214	2
Total revenue, net of interest expense	20,760	18,017	16,763	16,504	37,523	34,521	9
Provision for credit losses	195	201	3,469	3,324	3,664	3,525	4
Noninterest expense	10,522	10,388	7,191	7,407	17,713	17,795	—
Income before income taxes	10,043	7,428	6,103	5,773	16,146	13,201	22
Income tax expense	2,561	2,813	1,556	2,186	4,117	4,999	(18)
Net income	$ 7,482	$ 4,615	$ 4,547	$ 3,587	$ 12,029	$ 8,202	47
Effective tax rate [1]					25.5%	37.9%	
Net interest yield	2.35%	2.05%	3.97%	4.18%	3.78	3.54	
Return on average allocated capital	62	38	18	14	33	22	
Efficiency ratio	50.68	57.66	42.90	44.88	47.20	51.55	
Balance Sheet							
Average							
Total loans and leases	$ 5,233	$ 5,084	$ 278,574	$ 260,974	$ 283,807	$ 266,058	7%
Total earning assets [2]	682,600	651,963	279,217	261,802	717,197	686,612	4
Total assets [2]	710,925	679,306	290,068	273,253	756,373	725,406	4
Total deposits	678,640	646,930	5,533	6,390	684,173	653,320	5
Allocated capital	12,000	12,000	25,000	25,000	37,000	37,000	—
Year end							
Total loans and leases	$ 5,470	$ 5,143	$ 288,865	$ 275,330	$ 294,335	$ 280,473	5%
Total earning assets [2]	694,676	675,485	289,249	275,742	728,817	709,832	3
Total assets [2]	724,015	703,330	299,970	287,390	768,877	749,325	3
Total deposits	691,666	670,802	4,480	5,728	696,146	676,530	3

[1] Estimated at the segment level only.

Figure 7: Bank of America's Customer account in year 2017-2018, by Bank of America

Consumer Banking

(Dollars in millions)	Deposits 2019	Deposits 2018	Consumer Lending 2019	Consumer Lending 2018	Total Consumer Banking 2019	Total Consumer Banking 2018	% Change
Net interest income	$ 16,904	$ 15,939	$ 11,254	$ 11,086	$ 28,158	$ 27,025	4%
Noninterest income:							
Card income	(33)	(33)	5,117	5,135	5,084	5,102	—
Service charges	4,217	4,298	2	2	4,219	4,300	(2)
All other income	832	762	294	429	1,126	1,191	(5)
Total noninterest income	5,016	5,027	5,413	5,566	10,429	10,593	(2)
Total revenue, net of interest expense	21,920	20,966	16,667	16,652	38,587	37,618	3
Provision for credit losses	269	195	3,503	3,469	3,772	3,664	3
Noninterest expense	10,682	10,657	6,936	7,015	17,618	17,672	—
Income before income taxes	10,969	10,114	6,228	6,168	17,197	16,282	6
Income tax expense	2,687	2,578	1,526	1,572	4,213	4,150	2
Net income	$ 8,282	$ 7,536	$ 4,702	$ 4,596	$ 12,984	$ 12,132	7
Effective tax rate [1]					24.5%	25.5%	
Net interest yield	2.40%	2.34%	3.80%	3.97%	3.81	3.77	
Return on average allocated capital	69	63	19	18	35	33	
Efficiency ratio	48.73	50.83	41.61	42.12	45.66	46.98	
Balance Sheet							
Average							
Total loans and leases	$ 5,373	$ 5,233	$ 295,562	$ 278,574	$ 300,935	$ 283,807	6%
Total earning assets [2]	703,444	682,592	296,051	279,217	738,770	717,189	3
Total assets [2]	735,232	710,925	306,169	290,068	780,676	756,373	3
Total deposits	702,908	678,640	5,368	5,533	708,276	684,173	4
Allocated capital	12,000	12,000	25,000	25,000	37,000	37,000	—
Year end							
Total loans and leases	$ 5,472	$ 5,470	$ 311,942	$ 288,865	$ 317,414	$ 294,335	8%
Total earning assets [2]	724,536	694,672	312,684	289,249	760,137	728,813	4
Total assets [2]	758,385	724,019	322,717	299,970	804,019	768,881	5
Total deposits	725,598	691,666	5,080	4,480	730,678	696,146	5

[1] Estimated at the segment level only.
[2] In segments and businesses where the total of liabilities and equity exceeds assets, we allocate assets from *All Other* to match the segments' and businesses' liabilities and allocated shareholders' equity. As a result, total earning assets and total assets of the businesses may not equal total *Consumer Banking*.

Figure 8: Bank of America's Customer account in year 2018-2019, By Bank of America

In terms of connection through the digital support in which Erica is integrated, Bank of America reports that there has been a clear growth in terms of active users from 2018 to 2019, from 27 million to 38 million (Bank of America, 2018, 2019). During 2018 there were more than 6 billion logins on the mobile application, while in 2019 there were more than 10 billion logins, almost 40% more than the previous year. In terms of customer interactions, Bank of America reports that in 2019 interactions per day rose to 28 million, with 81% of these interactions

occurring via digital channels (Bank of America, 2018, 2019). The total number of customers also increased significantly compared to the previous year, a sign that technological innovation is an important driver. In 2018 they amounted to 55 million clients while in 2019 these are more than 66 million. If we look at sales, however, in 2018, 25% of sales to customers were made through digital support, while in 2019 these grew to 29%. Bookings for physical consulting through digital support also grew by 19% between 2018 and 2019, reaching about 2.3 million (Bank of America, 2018, 2019).

Moreover, a crucial result that has been achieved through the use of Erica is the age of the people using the platform:

- 15% Generation Z
- 49% Millenials
- 20% Generation X
- 16% Boomers

Jatin Metha, Artifical Intelligence expert and consultant at Aiquido, a company specializing in technology and artificial intelligence (Aiquido, 2020), described Erica as "The best example of a virtual assistant in the world of finance", as can be seen in his article published in June 2019 (Metha J., 2019).

2.2.1.3. Italian Banking System (Monte dei Paschi di Siena – MPS)

I managed to have an exclusive interview with a senior consultant at the Italian Bank "Monte dei Paschi di Siena" , which is the oldest financial institution of the world, who gave me some general and some more specific information regarding the integration of digital systems in the Italian banking system.

In relation to the "digital revolution" that is affecting Italian banks in the last decade, he explains, it is necessary to make some observations.

First of all, Italy is a country where traditional banks have always been highly regarded by customers. It is well known that it is the country with the highest per capita savings rate in the world. Households hold more mutual funds, bonds, government bonds and shares than any other country. The customers of the banks in Italy are mainly depositors and the wealth is in the hands of a category of people who have had the opportunity over time, when the rates were even 15 - 18% to accumulate an important wealth consisting of the financial instruments mentioned

above. This very wide range of clients, with modest knowledge in the financial field, has always entrusted the consultants appointed by the banks to this service with the choice of managing their savings. Institutions have strongly pushed for the "loyalty" of customers in order to obtain the exclusivity of relationships and to increase the masses administered and managed.

The recent difficulties of the Italian banking system have only relatively threatened the hegemony of traditional banks. he reluctance of investors and savers to use the new digital technologies, which finds its reasons in the poor computer literacy of the older population, has not allowed and still does not allow the take-off of digital banks to the same extent as in other more evolved states that have a much younger average age of bank customers, such as Bank of America and BBVA.

Having said this, it is possible to affirm that the use of innovative startups has not always represented an element of strong development of the underlying traditional banks.

In some cases, expectations have been severely disappointed by the low interest shown by customers in digital platforms. One concrete case is that of Widiba, online bank of MPS. When the Widiba project was presented in 2013, Fabrizio Viola, then CEO of MPS and Andrea Cardamone, number one at the start-up, defined the new bank as "one of the pillars for the relaunch and repositioning of the business model of the MPS group, a pivot on which the bank would bet a lot for the relaunch of the business and the return to profitability".
In the 2015-2019 business plan, the digital bank was to achieve four objectives:

- **To have a network of over a thousand financial advisors**
- **Exceed the 500,000 customers**
- **Reaching 20 billion in investments.**
- **Achieve these goals within 2 years.**

As of August 2018, the bank had 591 advisors and 235,000 customers. The assets intermediated 8 billion and the losses in three years amounted to 46 million euros. An extremely disappointing result.

The social strategy put in place appears less fruitful than expected and, as mentioned above, clashes with a reality that still sees wealth concentrated mainly in the hands of 50- and 60-year-olds, not exactly inclined to use social networks and not exactly the target customers of online banks (Timpone, G., 2018).

From these premises we can understand a bank in great difficulty as MPS, saved only with public money several years ago, despite having an innovative start-up and one of the first launched on the Italian market has not managed to impress on the new model of bank the expected acceleration. The reason, certainly related to the generation problem, is also determined by the collapse of confidence of MPS customers on the possibility of the bank's exit from the very strong crisis. The solidity and strength of the traditional bank, therefore, remains the fundamental element to achieve consensus. Not necessarily a policy of strong digital technological development can make up for the image shortcomings reported by the media.

As proof of this, smaller banks, which give of themselves a strong perception of financial solidity, while equipping themselves with digital tools to allow their customers to use the web, have decided not to make huge investments on their digital platform, sensing, for now, a low interest of their retail customers towards this form of relationship.

Cassa di Risparmio di Ravenna for example, a bank with a strong but limited presence in the province of Ravenna, with its great strength of penetration into the local market and with a certainly enviable financial situation, has not changed its commercial policies by making the image of capital solidity prevail over the image of a bank with a strong financial position.

Obviously, the commercial policies of each bank must have concrete references to the type of masses they administer. The exponential growth of the volumes of the masses of deposits and clients makes us understand how, until the complete computer literacy of the clientele is completed, digital banks will always have a marginal role with respect to traditional banks. Making expensive investments in this sector will bring economic returns only when an important generational change will bring with it a new model of relationship with one's own bank (Interview, 2020).

2.2.1.4. Competitive advantage

As shown by the comparison between the two business models of Bank of America and MPS, the advantages for those who have been able to redesign their business model taking into account the technological support provided by smart assistants are considerable. With the support of the two reports produced by Bank of America and a report developed by the consulting firm Deloitte, we now try to summarize them in this paragraph:

- **Cognitive Engagement**: improve the concept of the customer's request in order to offer customized solutions. Once the customer has interacted with the smart assistant several times, the smart assistant is able to provide solutions tailored to the customer's requirements as well as to the operations carried out in the past (Deloitte, 2020).

- **Time saving**: as we have seen in the Bank of American data, 38 million users perform online transactions without the need to interact with a physical subject. In addition, even if there was a need to interact with a physical subject, the appointment could be easily arranged through digital support (Bank of America, 2019).

- **Improve in security**: voice recognition can reduce the risk of fraud (Deloitte, 2020).

- **Increased customer confidence**: all 3 advantages just presented lead to the definition of an additional point. Client confidence in their institution is a key factor in maintaining client assets in a bank's portfolio. As we have seen in Bank of America's report, the number of clients who have adopted digital support is constantly growing, which means that the service obtained reflects the advantages previously presented (Bank of America, 2019).

However, as emerged during the interview with the Senior Consultant at Monte dei Paschi di Siena, the innovation factor is closely linked to the public in front of us as well as the culture and the actual service that is asked for.

In countries that are even less developed in the use of digital communication, the elements that are at the basis of the success of a credit institution are to be found more in the financial solidity and image that the bank gives of itself rather than in the digital tools it has equipped itself with. But the future will be different and it will be important that the adjustment takes place at the right time, when customers have evolved and before banks more ready for change can take away an important slice of the market (Interview, 2020).

2.2.2. Sales and Retailing

As we have seen for the finance industry, the sales and retailing industry can also count on the support of voice enabled speakers in order to increase its sales. Since, according to Voicebot's 2019 report, the number of smart speaker owners is increasing exponentially every year, brands have an obligation to adopt a strategy that puts them ahead of their rivals taking advantage of this trend (Kinsella, B., Mutchler, A., 2019). The trend is now consolidated, as confirmed by Medium in an article released in April 2020, and in the future growth opportunities will be limited with the entry of new competitors (Medium, 2020).

If one wonders what could be the reason for which this new approach is necessary, Medium has also drawn up a list of potential reasons why no brand can afford to give up the introduction of a strategy based on the use of voice.
First of all, clients want an experience that is concrete, tangible, without interruptions or problems. Many times, when searching for a product or service, clients struggle to find the type of solution they are looking for (Medium, 2020). While in the past the consumer's attention was focused on the product itself, nowadays what really makes the difference is the customer journey, as Walker says in a report published in 2020 (Walker, 2020).

Nowadays the products available on the market differ little when it comes to performance or functionality (Lobzhanidze G., 2020). Even the price now seems to be irrelevant, unlike in the past. Consumers want to "reward" the companies that can best adapt to their tastes and preferences, so that product research takes less time and less stress (Walker, 2020) (Medium, 2020). Such advanced support has the potential to improve consumer satisfaction, a key element as we have seen in the case of the banking sector (Medium, 2020).

In addition, the speed with which a user performs a purchase has also become crucial because of the world we live in. People no longer have time to waste to inform themselves through friends and online reviews about what is the best product compared to another. They want an answer immediately and having to wait too long could make them ignore the product even before the customer knows it exists (Medium, 2020).

The process of comparing available solutions is also a factor that brands must pay attention to. In fact, since the smart assistant is able to adapt its results based on previous research, brands can take advantage of this feature in order to be in the right place at the right time (Digital Marketing Institute, 2020).

Finally, another reason to incorporate the use of voice into your business model is that voice will change SEO best practices. In fact, the way a client performs a search has totally changed compared to the past, as noted by the Digital Marketing Institute in a recent article (Digital Marketing Institute, 2020). Voice searches are more detailed and therefore longer than written searches, as we can see in figure 6. This leads to the fact that brands, in order to position themselves at the top of the list of results, must integrate all possible combinations of keywords into their system (Digital Marketing Institute, 2020).

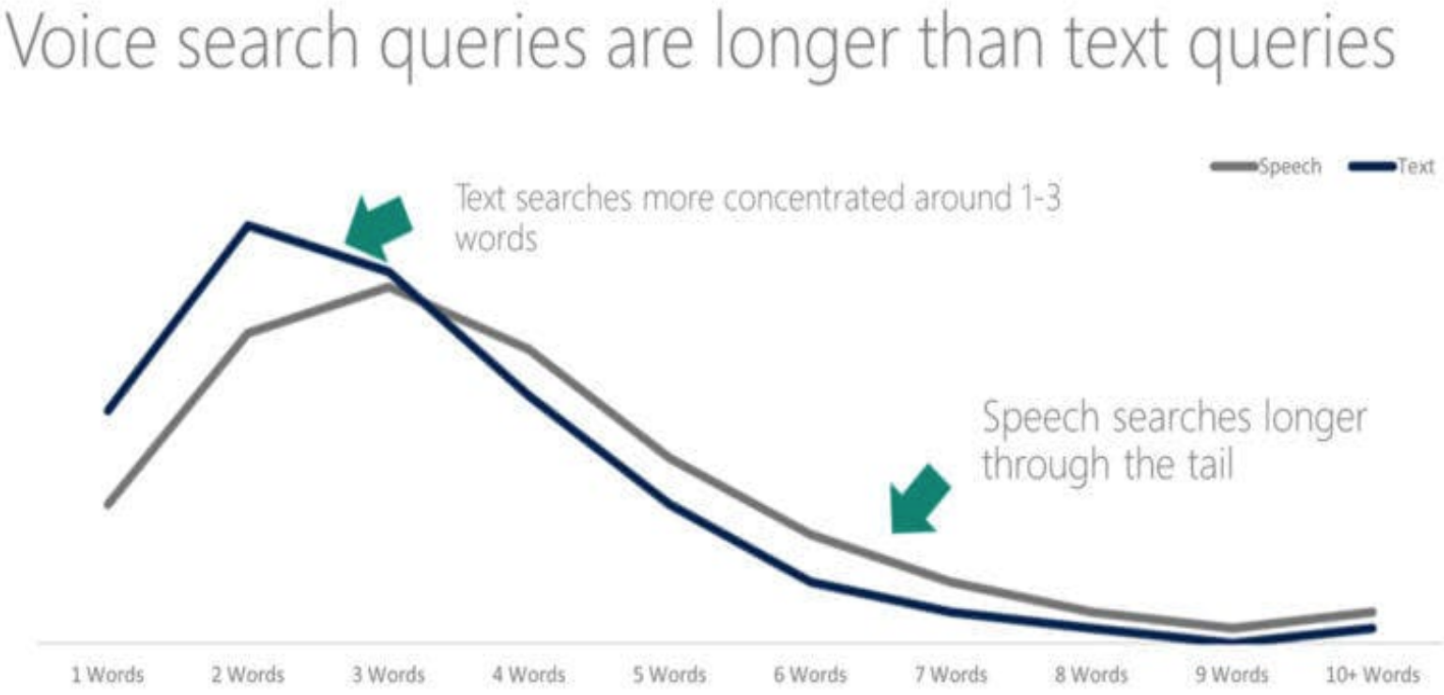

Figure 9: Difference in words searched via typing or speaking, by Digital Marketing Institute

As you can see through the graph, the number of keywords changes from 2, when the user enters the desired product in the search bar in writing, to 7, when the search is done in written form.

2.2.2.1. Business model Comparison

If we compare two examples of customer journey in which voice support is present and not, we can guess that there are 4 critical phases in which the use of voice can make the difference. Let's consider the two models in figure 10 and 11.

- In the **Awareness** phase, i.e. when the consumer is not yet aware of the solutions available on the market, brands can optimize the way they position themselves in the list of results in order to emerge against their opponents.

- Then, in the **Consideration** phase, i.e. when the customer is intent on deciding which solution best suits his preferences, the smart assistant can present the same result over and over again, influencing the consumer's final decision.

- In the third phase, during the **purchase**, the smart assistant acts as a real "agent", performing the purchase without the user having to select or press any button.

- In the last phase, once the purchase has been made, the smart speaker will always propose the previous solution as the first solution, and in case the customer has had a positive experience, this will lead him to the purchase of the same product, sealing a very high degree of **loyalty** to the brand (Medium, 2020)

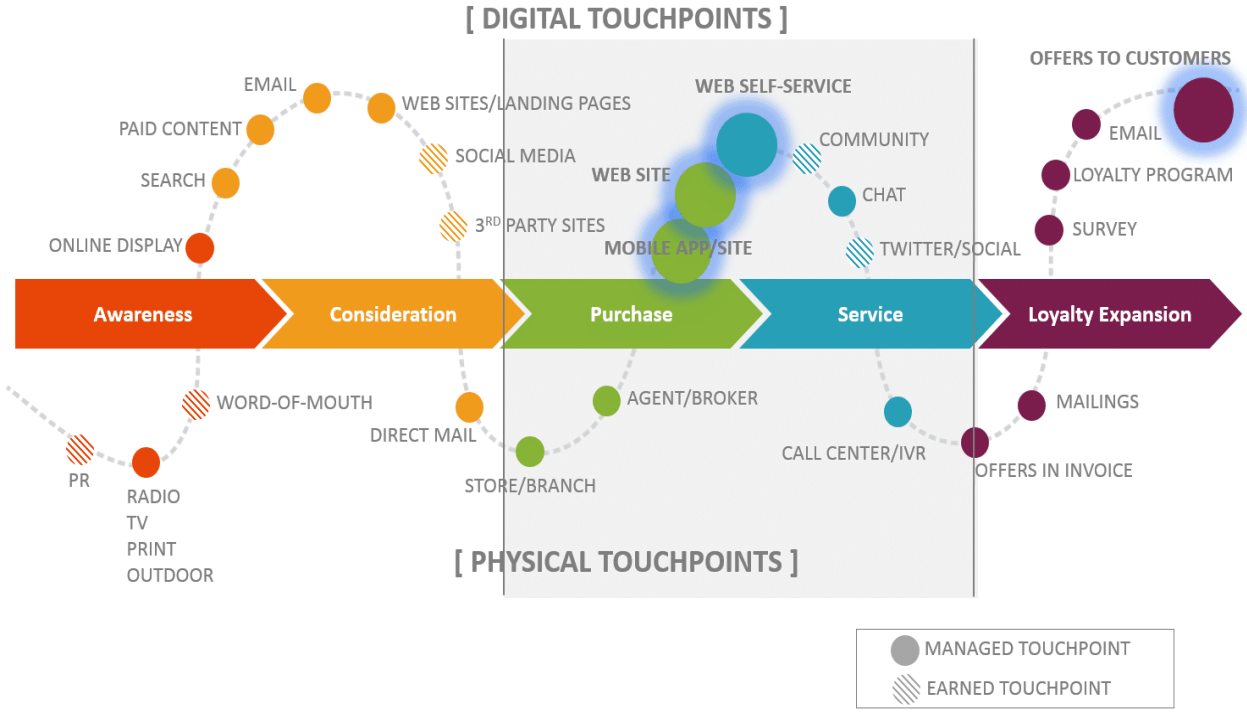

Figure 10: 4 phases of the customer's journey, digital vs physical touch points, by Medium

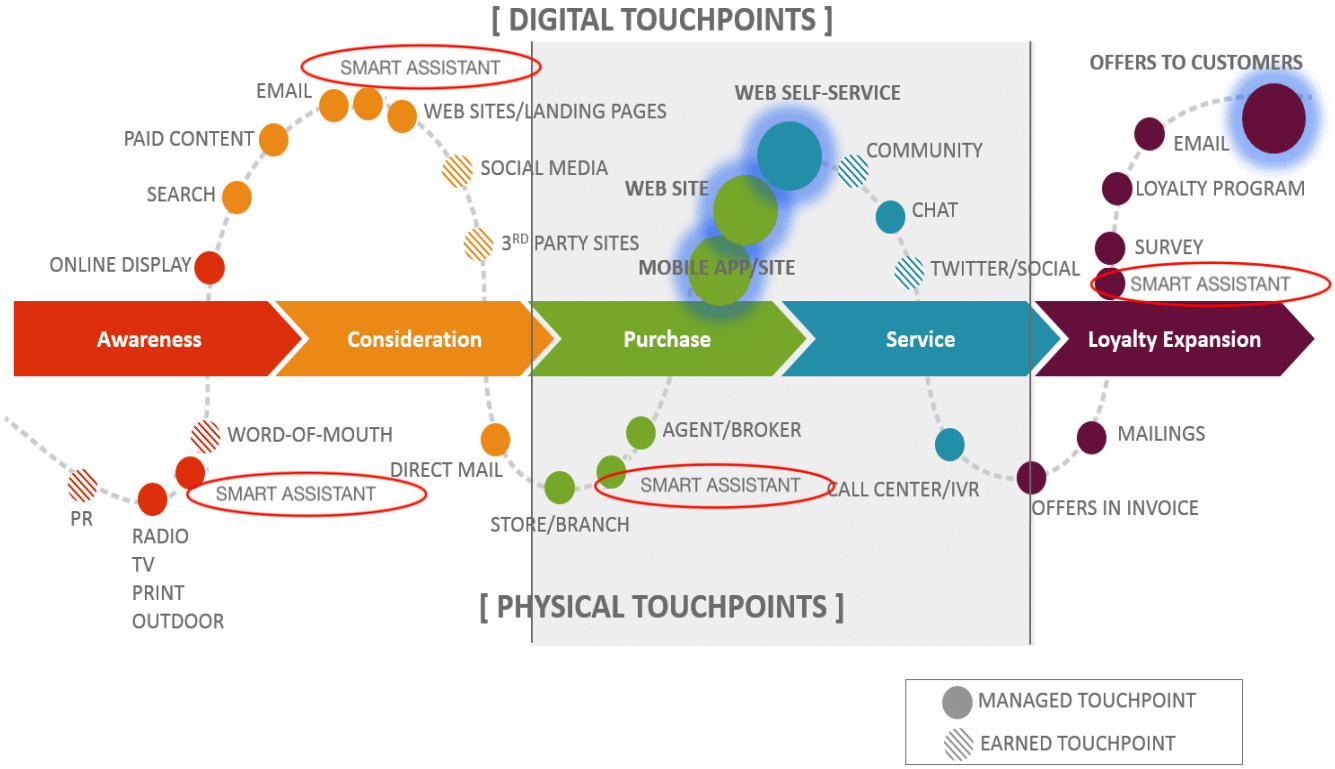

Figure 11: 4 phases of the customer's journey, digital vs physical touch points, with integration of smart assistant, personal production.

In the following paragraphs we will present how two leading brands in cosmetics and beauty have applied in a very different way these concepts with quite a different outcome.

2.2.2.2. Sephora

Sephora is a French cosmetics retailer within the LVMH brand portfolio. Through its innovative spirit, Sephora aims to educate its customers by helping them to immerse themselves in the world of well-being and health. Over the years, Sephora has built a reliable image through the use of quality products, excellent support and a consumer experience that is constantly being revolutionized (Sephora, 2020).

Sephora adopted its own digital First approach, as explained by its Executive Vice President, understanding that the customer's focus was shifting to other channels. Beth Laughton explains that the need to adapt to customer presence is of paramount importance in order not to lose ground (DeNisco A., 2018). For this reason, while competitors prefer to develop their sales departments, Sephora's focus has been on finding their customers right where they are: online. The company has therefore developed several Hi-Tech solutions that offer the user the possibility to customize their shopping experience through the use of Augmented Reality and Artificial Intelligence (DeNisco A., 2018).

With this assumption, at the end of 2018, the French brand announced that it would develop a collaboration with Google in order to offer its services in an innovative way, integrating the use of the voice available on Google home (Global Cosmetic News, 2018). Google Home will be installed at 10 Sephora resale points across the United States with the aim of renewing the consumer experience when the consumer is about to buy wellness products. This new experience has been defined as the "dream vanity experience" (Global Cosmetic News, 2018). The customer will have the opportunity to ask for advice through intelligent support and to apply the product in live mode, in addition to the possibility to buy it online or at the point of sale.

The voice support provides the ability to browse the Youtube platform, owned by Google, giving the user the ability to choose from more than 2,500 videos contained in the official Sephora channel. These videos contain tips on the various products available within the range offered by Sephora, followed by instructions on how to apply them (Iribarren, M., 2018). In addition, as always reported in the article

published by Voicebot, through the partnership, Sephora has also added the Google home Nest smart speaker to its portfolio of products that can be purchased online and physically (Iribarren, M., 2018).

Apparently, both companies have major interests in the partnership. On the one hand, for Sephora it means being able to introduce technological support as an integration to the service it offers in order to take a step ahead of its competitors, on the other hand, Google wants to strengthen its presence in the wellness market, also considered a weak point for the Hi-Tech company (Iribarren, M., 2018).

However, this would not be the first time that Sephora has identified the integration of voice use by the user as a competitive advantage over its competitors. In fact, as told by Voice bot in his long article dedicated to the brand of cosmetics and wellness products, Sephora had already developed some features that can be activated through the use of Google Home, although no collaboration had been stopped yet. Through the use of these features, Sephora had introduced the possibility for the user to book visits to its various centers, the possibility to entertain themselves with some quiz games and a portfolio of podcasts dealing with the topic of wellness (Iribarren, M., 2018).

As a novelty, in addition to the possibility to find dedicated videos on the Youtube virtual platform, Sephora's personal assistant, Sephora Skincare Advisor, is also able to provide advice regarding the various products or the stores closest to the user (Sephora, 2020).

As the female gender when browsing the Youtube platform is generally interested in videos containing advice on the application of make-up and specific products, as reported by The Washington Post newspaper in an article dating back to 2018 (The Washington Post, 2018), it can no doubt be said that the move made

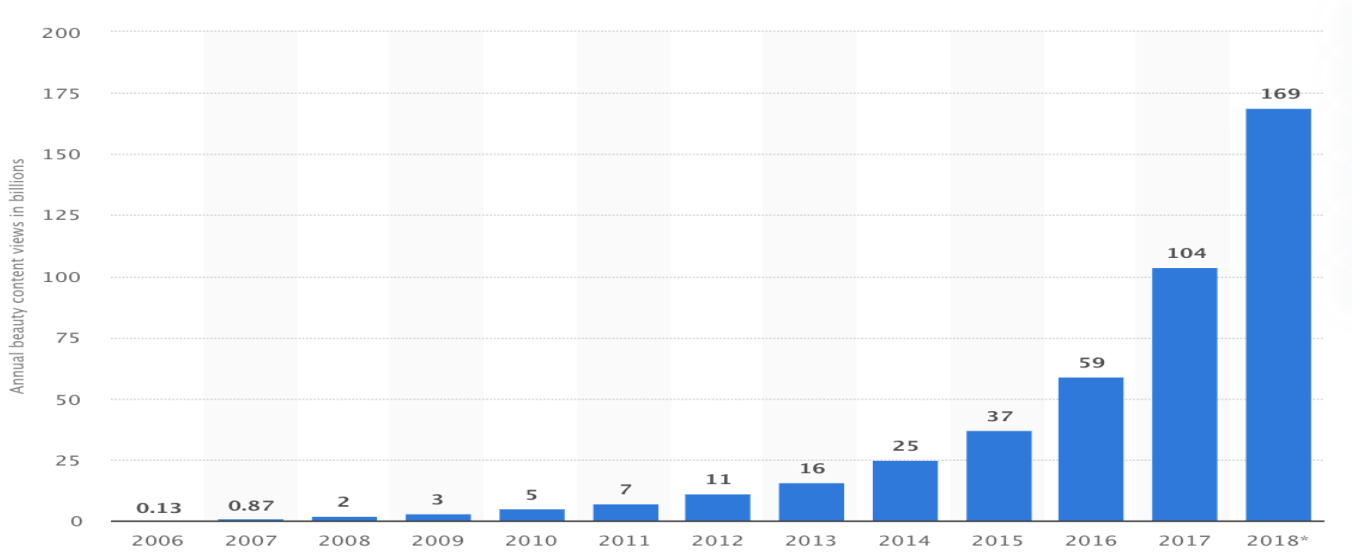

Figure 12: views of videos containing content related to the topic of beauty 2008 – 2017, by Statista

by Sephora's management has proved successful. According to a Statist produced by Statista, the views of videos containing content related to the topic of beauty have grown exponentially, rising from 2 billion in 2008 to 169 billion in 2017. (See Figure 12).

Unfortunately, due to lack of data it is not possible to attribute a numerical value to the innovation brought by Sephora Skincare advisor and therefore we will limit ourselves only to an analysis of the revenue of its ecommerce. In the chart below produced by EcommerceDB, which takes into account the revenue from 2016 to 2019, in anticipation, you can see that they have a growth rate of 25% on average, from $ 580 million in 2016 to 1.2 billion in 2019. Since the voice assistant function was introduced at the end of 2018, it is reasonable to assume that this functionality has actively contributed to this result.

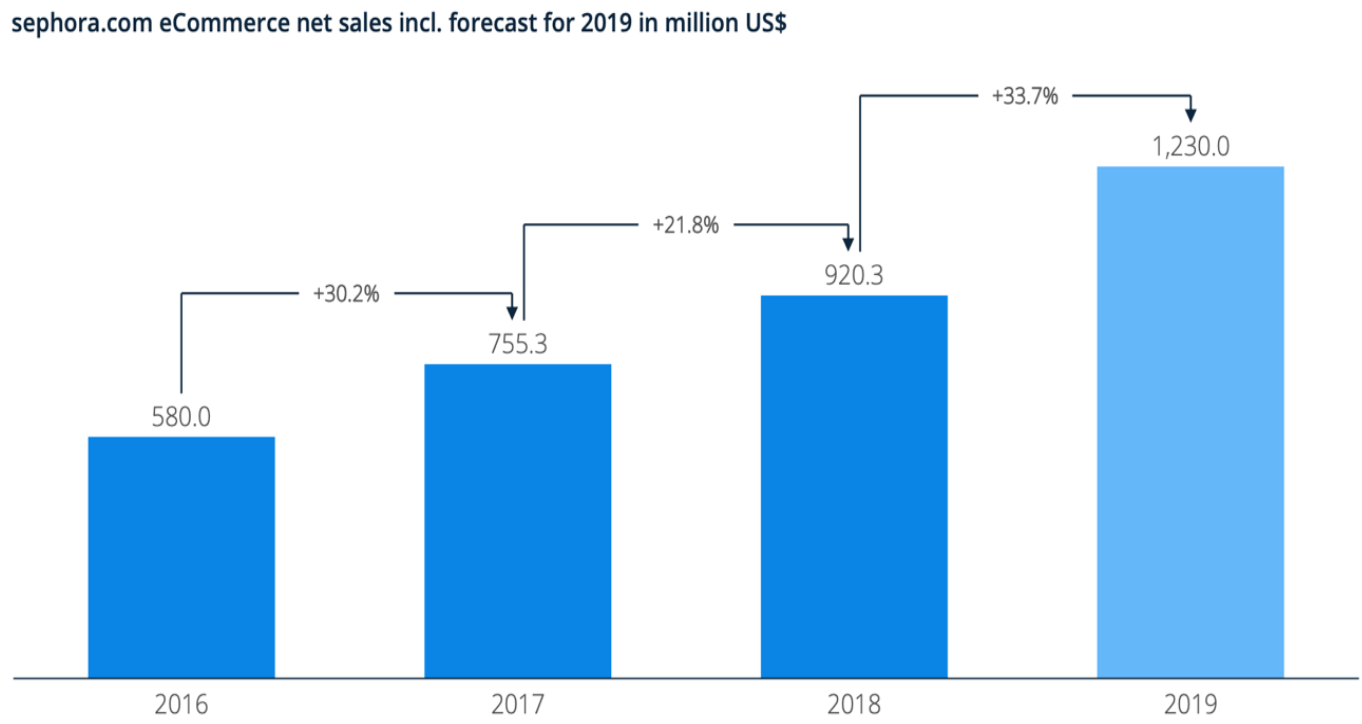

Figure 13: Sephora eCommerce revenues, forecast year 2019, by EcommerceDB

2.2.2.3. Alès Group (Lierac, Phyto, Jowaé brands)

Alès Group is a Group that operates in the cosmetic and wellness products sector. Its portfolio includes brands of the caliber of Lierac, Phyto and Jowaé(PatrickAlès, 2014). In 1989 the Group became international and began exporting to the following countries in chronological order: to New York, "then to Germany (1991), Italy (1996), Belgium (1997), England (1998), Canada (2001), Spain (2004), Poland (2005), Portugal (2008), Switzerland (2009), Greece (2010) and finally Hong Kong (2011)". (C. de Girval, 2019).

Over the years, the group has incorporated various entities to either acquire know-how or to strengthen its market position (Alès, P. 2014). However, since 2010, there has been a succession of strategic errors. This led to a plunge of sales in 2017 and a drastic slowdown in flow of inventories.

The graph shows the Group's sales from 2012 to 2017. Although there are no graphs showing the most recent cash flows, from 2017 to 2020 the company continued to lose ground to its competitors, recording large and constant losses. In 2017 the company recorded a loss of 15.4 million euros, in 2018 this loss amounted to 39 million euros, while in 2019 it even grew to almost 100 million euros (D. Chapuis, 2020).

In 2018, to make up for the deficit of the Lierac range, Alès Groupe is creating a new entry-level brand (price between €12 - €25), Jowaé, inspired by Korean cosmetics. According to Romain Alès, son of founder Patrick Alès, this low-cost segment represents 20% of the total market and is very promising (has already more than doubled in 5 years). Moreover, "in one year, in Italy, Joawé already has 1% of the market" (D.Chapuis, 2018).

There are two main reasons for these huge losses: firstly, as the activist fund CIAM suggests, which holds 3% of the capital of Alès the group has fallen behind in terms of digital and marketing." Having failed to grasp the importance of reaching its customers in the new distribution channel such as ecommerce, the company would have made a misstep. Moreover, even the support of technological tools was not fully exploited and this contributed to increase losses, as reported by Patric Puy, chairman of the management board of Alès (D.Chapuis, 2020)

Today the company is reinventing its business model. Due to the strong competition in the sector, Alés has decided to strengthen its presence through the use

of digital marketing and digital communication, introducing new technologies such as voice support and applications using augmented reality (D.Chapuis, 2018). As a matter of fact, as Profession Bien-Être explains in a recent publication, the use of social media and media that allow the user to interact seamlessly with brands are probably the strategies that will lead to victory in the long term (Profession Bien-Être, 2020).

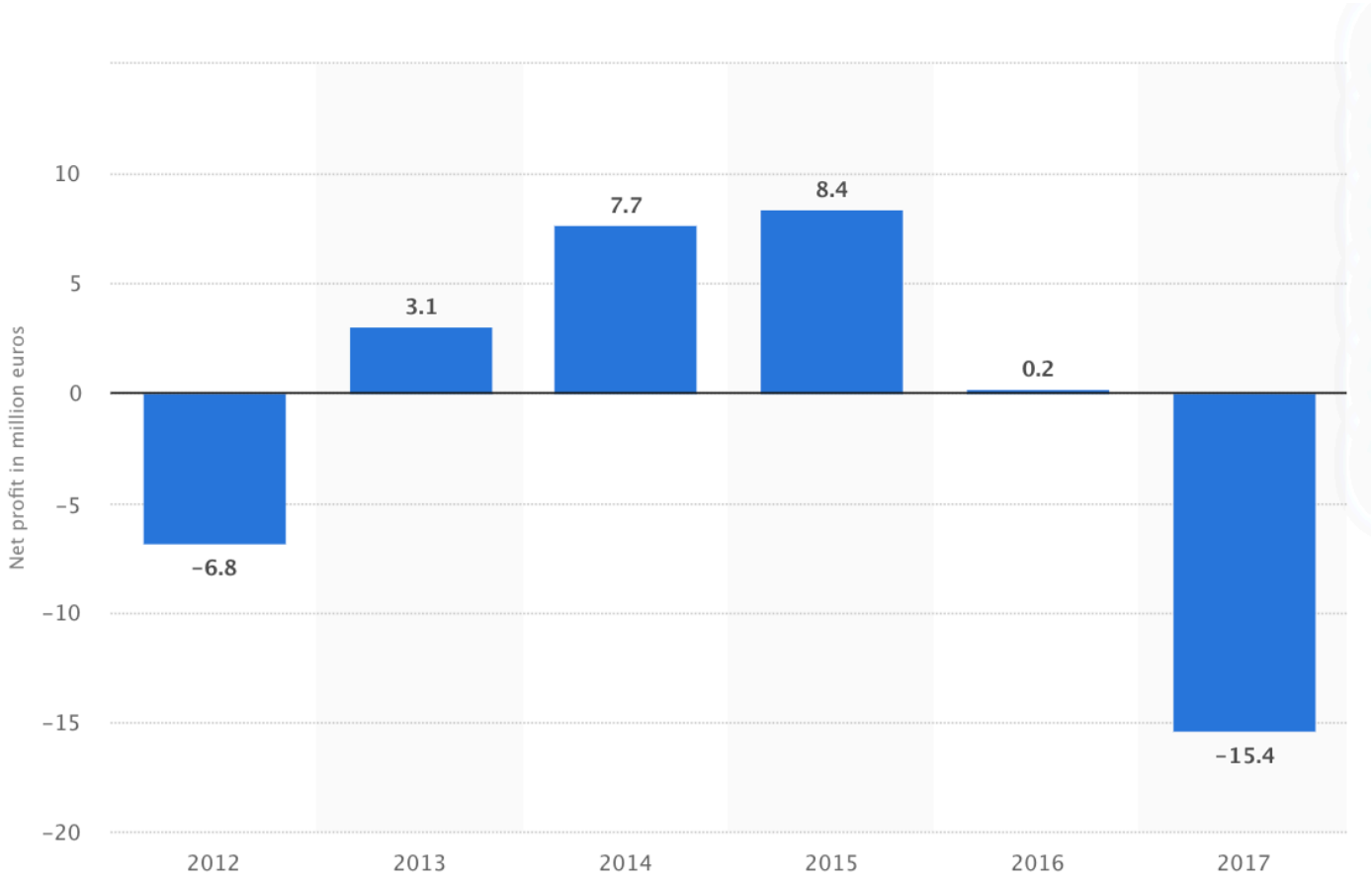

Figure 14: Net profit in million € for Alés Group, by Les Echos

2.2.2.4. Competitive Advantage

Having to make a comparison with Sephora, whose digital strategy seems to be consolidated and which is bringing benefits, and Alés group, whose failure to implement a digital strategy has led to considerable losses, there are some aspects that are particularly striking:

In 2020 it is important for a brand to establish a multichannel approach to its users. Since the majority of users gather information and buy online, it is necessary that the presence is consolidated on every channel available. As we have seen for Sephora, the brand provides information about trends and its products not only on its online website, but also through its Youtube channel. Through this, Sephora is able to redirect its users to its website providing a smooth and uninterrupted experience.

Creating an experience that goes beyond the purchase of a simple product is also the key to success. A customer who is included and attracted through the use of technological tools such as augmented reality and voice support will have an experience he will hardly forget and for this reason is more likely to return to buy from the same brand.

Moreover, as Sephora has shown, partnerships with brands from different sectors can bring a considerable advantage. The partnership with Google in fact, for which Sephora has promised to sell the smart speaker in its stores and ecommerce , will allow Sephora to be a privileged partner in having access to all the new technologies that Google will bring to the market, offering the brand a sizeable competitive advantage.

2.2.3. Healthcare

The third field that I want to take in analysis because of the great possibilities of development is that of medicine. The market of medicine has a global value of around 8 billion dollars (Walsh, E., 2019), which is why innovation is possible and is an attractive factor for the various competitors operating in this market.

Strictly speaking, talking about voice marketing in the field of healthcare services is probably reductive and somehow misleading. However, it should be seen

in a context where health services are also confronted with the problem of constantly improving the relationship between cost and quality, having to provide the community with the best services at the most competitive costs and therefore having to deal with the laws of the market, even for public institutions.

When you apply the concept of artificial intelligence to the medical sector, it is more natural to think of advanced robots that can perform difficult operations in which the human hand would have a higher rate of error (Russel, S., 2019). Nevertheless, the voice is a support that could bring an added value in the provision of health service, as stated John Browstein, chief innovator officer at Boston Childcare's Hospital, during an interview given to IT Healthcare news (Sullivan T., 2018). During the interview in fact, Brownstein explains that the use of voice support could be integrated in the whole process that a patient undergoes when he arrives at a clinical institution, from triage to the booking of the visit (Sullivan T., 2018).

In a Forbes publication written by Reenita Das dating back to 2018, a writer in the field of medicine for over 10 years (Das R., 2018), the journalist provides an overview of what could be the major trends that relate Artificial Intelligence and the medical field. One of the forecasts estimates that, with regard to voice interaction supported by smart assistants, the medical field has all the prerequisites to develop a sub-ecosystem in which the applications that could generate tangible added value are varied (Das R., 2018). We try to analyze them now through the support of a publication by some researchers in the medical field regarding artificial intelligence.

- **Medical report transcription**

The use of voice recognition in the medical field is useful first of all in case of transcription needs, says Zhao in his article "Speech-Recognition Technology in Health Care and Special-Needs Assistance". One of the disciplines in which a benefit from the introduction of speech recognition has been documented is radiology (Zhao X., 2009). Since radiologists find themselves in the situation of having to handle a high number of cases every day, especially in the case of chest examinations, explains the author, the introduction of this support has shown considerable benefits in terms of time and efficiency (Zhao X., 2009). Due to the high number of data that the doctor has to handle during examinations, it was considered necessary to introduce a support that could perform this data collection work in his place (Warfel T., Chang P., 2004). While in the past it was necessary for a user to manually check the transcription in order to avoid grammar or accent errors, nowadays the technology has reached a level where this process is no longer necessary (page 3).

- **Natural language processing to read medical reports**

Natural language processors are complex speech recognition tools that are able to read and analyze large data sets concerning the situation of various patients in order to capture elements that were not previously discovered or simply ignored because they were considered superfluous. Thanks to this support the algorithm is able to highlight structural problems that have not emerged during human analysis. This is an important tool that can lay the foundation for the discovery of new disorders and pathologies (ForeseeMedical, 2020).

- **Natural language processing to transcript conversations**

There are companies that are developing powerful tools that aim to transcribe the conversations between doctor and patient in different forms. The objective is obviously to report every single piece of information in as much detail as possible without the doctor himself concentrating on that task, devoting 100% of his or her focus to the patient himself. These transcripts are then recorded in electronic databases and can be used for future treatment and analysis in order to save time and have benchmarks (Mejia N., 2018).

- **Natural language processing to conduct AI conversations**

When it comes to conversation, forgetting the pun, there are several uses where speech recognition support can be useful:

- **Symptom control and information dissemination**

Through the use of chatbot you can recognize your symptoms without having to try to understand your problem through the use of high technical jergon. For example, a parent can introduce their child's symptoms in order to receive first aid tips and other useful advice (KeyReply, 2020).

- Medical triage and escalation according to the severity of the codes

There may be cases where the recognition of some symptoms through language recognition may favour some cases rather than others. If, for example, a

patient experiences pain in a particular area of the body that may be a more serious intrinsic problem, the Recognition System may facilitate the case by giving him/her a more urgent code. If we take for example the recent case of hospital overcrowding during the Covid-19 pandemic crisis, all those cases that have been given a code that is too high in relation to the real risk, this problem could be used through the use of the support just described (KeyReply, 2020).

- Reservation of a specialist visit

Through the use of intelligent support, it would be possible to book specialist visits without the need to hire additional staff. This could help to reduce costs in order to use the budget more efficiently (KeyReply, 2020).

- Engagement with the patient after the visit

Being able to monitor the patient after the visit is also an important and very often overlooked aspect. However, monitoring the patient's condition after they have been given medication, sometimes with side effects, is another benefit (KeyReply, 2020).

2.2.3.1. Boston Childcare’s Hospital Case Study

During 2016, Boston Children's Hospital released an application compatible with the Amazon Echo Alexa smart speaker called KidsMD. The purpose of the service is to offer support to parents by providing medical advice regarding their children's condition in order to ensure constant support without the need for a doctor directly, especially in less serious cases (Mutchler, A., 2017).

In terms of measurable results, as reported by Devin Nadar, partnerships manager for innovation and the digital health accelerator at Boston Children's Hospital, the application has received more than 100,000 patient interactions, an increase of about 200 active users per week (Metrock B., 2018). With regard to use by physicians, only 16% of respondents stated that they do not support the use of voice in applied medicine (Lindenauer S., 2018), 48% would be willing to deploy the technology in their clinical setting, 36% were undecided, many of whom cited a lack of familiarity with how the technology could support patient care (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2018).

The application had been in use for almost two years at that time and this activity demonstrates the interest and satisfaction achieved by users. In addition, given the increase in sales and market share obtained by Google, Nadar tells how in reality, even though the application was made usable only through Amazon Echo Alexa, in order to extend the support of Boston Childcare's Hospital to more patients, the company was also thinking about including Google in its business model (Metrock B., 2018).

However, the application was only defined as an initial step in introducing the use of voice within the business model adopted by the hospital. In fact, Matt Murphy, innovation lead at the institution's Innovation and Digital Health Accelerator explains that the plan is to integrate more functionality to improve the support provided by Boston Childcare's hospital (Mutchler, A., 2017).

Boston Children's Hospital is currently using the various functions of Alexa in order to help its physicians with specific protocols to follow, doses for each drug and to better organize the work between the various staff (Mutchler, A., 2017). These are some of the examples that help to reduce research time. In addition, again with reference to the reduction of work time, Boston Childcare's Hospital has developed a feature with which you can track the various machines located in different rooms (Mutchler, A., 2017).

One of the major benefits that doctors have reported following the insertion of voice support, as has also been explained earlier in this text, is the ease in searching and analyzing information within the hospital's databases (page 43) (Mutchler, A., 2017).

Through the support of the Harvard publication mentioned above, we try to identify the fields where the use of voice support is really bringing value:

- **The intensive care unit.**

In the preparation of an operating room there are several needs, some related to the safety of the operation, such as the sterilization of instruments and infection control, and others more related to the organization of procedures, such as always having immediate access to all the information you need. Childcare's Hospital in Boston developed an accelerator in 2016 to help doctors with these needs (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2018). However, even though the process was still in its infancy, the hospital said the accelerator was already successful, showing that its inclusion in the hospital's business model could bring incredible

benefits in the future. In its first year of use, the accelerator supported more than 300 doctors, researchers and administrators, as reported in a Harvard Business Review publication, being incorporated into about 25 departments within Boston Childcare's Hospital (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2017).

In addition, she has contributed to the improvement of 9 projects, three of which have become startups raising more than $2 million in venture funding (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2017). In general, IDHA's voice deployment in the Boston Children's ICU allows physicians to obtain logistical and administrative information that can be enormously supportive when one considers the time it takes to obtain it with a traditional approach. In fact, by asking voice support questions like “which room can be used for patient X” or “where can I find the machine for a heart analysis?”, Doctors don't need to make phone calls or search the hospital records themselves. In terms of benefits, they have said that the main advantage is the reduction in the time needed to perform that operation, sometimes even minutes, but which can prove crucial in an environment like health care. (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2017).

- **Organ transplant.**

Boston Children's is one of the pediatric surgical hospitals where thousands and thousands of transplant operations are performed every year. With the support of IDHA, the team has been able to improve and speed up the entire process of evaluating available organs prior to an operation. The result achieved through voice command support, as stated by a physician not mentioned in the article, was an effective streamlining of the waiting list and donor check-up operation (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2018).

- **Home health decisions**

Voice interaction, as we have reported on page 40, can also be particularly useful in the home environment in terms of decision support. As described in the introduction to the paragraph, KidsMD, currently available on the Amazon platform, provides first aid information to parents seeking advice on caring for their children. The information is taken from the Childcare's Hospital database in Boston and communicated directly to the user without the need for direct interaction with doctors, especially in mild cases (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2018).

One of the problems that initially emerged was related to the background noise related to the device, as we saw during the presentation of the smart speaker Amazon Echo Alexa. However, this gave the management the opportunity to optimize the use of the voice assistant in the most efficient and necessary places, so that the assistant did not provide misleading information (Comstock J., 2016).

In addition, as revealed in an interview conducted by Harvard Business Review, the biggest problem found is actually attributed to the security of the information collected. Of those who participated in the interview, 55% said they were more or less convinced of the reliability of the information provided by voice support, while only 7% said they trusted the information blindly (Small E. C., Nigrin D., Churchwell K., Brownstein J., 2017).

In terms of actual patient support, however, Boston Childcare's hospital had to wait until it complied with the data protection regulations. In fact, in 2016, voice platforms were not yet HIPPA compliant, so it was not yet possible to access patient data through voice command support (Mutchler, A., 2017). This compliance came in 2019, when Amazon announced that it had developed a service from Alexa that was compliant with current regulations regarding the processing of sensitive patient health data (HIPPA), a critical point to ensure the effectiveness of voice support. Six institutions have developed voice support to help doctors do their work, including Boston Childcare's Hospital (Kinsella B., 2019).

2.2.3.2. The Italian Healthcare System

Although the integration of technology into the hospital system has already achieved significant goals in countries such as the United States and China, in Italy the process seems to be slow and full of obstacles. However, as we have just analyzed in the case of Boston Childacre's Hospital, these goals could be strategic in terms of digitization and innovation in healthcare and eventually result in better services and lower costs for the patient.

In the case of projects that help patients when they are not physically at a clinical institution, as in the case of KidsMD seen on page 53, from 2018 to today it seems that only 45 hospitals provide adequate support, and in any case not through a smart device that provides advice through voice interaction. In fact, with regard to the integration of Artificial Intelligence as a support for doctors as well as patients, during an analysis in 2018 it was found that only 5% of the facilities that had taken

part in the national survey had such equipment available. In 2020, through the same analysis, the realities that provide such a service rose to 16%, still too few considering the potential they could bring (NetConsulting cube, 2019).

Strongly related to this matter is the one related to BigData, fundamental pillar in order to efficiently use Artificial Intelligence. However, even in this case the structures analysed which have databases that could speed up the information search processes are only 7% in 2019, with an increase to 15% in 2019 (NetConsulting cube, 2019).

With regard to the question of booking visits, however, simplified by the introduction of smart assistants in the case of Boston Childcare's Hospital, this service is guaranteed only in writing and in any case only 11% of the companies taken into consideration (NetConsulting cube, 2019).

The concept of voice insertion, on the other hand, identified as an optimal solution to ensure security in the exchange of information between patients and doctors, is still seen as insufficiently developed and reliable to be integrated to bring value. Although the aim is precisely to improve communication between the two parties and efforts are being made to find an appropriate solution, it seems that ideas on how to do this are not yet clear. This is a concept of the utmost importance and relevant to the delivery of quality services. However, in order to build a process of digital transformation it seems necessary to create a general committee to oversee the various processes in order to develop in the right direction. The report mentions the need to create targets that are measurable through KPIs in order to keep track of moves and their development as well as the costs and the brought results (NetConsulting cube, 2019).

The objective would therefore be to measure not only the actual result obtained by the patient, defined as PROMs (Patient Reported Outcome Measures), but also as PREMs (Patient Reported Experience Measures), the real experience lived by the patient.

However, when asked about the degree of satisfaction achieved after using certain services, first of all receiving advice on medicines or distance visits, only 22% answered positively (NetConsulting cube, 2019).

2.2.3.3. Competitive advantage

We analysed through the presentation of the two opposing cases, the excellence of Boston Childacare's Hospital and the inexperience of the Italian healthcare sector, how the introduction of technological systems (and in particular voice enabled systems) can bring concrete benefits to support the routine of patients and doctors, to provide better services and to bring costs down.

In terms of the reduction in time of the various processes, it has been analysed how the support of a smart assistant can help doctors and patients to retrieve information in a shorter time compared to traditional methods. Considering the importance of the time factor within a health clinic, this is probably one of the most important benefits that can be achieved. In addition, if we take the KidsMD application case under analysis, parents can get in touch with the smart assistant in order to obtain information about the first aid care to be given to their children, avoiding to physically present themselves in the various institutions. As we experienced during the recent Covid-19 pandemic where delays in operations were also due in part to unjustified hospital crowding, this support may avoid such circumstances in the future.

As far as cost reduction is concerned, on the other hand, it will be possible to save on all the non-essential personnel used for secretarial services and visit bookings as they will be replaced by smart assistants and virtual databases. It will therefore be possible to invest this budget for more critical and necessary resources to carry out the main services that are provided in care institutions, such as research and machinery.

In addition, as presented in the report provided by the Healthcare System of the Italian territory, the introduction of voice supports, if well compliant with the regulations in force regarding the protection of sensitive data regarding the patient's health, will be able to guarantee greater security in relation to patient data.

2.2.4. The way ahead: why market leaders must integrate language technologies in their business models

Due to the relentless increase of smart speakers in consumers' homes, one thing is now certain: the use of voice within our daily routine to interact with electronic devices is an established trend that will just increase in the future. As a consequence, it will change the way brands come into contact with their prospective customers and will require a forward-looking strategy to take full advantage of these technologies and win larger market shares.

As we analysed earlier, there are multiple reasons why brands need to fully embrace this new trend. Let us summarize them below:

- **Voice technology will continue its growth**

Since its introduction, voice interaction has simplified the lives of its users. Even if initially the technology used was not as effective as it is today, having achieved an accuracy rate in excess of 95%, as we have seen on page X , the trend was established and more and more processes became leaner faster, simplifying the lives of an increasing population of users. While in the beginning these smart assistants were only available on media such as mobile phones and computers, which were offered in the complete package with its hardware, the fact that today there is a whole market dedicated to smart speakers and assistants says a lot about the potential that these devices can achieve. As we write this text, there are more than 45 million voice enabled devices in America that are revolutionizing consumer habits. In fact, as reported by the Digital Marketing Institute, 50% of research prior to purchasing, in 2020 will be done through the use of voice (Digital Marketing Institute, 2020).

- **The introduction of voice will revolutionize the consumer experience.**

It is wrong to think that introducing the use of voice technology into a business strategy is just a way to keep up with your competitors. In fact, through this innovation brands have the opportunity to completely revolutionize the consumer experience, which as we have seen on page 35, has become even more important than the product itself. Since the various smart speakers will keep track of past interactions, both positive and negative, it is important that brands are positioned

in the right place at the right time, and above all that they are ready if they want to develop a certain degree of consumer loyalty.

In addition, the time factor is also crucial, as described above. Our lives are much more eventful than in the past and you don't even have time to look for something in writing anymore. On the contrary, voice is a tool that allows us to get right to the heart of the matter, saving time and being more precise.

- **The introduction of the voice will change the way we look for information.**

Another reason why brands need to adapt to this technological revolution is because when we talk, we express ourselves differently than when we write. As we have seen previously on page 33, there is a substantial difference in the number of keywords we use when we express something by voice or in writing.

Also, the accuracy with which we can express concepts changes from written to spoken, and brands will need to keep track of these differences when trying to position themselves in the results list of a research in internet. In fact, the first results of the research will be the ones that will reach the highest number of hits and it is therefore necessary that the strategy used allows to reach the best positioning. If when we search with the mouse and keyboard we do not go beyond the first page of results, from which also derives the joke "the best place to hide a corpse is the second page of Google results" (Fry, H., 2018), through the voice search we will not even go beyond the first pair of results, since it is assumed that they will be the ones that best fit our search.

- **Customers will go to your competitors if they don't find you.**

As far as positioning with respect to competitors is concerned, it is natural to assume that if potential new customers looking for voice enabled platforms will not find you, they will go to competitors who have successfully made the new technology available on their IT platforms. Since it is already difficult to win the competition for a new customer when it is possible to be noticed with the same chances as your competitors, not taking part in the competition can certainly have negative effects on performance and brand presence.

- **Learn from the best**

The best way to integrate voice marketing into your business model is to take inspiration from those who have already done so. We have previously offered an overview of the best brands in the world that have already integrated voice marketing into their business model or are in the process of doing it. We've seen how Boston Childcare's Hospital is able to provide remote support for parents who are in doubt about how to care for their children, we've seen how Sephora captures new viewers and potential clients through a mix of technology such as smart assistants and video streaming, or how Bank of America is able to make its clients' and partners' transactions more secure and tailored to their needs. There are still many areas where voice support can become an important competitive advantage for those who will integrate it effectively into their business model, and a long-term vision is key to understanding and capturing them at the right time.

3. A practical case: language, voice technology and adaptive learning applied to the business of learning

This part of the thesis will be developed in two steps. In the first one I will present two companies that provide educational services and have already integrated voice enabled systems into their business model. In the second I will propose a business model which integrates an innovative AI technology learning adaptation - and I will explain how this can bring an additional competitive advantage. The objective of the project is to teach foreign languages in an innovative way and to reach this objective it has built its strategy around the integration of both of the technologies mentioned here into its IT platforms.

Teaching with the support of digital media has received numerous awards and considerations because of the value it can deliver to the learner (Cit) The integration of technological media into the various teaching methods significantly improves productivity, increases engagement and motivation and ultimately accelerates the process of learning. (U.S department of education, 2020). It is widely accepted that these technologies can be actively used in both teaching and learning contexts (U.S department of education, 2020).

In the field of voice recognition, some tools have been developed offering support in the field of active learning, a system through which the student interacts with the material provided through reading, writing, communication and reflection (Carr R., Palmer S., & Hagel, P. 2015). These tools, Instant Note Capture (INC) and Virtual Oral Recitation Examination System (VORE) are built around AI technologies and allow the student to interact with the System with his own voice. As Dittrich Tobya and Star Sequoia explain in their publication for Portland Community College "Introducing Voice Recognition into Higher Education", Instant Note Capture (INC) allows the student to live a totally new learning experience, making classes more dynamic and interactive (Dittrich T., Sequoia S., 2018). The Virtual Oral Recitation Examination System (VORE), allows the teacher to examine and evaluate student performance based on keywords (Patent No. 7,296,218 B2)(Patent No. 7,562,288 B2).

3.1. Information gathering about companies or institutes which have already implemented voice marketing in providing learning services

If we look at companies or institutes that have integrated the use of voice support into their business model, we can find some interesting examples of application of the ideas we will analyse afterwards. For some of them we will also present a business model canvas with a focus on the key elements they propose.

- Britannica

Encyclopedia Britannica has been innovating the field of learning for over 250 years. With the support of the various services that Britannica offers, it has gained a strong reputation and a primary place in the learning business. In fact, by providing products such as curriculum products and professional readiness training, Britannica has helped to develop new methods of teaching and studying not only for students, but also professors and professionals. As explained in a District Administration article, there are many professors interacting with students with smart media that can be activated by voice control (Crist C., 2019).

The products developed by the company specializing in learning are of all kinds, from video games to video and storytelling, all of which can be controlled through the use of one's own voice. The various products are reviewed on Amazon, as they are powered by Amazon Alexa, with a very positive feedback from users (Valenza J., 2019).

One of the great achievements of Encyclopaedia Britannica is undoubtedly Guardians of History, probably inspired by the name of The Guardians of Galaxy, a voice-activated video game that allows students to learn about the events of the past through exciting adventures. Thanks to the possibility of interaction with the video game, students are favoured in the development of skills such as how to make decisions and more theoretically on the knowledge of the great events of the past, as reported by School Library Journal (Valenza J., 2019).

Each game session includes approximately 40 minutes of recorded lesson time plus 20 minutes of player interaction. These are actually variable because being this an interactive game, the duration is influenced by the player's own choices. In fact, as Ryan Bond, Britannica's Director of Digital Consumer Products,

explains during an interview with Joicy Valenza, the game is structured to follow a tree-like pattern in which each decision involves a succession of other related decisions (Valenza J., 2019). In addition, the player is provided with additional support tools such as primary resources and free e-books through which he or she can learn more about the concepts he or she has just learned (Valenza J., 2019).

"Voice experiences and games are two of the ways that technology, used effectively, can help enhance learning inside and outside classrooms," says Karthik Krishnan, global CEO of the Britannica Group. "Britannica is excited to blend modern-day technology and science fiction concepts to continue to inspire curiosity and the joy of learning in new and engaging formats" (Britannica Digital Learning Website, par 5).

Encyclopedia Britannica also provides a small practical guide in which it explains concretely in which areas it has been possible to introduce voice support in order to promote learning.

Basic Skills

Math: In the field of mathematics there are several ways in which the smart assistant can provide support to the student: for example, it is possible to ask for the generation of random numbers, random events, and to ask for the correction of simple exercises.

Language Arts: Through voice support it is possible to improve communication and vocabulary knowledge in some languages.

Weather: It is also possible to teach the weather forecast and vocabulary to describe it.

Classroom Tools

Music, Time and Reminders: The teacher can create a pleasant atmosphere in the classroom by setting alarms, alerts for students and soft music to facilitate study.

Custom Learning Skills

"Alexa Skill Blueprints": through the Alexa Skill Blueprints service it is possible to create personalized programs and games with which students can interact with each other with the support of a smart assistant. This can be useful to store notions, new words and foreign languages.

Listening and critical thinking

Voices: it is possible to ask students to recognize the various roles within a voice track, such as interviewers, narrators, presenters.

Sounds: it is possible to have students identify the sounds of everyday life.

- Babbel

Babbel is a foreign language teaching platform available on mobile devices and computers. The service is subject to a fee and it is possible to register with different subscription formulas depending on the duration of use, which ranges from one month to one year (Babbel, 2020). Babbel has been recognized as one of the most innovative companies operating in the field of learning by experts such as TechCruch and Fast Company. The application has more than 1 million subscribers worldwide and is ranked first among foreign language teaching applications (Online Course Report, 2020).

In terms of product, the company has been good at adapting to the needs of its users, developing 15-minute mini lessons that can fit into the schedules of anyone with a real desire to learn a new language. (Netguru, 2020).

When we look at Babbel's efficiency, it can be said that the company has done an excellent job in terms of making the product efficient and delivering good results. In fact, according to a study by the University of New York City that looked at a group of inexperienced students struggling with learning Spanish, it was found that 73% of students were able to have a conversation in Spanish after just 15 hours of class (Bradley L., Al-Sabbagh K., Vigmo S., 2018).

Furthermore, from a target audience point of view, while initially the company focused on a B2C service, where users were normal consumers, Babbel also turned to the B2B market, developing relationships with brands such as Huyndai, Paramount, N26, Bacardi, Pirelli, and Acer (Netguru, 2020).

However, if we look in detail, as explained in an Online Course Report article released in 2020, although Babbel can be considered a leader in this market providing teaching experiences of considerable value, it also has some flaws:

- **Lack of curricular direction**: Although it may seem that the offer of a multitude of programmes and curricula could be considered a strength, the insufficient structure and organization of the proposed programmes is a weakness. In fact, since the platform does not follow a precise structure adapted to the student's needs, the student may find himself disoriented as he continues his studies.

- **Limited explanations:** From the point of view of grammar, often an important obstacle in the learning path of a foreign language, Babbel does not offer a real error correction but rather an option similar to "try again, you will be luckier".

- **Speech recognition is weak**: Although the most effective form of learning learn pronunciation and vocabulary is dialogue, Babbel has a really inadequate speech recognition system when you think about the ambitious goals the company sets itself. In fact, if you think that the online language learning market will reach 20 billion in 2026, as Ola Prejs explains in an article for NetGuru (Prejs O., 2019) in order to remain leader of the market and be able to seize an even larger market share, Babbel will necessarily need to invest heavily in improving its Speech Recognition System to be able to capture all the information spoken by users much more efficiently (Online Course Report, 2020).

Personally, after 6 months of using the Babbel application I can provide my personal analysis of this tool. As far as the pros are concerned, I think the application definitely has a good entry price. In fact, after having paid €59 for a 6-month subscription, I was given an additional 6 months with the aim of establishing a well-defined relationship with the brand. The application has a very simple and intuitive interface. The programmes are divided according to the various levels, beginner, advanced and expert, and offer a range of topics that allow you to express yourself on concepts concerning the daily routine and trends of the moment. From the point of view of grammar, I agree with the analysis provided by Online Course Report, because I realized that there are actually no clear explanations regarding rules and forms to use. However, from the point of view of the efficiency of the learning process, I must say that through Babbel I learned in a relatively short time fundamental concepts and vocabulary in the French language, which quickly improved my command of the language and allowed me to interact more easily with the various interlocutors in everyday life. As far as speech recognition technology is concerned, I agree that Babbel can certainly improve its systems in order to reduce the error rate and offer an adequate service. Many times, I have seen my pronunciation rejected for no reason and this factor risks making the user nervous. In addition, I believe that

offering the same standard teaching system for all the users is not the right approach when entering the world of education. I strongly believe that every subject is unique to himself and for this reason it is necessary to adapt the program offered to the user's skills. This would improve the learning process and ensure success for all users. In the following page it is shown the Business Model Canvas adopted by Babbel.

3.2. A proposal of a business model for the provision of learning services actively integrating voice marketing

After a careful analysis of the services that are now available in the field of education via online platforms, it has emerged that even the most widely used of the applications, Babbel (page 56-57-58), has failed to integrate a sufficiently high level of technology to enable the user to exploit the full potential that voice technologies can offer. I am convinced that the prospects and possibilities for development in this field are huge and that is why I have developed a project that, based quite substantially on the use of the various language management techniques and in particular voice recognition, would be able in the future to establish itself as a major player in this field and gain a real competitive advantage over all other competitors.

First of all, I think that the project should include collaboration with some strategic partners, especially in the field of technology. As we have seen in many parts of this dissertation, the two leading brands in the smart speaker and voice recognition market, respectively Amazon Echo with Alexa support and Google home assistant, have reached a level of accuracy close to 95%. Such precision would help to offer a technologically more advanced service than that of competitors. In addition, these two brands are also in a position to allow access to an unlimited amount of information that would help to offer the most detailed and personalized user experience possible. On the other hand, I also believe that such collaboration would bring a competitive advantage to both parties. Indeed, as we have seen in the case of Sephora, Google or Amazon could have an interest in entering the world of education and through the partnership this would be possible.

I then consider the possibility of using some speech recognition-based systems, such as the Instant Note Capture (INC), which allows for more dynamic user interaction, and the Virtual Oral Recitation Examination System (VORE), which allows for user assessment based on the keywords spoken. These technologies are currently under patent protection but would also be accessible through a partnership agreement.

A further advantage would also be the introduction of an additional technology called adaptive learning, as presented on page ( ) . In practice, the smart speaker would be able to adapt to the student's level, offering personalised communication on learning skills and learning speed.

3.2.1. Specific project with Voice marketing applied to learning

In this paragraph I will explain how the project would take shape. As we have seen in the case of Encyclopedia Britannica, it is possible to develop an application in co-ownership with Google or Amazon that interfaces with the end user. The user can buy the smart speaker and then download the application inside it. Different pricing plans will be available, depending on needs and availability, and a trial period will also be provided during which the user can verify the efficiency of the service.

In practice, the smart speaker would function as a private teacher, equipped with all the information present in the Google or Amazon database and through them can offer detailed lessons on each topic. Through collaboration with the most established teaching institutions on the market (e.g. British Council, Alliance Française, Instituto Cervantes), the various teaching modules will be integrated and accessible through voice interaction.

Through the two patent protected systems, INC and VORE, the smart assistant will be able to provide an assessment to the student based on the keywords spoken. Since, as we have seen, the smart speaker is able to offer customized information based on our past interests and interactions, the smart speaker will be able to identify the topics where the user has encountered the most difficulties in order to deepen them and improve the student's performance.

Moreover, an additional level of customization can be achieved by integrating a learning adaptation system that allows the smart speaker to adapt its skills to those of the student. Since each subject has learning times and techniques that differ from those of any other, I think this feature is necessary in order to be able to offer a service that differs from the standard services available on the market.

3.2.2. Critical analysis of the business model and expected benefits and results

My analysis of the substantial benefits of the proposed system will be set out here. There are a number of elements in particular which will make it distinctive from existing systems on the market and offer tangible added value to those who will use this service.

- Personalised learning

The user will have the possibility to obtain a fully customized service according to his needs and capabilities.

- Abundance of information

The application will have the possibility to access a wide base of information and documentation, coming from the most authoritative sources, and to be updated through a network of interconnected information available online. This will reduce maintenance costs and shorten waiting times. At the same time, the service will offer new and more up-to-date functions.

- More in-depth student involvement compared to traditional methods

The application will make the study of the language a new experience that has never been made available users until now. With the use of the most advanced language technologies and in particular the smart speaker the student will be totally involved in his learning sessions, with a real interaction and dialogue with the system. With the simulation of real conversations, the improvement of the fluency of conversation and pronunciation will be much faster and more substantial than conventional online learning systems already existing on the market.

- Making students more aware of their own abilities

Through the method of evaluation on the basis of recognition of key words the student will become aware of his or her progress in learning, of the notions he or she already possesses and uses correctly, and on the other hand of what are his or her shortcomings and the notions he or she needs to review or deepen.

3.2.3. Competitive advantage that the model can bring about

- **New method compared to competitors**

According to my assessment and the analysis of the systems already in place, I believe that this project can have considerable added value because of its uniqueness. In fact, there are currently no platforms on the market that provide a learning model totally based on voice interaction and that allow the user to simulate a conversation with a counterpart able to respond on a human level. This feature in particular will be able to stimulate great interest from users and in particular from those who have not been able to obtain appreciable results with traditional systems. The innovative features, if well supported by synergies with technological partners, will also be able to attract new users who are not looking to improve their language skills but who will feel attracted by the novelty of the system. This will allow for the acquisition of ever-growing parts of the market to traditional competitors.

- **New technology compared to competitors**

From the point of view of the technology adopted, a system able to adapt to the student's skills and to simulate real interactive sessions is probably the closest to the accompaniment of a private teacher. An appropriate pricing policy will make the system clearly competitive with the cost of a private teacher. With the constant progress of language technologies in AI, the features and performance of the system will be able to evolve constantly offering new functionalities to the user.

- **More information than competitors**

The possibility of being able to rely on content managed by Amazon and Google and to multiply synergies with these major technological players would make the platform constantly updated.
The strength of these IT giants is the huge number of users who access their platforms on a daily basis. This offers considerable potential to reach a large number of contacts. This would also open up the possibility of advertising actions within the platform itself, increasing profits.

4. Conclusions

As explained above, this thesis aimed at analysing the results achieved so far in the integration of voice support technologies within the applications commonly used by consumers, mainly to acquire goods and services. As the consumer's purchasing experience has been recognized to be even more important than the product itself, due to the increasing homologation of products and standards to be met, suppliers of goods and services must strive to introduce techniques and tools that make the consumer experience increasingly attractive and rewarding. The introduction of AI technologies such as smart speakers and smart assistants into the purchasing platforms of the most competitive and cutting-edge companies is one of the most interesting examples of these trends.

As we saw during the first part of the literature review, the study of these techniques and their application in various fields of business and especially marketing, was prompted by the observation that devices such as smart speakers and smart assistants are increasingly found in our homes. Thanks to the continuous improvement of language processing and speech recognition technologies, which thanks to the progress made by AI in this field, now reach accuracy rates of over 95%, these supports are increasingly used practically daily. Major IT giants such as Google, Amazon or Facebook invest enormous sums and employ their brightest minds in the development of these technologies.
Feedback and user satisfaction are constantly increasing. As time goes by and the number of "IT literate" users increases, these tools will progressively become ubiquitous in many fields of daily activity.

The most competitive and cutting-edge companies in their fields have long understood that these technologies are necessary to remain market leaders and try to increase their business volume.
As proof of this, we have taken into consideration the best achievements made by some leading companies in their field. The analysis allowed us to determine what are the characteristics of the systems implemented by these companies, what are the advantages of these systems and how these characteristics can be transposed into similar systems in other business sectors. All complemented by timely analysis of the benefits and competitive advantages that can be expected.

We analysed how the use of smart assistants and voice enabled devices in the banking sector has contributed to improving performance in terms of customer transactions and in particular investments. By receiving advice tailored to their

needs, thanks to the incredible amount of data these applications can draw on, Bank of America's clients were able to record a significant increase in profits from targeted investments and deposits. The obvious result has been a significant increase in the bank's profits.

In the healthcare sector, the case of Boston Childcare's Hospital highlighted the savings in time and money from which both institutions and patients are able to benefit.

The analysis then moved on to the sales and retailing sector, which is probably the area where the introduction of voice marketing has already produced the most remarkable results and yet has the greatest scope for progress. It was found that the introduction of voice support in online sales channels can increase a company's business volume, offering greater presence and visibility, a greater commitment to satisfying consumers' wishes and thus guaranteeing their loyalty.

Finally, the field of online educational services was analysed. It was examined how two major players in online education, British and Babel, have already integrated voice support into their educational services.

The systems developed by Britannica and Babel have been able to revolutionize the online education business, providing a completely new service compared to competitors' offerings. Considering that their platforms are still far from perfect, the margins for improvement are enormous and promise even more remarkable future developments.

Based on the observation that despite the successes achieved by these two companies there is the possibility of creating even more powerful and competitive platforms, we presented a project for a platform for the provision of online language learning services.

We tried to provide an example of a business model that would push the use of AI based language technologies even further forward. In fact, one of the winning points of the project is the use of even more innovative and reliable technologies, as this was reckoned to be the main weakness of Babbel, a leader in the field of foreign language teaching through an online platform. Moreover, the addition of an Adaptive Learning System that allows the platform to adapt to the student's needs is also a significant competitive advantage and provides benefits in terms of cost and effectiveness of the learning process.

It is necessary to clarify that, since one of the problems related to the adoption of these technologies is related to privacy with regard to the protection of

users' data, the system must take into account the international regulations governing data protection and in particular the General Data Protection Regulation of the European Union.

As the subject matter is relatively young, the information provided is often accompanied by assumptions and forecasts, which as such can be contradicted in the future.
However, if the past can serve as a guide to predict the future in some way, it is clear that the progress made so far and the benefits to the companies that have been forward-looking, there is a good chance that all these predictions will come true, in a more or less short time.
If anything, the part that is most uncertain for progress in these areas is the legislative and environmental context that will have to be developed in order to reconcile the needs of the consumer and civil society with the economic objectives of companies in a reasonable way.

Finally, we must remember that the exceptional social context in which we find ourselves due to the health emergency did not allow us to develop the theoretical project in sufficient depth, since the experts in the fields involved that we intended to consult were unavailable.

For this reason, the business plan example presented is basically a theoretical model that has not been directly confirmed by the specialists in the field, who could not even enrich the project with further indications and advice. However, we believe that the project, based on a number of assumptions and consolidated data, contains all the prerequisites for a successful outcome.