# CONTENTS

DIFFERENT PERSPECTIVES 4

PREAMBLE 7
INTRODUCTION TO CHATBOTS 8
CHATBOT AND CONVERSATIONAL AGENT? 9
THE HISTORY OF CHATBOTS 14
THE FIRST AREAS IN WHICH CHATBOTS HAVE BEEN USED 16

HEALTHCARE CHATBOTS 22
ILLUSTRATIONS OF HEALTHCARE CHATBOTS 24
Medical information 26
Prevention 27
Pre-diagnosis and referral 27
Hospital and outpatient care 28
Support for chronic diseases 29
Compliance and information on treatment 30
Conclusion 31

SPECIFICITIES OF HEALTHCARE CHATBOTS 34
The Data 34
The Personality 36
The Technologies 37

EVALUATION OF HEALTHCARE CHATBOTS 40
Assessment update 41
Methodological approach 42
› Illustration by a study of acceptability and how the innovation is perceived 42
› Key factors in the evaluation 45

CONCLUSION 49
We opened 39BIS, our e-health laboratory, in December 2017, with the aim of accelerating our e-health projects, to design, develop and demonstrate the value of e-solutions for patients, caregivers and healthcare professionals.

In the course of the e-health projects that we are carrying out at 39BIS, we wish, with the «39BIS Files», to share our experiences and all that we have learnt with the various stakeholders in the sector, in a spirit of collective intelligence. In this way, we want to help enrich discussions and establish good practices for the benefit of all, in a field that is in full swing but for which there are still many issues to be resolved.

This new «39BIS File» on «Healthcare Chatbots» presents our first feedback on the development of a chatbot for healthcare. After having undertaken a preliminary study at 39BIS on how to reinforce the proper use of so-called «complex or restricted» prescription drugs, we began developing a chatbot (or conversational agent) in 2018, since we are convinced that the technology can address this issue.

We are currently pursuing the development of our chatbot as an experiment, jointly with healthcare professionals, and have chosen two partners with complementary expertise: Orange Healthcare and the start-up, Kap Code.
The word “chatbot” comes from «chat», which means informal conversation and «bot», short for “robot”. Chatbots are software that simulate a human conversation such as Siri for Apple or Alexa for Amazon. The interaction between people and chatbots can be oral and/or textual. To respond appropriately to a user request, the chatbot must be able to understand the meaning of the question, then find the answer and deliver it.

While this technology first experienced a strong boom in banking and commerce, the use of chatbots is beginning to emerge in healthcare and can provide many services to both health professionals and patients. There is a wide range of potential applications, including medical information, prevention, outpatient follow-up, and patient support or pre-diagnosis.

The work we have already done and our discussions with health professionals suggest that the technology could be put to good use on a routine basis, particularly in the search for treatment information.

As the development of healthcare chatbots has only just begun, there is, to our knowledge, no authoritative book on the subject. We wanted this «39BIS File» to serve as a white paper on healthcare chatbots, a first sharing of experience on the subject that could be useful to everyone. There are indeed many aspects and requirements that are specific to the development of a healthcare chatbot, and they are described in this white paper. In particular, a healthcare chatbot must have a high rate of relevant answers and provide accurate, scientifically verified and referenced information, because when it comes to health, anything less than entirely reliable is not acceptable.

There is enormous potential for chatbots to improve the quality of healthcare services. Far from worrisome fantasies, chatbots could, on the contrary, improve the human dimension of care, by taking on simple tasks at each stage of the patient’s care pathway and thus freeing up time for healthcare personnel. At the same time, we are aware of the ethical aspects raised by these new e-health models, and intend to integrate ethical principles during the design phase.

We hope you enjoy reading it!
Orange Healthcare, a subsidiary of Orange Business Services, helps healthcare stakeholders modernize with digital solutions for a more efficient healthcare system.

Healthcare information systems are becoming increasingly complex due to the wide range of digital tools available to patients and healthcare professionals. In this context, we believe that a chatbot can be a relevant solution that facilitates the use of digital tools. Like intelligent search engines, a chatbot can, for example, enable someone to find relevant information quickly. In a company like Orange, there are many areas in which to experiment and chatbots are developed for both external and internal use.

We accepted Sanofi’s request, realizing that the development of a healthcare chatbot involves specific requirements, in particular to meet safety and privacy standards for health data, but also safety and efficiency standards concerning the use of the device (patient follow-up, diagnosis, medical appointment). Finally, in healthcare, it is essential to evaluate the risks and limitations of chatbots.

The very useful experimental work we are doing with Sanofi and Kap Code is clarifying the role that information technologies can play in the healthcare context.
New digital technologies are driving a gradual transformation of the health sector, aimed at meeting the major challenges of the healthcare system, namely ageing of the population, development of chronic diseases, medical desertification and the demand for personalized medicine. It could facilitate the coordination of care, administrative management and in some cases diagnosis and therapy.

According to the WHO, e-health is defined as «digital services for the well-being of the individual» et «the use of tools for the production, transmission, management and sharing of digital information for the benefit of both medical and medico-social practices».

Today, the health sector is using more and more connected tools: wearable devices (bracelets, watches or any connected clothing), telemedicine, shared medical data, virtual or augmented reality, etc.

Thanks to these new tools, patients can play a more active role in their care pathways. Since September 2018, in France, health professionals have been able to perform telemedicine diagnosis and healthcare that reduce unnecessary travel and that are reimbursed by social security. Remote expertise engenders coordinated management by professionals.

These new approaches to health care management and the care pathway require thorough evaluation of the digital tools they use. In this white paper, we wanted to focus on chatbots.
INTRODUCTION TO CHATBOTS
CHATBOT AND CONVERSATIONAL AGENT?

Several million people enter keywords every day in search engines such as Google and then have to choose from a list of results, usually in the form of web pages in which it is again necessary to search for specific information.

A chatbot is a software robot that can reproduce natural language and interact with an individual through automated conversations. Chatbots allow you to receive a unique answer or a service.

In the literature, chatbots and conversational agents can be distinguished according to their level of understanding of natural language, the former using keyword or rule engines instead, while the latter are based on machine learning. We shall use the term chatbot in its generic sense in this white paper.

The operating model of a chatbot is always the same, whatever its scope, its theme and its level:
- Users formulate their queries in natural language via a voice or text interface.
- The chatbot receives the request and its engine interprets it to understand it.
- The chatbot provides a unique and qualified answer to the user’s query. The answer may be generic (i.e. the same for everyone), contextualized (adapted to the context, for example, at a given time and place) or customized (adapted to users, for example, by providing them with their bank balance).

Chatbot technologies are available via several digital channels: text (SMS, web pages, messaging applications, mobile applications, etc.) or voice (connected speakers, embedded services on mobiles, etc.).
THE CONVERSATIONAL AGENTS ARE ACCESSIBLE...

ON DIFFERENT APPS

Facebook messenger
Web chat
Slack

ON DIFFERENT DEVICES

Have a dialogue on a smartphone
Chat on a computer
Have a discussion on a tablet
Have a dialogue with a robot

EXEMPLE:
I would like to make an appointment with a dermatologist near me.

You can make an appointment directly with a dermatologist in Leicester by clicking on the following link to consult the available appointment times.
The answer is adapted and users are supported, here in their attempt to make an appointment.

CONCIERGE
To consult the list of dermatologists around you, visit nhs.uk
The answer is the same for all users.

ASSISTANT
Doctor William Harry is available tomorrow at 6:15 pm, at 12 Watling Street in Leicester. Would you like me to make the appointment for you?
The system requests an appointment on behalf of the user. The user delegates his/her actions.

ADVISOR

THERE ARE 3 TYPES OF CHATBOT:

**ASSISTANTS**
which provide the user with a predefined answer like in a page for “Frequently Asked Questions”.

**CONCIERGES**
(akin to a hotel receptionist), which provide a contextualized response and facilitate a service to the user, for example by explaining the steps of an action to be taken.

**ADVISORS**
which integrate customized answers to complex requests with automated processes to perform certain actions.

**EXEMPLE :**

I would like to make an appointment with a dermatologist near me.

**ASSISTANT**
To consult the list of dermatologists around you, visit nhs.uk

The answer is the same for all users.

**CONCIERGE**
You can make an appointment directly with a dermatologist in Leicester by clicking on the following link to consult the available appointment times.

The answer is adapted and users are supported, here in their attempt to make an appointment.

**ADVISOR**
Doctor William Harry is available tomorrow at 6:15 pm, at 12 Watling Street in Leicester. Would you like me to make the appointment for you?

The system requests an appointment on behalf of the user. The user delegates his/her actions.
Chatbots use three types of data to do their job:

- **Utterances (or variations),** which refer to sentences or questions asked by the user and pre-programmed in the chatbot. For example: "What will the weather be like tomorrow in Blacklunans?"
- **Intentions (or intents),** which refer to the theme of the variation (i.e. weather forecasts in the previous example).
- **Entities,** which are contextual elements contained in the variation. The most common entities are places, dates, proper names (persons, companies, etc.). Some variations may not contain entities. In the previous example, we have a time entity (tomorrow) and a place entity (Blacklunans).

Natural Language Processing (NLP) refers to the techniques used by machines to perform various language operations. In the field of chatbots, the two NLP techniques we are interested in are NLU (Natural Language Understanding) and NLG (Natural Language Generation):

- The role of NLU is to determine, for a variation (i.e. a question asked by the user), the intention and possible entities associated with it. Thus, NLU mimics the human faculty of understanding the meaning and context of a sentence by reading it.
- When talking to someone, it is not enough to understand the person you are discussing with, you must also move the conversation forward. The role of NLG is to generate an answer or a new question.
Significant technical expertise is required to implement Natural Language Understanding (NLU). In some cases, especially when it comes to distinguishing subtle nuances, the use of buttons may be an alternative for guiding the user towards a precise choice and improving the understanding of the chatbot. The chatbot then proposes several alternatives from which the user can choose by clicking on one.

**IN NATURAL LANGUAGE:**

What type of payment method do you want to use?

Enter your answer here

**WITH BUTTONS:**

What type of payment method do you want to use?

- Credit/debit card
- Paypal
- Bank transfer of funds

It is also possible to use a mixed approach (natural language/buttons)
THE HISTORY OF CHATBOTS

Chatbots are in the spotlight today, but the first chatbot emerged in 1964 with ELIZA.

Several chatbots have been tested to try to understand and reproduce the human ability to conduct a conversation, through research on artificial intelligence in computer science. Other noteworthy chatbots were then created with Jabberwacky in 1982 and A.L.I.C.E. in 1995 for example.

Since 2010, the web giants have been launching smart assistants for smartphones and PCs to improve the user experience. The best known is Siri, launched by Apple on the iPhone in 2010. Then there was Google Now in 2012, Cortana at Microsoft and Alexa at Amazon in 2014.

Since 2016, chatbot solutions have been multiplying, particularly on Facebook Messenger, thanks to the simplification of chatbot technologies and implementation tools that anyone can use.
The ELIZA chatbot developed by Joseph Weizenbaum

Eliza takes the form of a psychotherapist who answers the user with questions. It works by detecting keywords in the user’s message and asks a question about them from a pre-recorded list of syntax formulas. The interaction between the chatbot and the user is therefore credible in a specific context, that of a dialogue with a psychotherapist.

The Jabberwacky chatbot developed by Rollo Carpenter

Jabberwacky was designed to simulate a conversation with a human being with the objective of also passing the Turing test. For Jabberwacky, there are no fixed rules or principles programmed; the system was designed to learn language and context through interaction with humans. It stores all conversations and comments from users and tries to use this information to find the most appropriate answer. The system is multilingual and able to learn and respond in the appropriate language if it has sufficient information.

The A.L.I.C.E chatbot developed by Richard Wallace

A.L.I.C.E is a more advanced version of ELIZA. It used for the first time the AIML language (Artificial Intelligence Markup Language) created specifically for natural language interactions.

Siri developed by Apple

Siri, launched in 2010, is the first mainstream consumer voice assistant integrated into a mobile operating system. With Siri, users can interact by text or voice to control a number of tasks on their smartphone in connection with certain mobile applications.
Chatbot technology has now reached a sufficient stage of maturity to be used for many different types of customer relations, which is the most frequent use of chatbots to date. They appear to be both a complement to call centers, and a way of saving money.

Launched in November 2017, Orange Bank is a 100% mobile bank that wanted to integrate from the outset the Djingo virtual advisor as the main channel for customer communication with the bank. Its main advantage is that it is available 24 hours a day and every day of the week. It enables the bank to respond to a strong demand from customers to get most of the information they need immediately, wherever and whenever they need it.

Since the bank’s launch, Djingo has handled more than one million conversations with Orange Bank customers (about 100,000 conversations per month), with a success rate for understanding around 75% . In more than half of the cases, the client obtains from Djingo all the information he/she needs, without another person intervening. In the other cases, the client is referred to an advisor at the Customer Relations Center. This automatically visualizes the history of the exchange and takes over by providing the customer with a more personalized relationship and a service with greater added value.

From this first use of Artificial Intelligence (AI) at Orange Bank, Djingo will evolve in the coming years into a virtual financial coach, anticipating customer needs by offering services in a proactive way.

This revolution in intelligent data processing raises key questions about controlling access to data and adapting the regulatory framework for the banking sector.

A second area of use is growing rapidly, that of chatbots for internal use within a company. This category has great potential for improving internal efficiency, but also for improving employees' working conditions, for example by simplifying their daily lives, two indirectly related factors.

During 2017/2018, various Sanofi teams conducted several experiments to provide new chatbot services to employees at Sanofi.

For example, a chatbot prototype has been developed to handle French employee leave processes for Human Resources departments. In its initial version, the chatbot is able to answer the most frequent and time-consuming questions for Human Resources support in France, such as postponing, deleting or modifying leave.

A survey on digital use shows a global increase in instant messaging (message volume, variety of services), with WhatsApp taking first place in Europe, WeChat in China and iMessage in the US.

With Machine Learning, in 2017, Google achieved word comprehension in 95% of cases (Google Machine Learning Word Accuracy).
SEVERAL FACTORS CONTRIBUTE TO THE INCREASE IN THE NUMBER OF CHATBOT ECOSYSTEMS

On the product side, in the United States, the Amazon Echo installed base exceeded 30 million products at the end of 2017\(^1\), the base of connected loudspeakers installed reached 50 million at the end of August 2018\(^2\), and was expected to reach 100 million worldwide\(^3\) by the end of 2018.

Surveys on digital use show a global increase in instant messaging (message volume, variety of services), with WhatsApp taking first place in Europe, WeChat in China and iMessage in the US.

With Machine Learning, in 2017, Google achieved word comprehension in 95% of cases (Google Machine Learning Word Accuracy)\(^4\).

Payment is becoming easier and easier, “frictionless”: when consumers from more than 16 countries were asked, they responded that more than ~60% of payments are made digitally. The payment channels of which the chatbots are a part, vary from country to country\(^5\).

---

1 - Internet Trends Report, KPCB May 2018
2 - CIRP Report, August 2018
3 - Media alert: Smart speaker installed base to hit 100 million by end of 2018, Canalys, 2018
4 - Google Machine Learning Word Accuracy
5 - Data from various sources
The web interface allows users to interact with the chatbot while it displays useful information. This interface is capable of:

- identifying employees,
- connecting to the HR system,
- reading the various days of leave taken by this employee,
- adapting its answer according to the employee’s position in a particular entity of the company.

This work made it possible to validate the usefulness of a chatbot for internal purposes.

As every chatbot is an evolving system based on the principle of learning from experience, this experiment demonstrated the importance of setting up a continuous improvement process. An agile methodology is therefore essential to ensure successful implementation. To promote use, the chatbot must be specific to a domain/use case and the range of intentions understood by the chatbot must ideally be broad in order to cover most of the users’ needs.

A third area of use is conversational shopping, with voice chatbots that are included in personal assistants such as Amazon’s Alexa, Google Home, or Orange’s Djingo. These assistants effectively create transactional chatbot marketplaces, allowing users to search, select and then purchase physical or digital goods. Similarly, for chatbots using native messaging applications such as Android or iMessage, once the payment has been controlled by dedicated modules (e.g. Google Pay or Apple Pay), users can choose what they want and buy it in a few interactions.
In your opinion, what is the maturity level of chatbot technology today?

Regarding maturity, it is important to bear in mind that chatbots are all about the use of natural language. In my team, we feel strongly about the difference between a chatbot that can replace FAQs in an “intelligent” manner, and a real conversational agent able to participate in multiple interactions to understand a question correctly, determine the context and identify an appropriate answer.

On the basis of a “cognitive journey” – the pathway towards increasing complexity – organised around at least three maturity levels, from “assistant” to “concierge” to the current “advisor” level, most deployed chatbots meet only the implementation requirements of an assistant capable of answering a simple question, and maybe personalising the answer to suit the context. The chatbot era is still in its early stages, in its infancy even, including for the largest players such as Amazon, IBM or Microsoft. Besides the aspects of proper natural language use and a chatbot’s ability to understand complex sentences, the use of these assistants is struggling to establish a strong Return on Investment (ROI).

In my opinion, we are reaching a stage when companies and users are ready for the second phase in which smarter “concierge”-like systems are implemented. I think it will take a few years before human-machine conversations are technically and economically viable.

What makes the development of mature conversational agents so complex?

Today, many players base chatbot learning on the identification of keywords in user questions. This approach is also reinforced by the fact that we are conditioned to perform keyword searches on the Internet. This methodology ultimately simplifies chatbot learning, making it possible to develop chatbots quickly, but their understanding of complex sentences remains limited.

To resolve the problem of generating real, structured human-machine interactions in which appropriate answers are formulated, several issues need to be addressed: what type of conversation do we want to establish? How should such conversations proceed? What information within the conversation is relevant to the original request? For example, if a person calls the Orange customer service to report the loss of their phone, they may enter into a level of detail that is incomprehensible to today’s chatbots, e.g. about the circumstances of their loss, or associated events that have nothing to do with their request for a replacement mobile phone. Most current chatbots are out of their depth in this type of situation.

Priming a mature conversational agent requires the use of tens of thousands of sentences for learning, which have to be designed with the close involvement of industry and linguistic experts.
How can the learning of such first-generation chatbots be facilitated?

This question can be addressed on two levels: how can chatbot capabilities be enhanced and how can learning be made simpler?

One solution to enhance the capabilities of first-generation chatbots is to consider opening them to external databases such as Wikipedia, YouTube or Google. The chatbot will then search for answers in these existing databases if the answer is not found within the scope acquired during learning.

One aspect not to be overlooked in chatbot implementation is the availability of reusable “chit-chat” modules, such as “what is your name?”, “what can you do?”, “how do you feel?”, “tell me a joke”, which improve user experience.

One of the ways to simplify learning is to use NLP (Natural Language Processing) tools to analyse conversations to identify any variation with the questions used in the learning phase.

How do we move from a prototype to an industrial solution for chatbot projects?

POCs (Proof of Concept) are relatively easy to launch for chatbots, and concern limited scopes for which the rate of fast understanding of intentions is high. But, if we want to achieve greater performance, for example by broadening the scope or by interfacing with the IS (Information System), things get more complicated. A much stronger involvement of industry experts is required.

The chatbot life cycle should also be taken into account as chatbots improve over time, and they should be built to allow for this continuous improvement.

In terms of technology, NLP engines that work for POCs will also work on industrialised versions, but scaling-up issues will have to be taken into account when designing the infrastructure.

A chatbot project is an ideal candidate for an agile approach: a simple but viable solution for the user should be rapidly implemented with a sufficient number of relevant answers, and this should be continuously improved and enhanced with the lessons learned from past interactions.

You mentioned the concept of ROI for chatbots. Can you tell us more about the economic viability of a chatbot?

Today, significant financial investment is required to deploy a suitable level chatbot, such as the Orange Bank chatbot based on IBM Watson and cited as an example of a disruptive AI solution by IBM CEO, Virginia Rometty. This kind of solution should be part of a long-term development strategy with a clear commitment to move towards a cognitive approach to customer service. It is also essential to define the ROI in relation to existing solutions.

In my opinion, one of the keys to a chatbot’s economic success is to be found in the field in which it is used. For there to be a fast ROI, very few digital solutions should have been deployed in the field already. For example, you would be hard pushed to make the business case for using a chatbot in telephone-based customer service if a natural language IVS (Interactive Voice Server) is already being used.

By contrast, the healthcare industry is a very attractive exploratory field for teams like ours.
Although healthcare was not the first sector in which experiments with chatbots have been carried out, since the beginning of 2018 we have seen the emergence of and experimentation with many different use cases in this field.

The chatbots thus try to handle several needs, such as personalized medical follow-up, communication and transmission of test results, dissemination of information, or even advice to patients or preliminary diagnosis.

It is in this context and based on the project initiated by Sanofi, in partnership with Orange Healthcare and Kap Code, that we are exploring in this white paper some practical cases of healthcare chatbots and the specificities of the healthcare sector. The white paper also includes our proposals for evaluating user perception of these new digital tools.
Our research was not meant to be exhaustive because we only tested those chatbots that were freely accessible. The inclusion of these chatbots in our white paper is not therefore based on their level of quality. In our study, we identified chatbots that were in the project phase and some that are already being used for everyday purposes. Some of them have resulted in scientific publications describing the use of the technology in healthcare, for a given subject.

It is important to note that, despite the emergence of many healthcare chatbots, there are still issues to be resolved about the maturity of the technology and of the uses. Nevertheless, studies have been undertaken and chatbots are being created, thus paving the way for gradual improvement of the technology to meet many needs.

We have identified certain trends in potential uses throughout the patient care pathway, as illustrated below:

ILLUSTRATIONS OF HEALTHCARE CHATBOTS

We conducted a web-based study to identify the different healthcare uses of chatbots around the world, focusing on French and English publications and tools.
POTENTIAL USES THROUGHOUT THE PATIENT CARE PATHWAY

- Medical information
- Prevention
- Pre-diagnosis and referral
- Hospital and outpatient care
- Support for chronic diseases
- Compliance and information on treatments
Today, health is one of the leading topics researched by Internet users on search engines, as shown by the increasing use of health forums and specialized sites for the general public in search of answers on a disease and on the different types of care. Although these forums and sites allow for a lot of sharing of experience between patients, the information is generally not subject to verification by health professionals, and is often free of any context, resulting in a risk of poor information or “misinformation” of the patient.

Access to quality medical information for all is therefore a major challenge in the process of empowering individuals to take responsibility for their health. New technologies, and in particular chatbots, drawing on verifiable information, would be perfectly adapted to meeting these challenging issues. Patients who avoid forums for the reasons mentioned above may have to consult documentation that is sometimes difficult to understand for non-health professionals. In addition to providing reliable information, the chatbot can use language adapted to the user’s level of understanding, depending on whether the person is a patient or a healthcare professional.

In oncology, WeFight has developed a chatbot called «Vik»⁴, which allows breast cancer patients to benefit from follow-up and advice, particularly on their lifestyle habits, as well as information about the disease and treatments. The chatbot also offers different access for patients and their families with a personalized service (possibility of subscriptions, drug reminders, etc.).

We can also note the example of the «Lybrate»⁵ chatbot developed in India. Published on Facebook Messenger, Lybrate answers simple health-related questions. It offers patients quizzes to perfect their knowledge. It can then refer them to an online consultation with a physician on the Lybrate website.

This technology could also be of interest to healthcare professionals who often need accurate, verified and referenced information to help them manage their patients. They sometimes lack the time to conduct complete searches through the mass of heterogeneous information offered by a search engine. Some databases are often difficult to handle, as their interfaces are not very user-friendly, and employ a technical lexicon that can vary from one medical specialization to another. A suitable chatbot would enable the health professional to be advised simply and quickly, safely and concisely.

---

⁴ - https://www.wefight.co/fr/accueil/
⁵ - https://www.facebook.com/lybrate/
PREVENTION

In the current health system, prevention is a major challenge for all ecosystem stakeholders both to improve the quality of life and to reduce the costs of treating diseases.

Many communication tools and campaigns have been deployed to convey prevention messages. Conversational agents can offer many advantages in the field of prevention, including virtual support combined with advice. For example, by working on daily eating habits, chatbots can help people adopt healthy lifestyles, as evidenced by a study on obesity prevention in Italy.

In January 2019, the Ramsay General Health Foundation launched a chatbot on its Facebook page. Dedicated to prevention, it allows for a dialogue with three virtual experts depending on the theme chosen: a stress expert, a nutrition expert and a tobacco addiction expert. The Ramsay General Health Foundation presents this chatbot as a tool that responds in a personalized way to users’ expectations by offering tailor-made programs, weekly check-ups, or an «alert» to accompany the user in the difficult moments of quitting smoking.

**Smart Alfred**, developed by the Betterise company and available in a Beta version on Facebook Messenger, offers users a personalized follow-up on different themes such as morale, alcohol, stress... Depending on the user’s answers, it pushes personalized content (recipes, physical activity tips, etc.) and allows users to compare their answers with their respective histories (number of glasses of alcohol drunk, morale during the last few days, etc.).

PRE-DIAGNOSIS AND REFERRAL

Digital technology and the resulting immediacy are omnipresent in today’s world, including in healthcare.

It sometimes takes a long time to obtain medical appointments, especially in certain regions where there are hardly any physicians. When patients have new symptoms, they want to quickly understand why. Web searches often yield multiple and sometimes unreliable results. Patients may feel overwhelmed and not know where to go, for example by going to the emergency room for a minor illness. To overcome this, several chatbots have been developed to provide answers and steer the patient in the right direction.

For example, the MedWhat chatbot, developed in the United States by healthcare professionals, is used to diagnose relatively minor illnesses (rhinitis,  } 6 - Addressing Challenges in Promoting Healthy Lifestyles: The AI-chatbot Approach, Ahmed Fadhil, Silvia Gabrielli
7 - https://medwhat.com/
sore throat, nausea, etc.) and to refer patients to a healthcare professional for an appointment if necessary. Once a diagnosis has been proposed, it has to be validated by a physician. This technology thus makes it possible to avoid long waits, and to accelerate the referral of patients to appropriate healthcare structures.

In China, Melody, a chatbot inside the Baidu Doctor app, saves the physician time in diagnosis by analyzing the patient’s symptoms before the appointment and offering the physician a preliminary diagnosis that he or she must then validate. The physician can access medical information while the patient can make an appointment directly via the chatbot. This example is a good illustration of the role this technology can play in setting up medical appointments: interviews prepared in advance enable the patient’s needs to be determined better so that they get care faster, which allows more time for personal exchanges between the patient and the health professional.

In France, a chatbot that offers a diagnosis or therapeutic orientation is considered to be a medical device. To date, we have not identified any chatbot that is recognized as such.

Far from wanting to replace the medical profession by proposing a diagnosis (the latter, once proposed, must always be validated by a physician), the chatbot is positioned as a facilitator of the patient-physician relationship: it is a communication vector.

In hospitals, the automation of certain administrative tasks via a chatbot can free the nursing staff to attend to their patients.

When preparing a consultation, a teleconsultation or an outpatient hospitalization; whatever the nature of the appointment, the physician or the reception structure will need to collect information about the patient. This information can be administrative or medical: medical history, current treatments, and the reason for the consultation… or any other information necessary for the smooth running of the appointment. Consequently, the implementation of a tool such as the chatbot can enable this data to be made
available ahead of time or even a concise analysis that can be quickly read and understood.

Some chatbots have been developed to facilitate the management of outpatients and the making of appointments with specialists\textsuperscript{10}.

For example, the Calmedica solution allows post-ambulatory follow-up for patients by SMS through a simple chatbot that replaces the next day’s call. This allows for patient follow-up, while freeing up healthcare personnel and avoiding patients being disturbed too much after their return home.

BotDesign’s chatbot is currently being tested at the American Hospital in Paris in the obstetrical anesthesia department to allow patients to monitor their health during their last weeks of pregnancy from their homes and prepare for hospital admission. The chatbot sends questionnaires to patients to collect data on their state of health, it provides the healthcare personnel with a structured synthesis and enables the management of appointment planning, patient referral and the transmission of medical records between health experts. Another test is underway at Toulouse University Hospital, in geriatric oncology\textsuperscript{11}.

However, there are still only a few examples of chatbots in hospitals and outpatient wards. In France, this is mainly due to the number and complexity of IT tools already used in hospitals, but above all, to their diversity.

\textbf{SUPPORT FOR CHRONIC DISEASES}

The personalization of the discussion and grasping of the context make the chatbot a good tool for therapeutic follow-up.

The Mara University of Technology in Kuala Lumpur, Malaysia, for example, has developed a chatbot called ViDi\textsuperscript{12} (Virtual Dietitian) that presents itself as a virtual dietitian to support diabetic patients. Regular conversations with ViDi, during which the chatbot will ask the patient a number of questions, enable it to provide the patient with appropriate dietary recommendations. In particular, this can enable more regular following of the recommendations and ultimately an improvement in the patient’s state of health.

\textsuperscript{10} - https://www.avaamo.com/ai-healthcare
\textsuperscript{12} - An architectural design of virtual dietitian (ViDi) for diabetic patients, Mara University of Technology, Malaysia

29
Our research has shown that there is a strong trend toward using chatbots in the field of mental health, and that there are many scientific publications on the subject.

The TeenChat\textsuperscript{13} chatbot test developed by Xi’an Jiaotong University in China has shown that teenage patients suffering from stress use this type of tool more regularly and are more honest about their feelings when chatting with a chatbot than when consulting a psychiatrist or psychotherapist. The corresponding scientific article highlights their familiarity with new technologies, but also the daily follow-up that reassures and supports patients about their condition.

In Sweden, a chatbot called Shim\textsuperscript{14} provides support for the mental well-being of vulnerable people. The corresponding scientific publication concludes that it is worthwhile to use chatbots in Mental Health. As in the previous example, personalized conversation is greatly appreciated by patients.

The Australian E-Health Research Centre, in collaboration with the Commonwealth Scientific and Industrial Research Organisation (CSIRO), developed a chatbot in 2018 to support autistic children in their development, by advocating non-verbal communication\textsuperscript{15}. With the help of the Asia-Pacific Center for Neuromodulation at the University of Queensland, they had already undertaken, in 2015, the development of a chatbot to support patients with Parkinson’s disease\textsuperscript{16}.

Therapeutic compliance is defined as the relationship between the prescriber’s recommendation and the patient’s actual intake of the prescribed drugs.

Today, compliance is a major public health issue. This is especially true for patients with chronic diseases who often have several drugs prescribed, some of which have to be taken throughout the person’s lifetime. Symptoms which are sometimes not apparent, therapeutic «fatigue», the number of daily doses and many other factors, explain why patients do not comply with prescription instructions. In fact, just reminding the patient to take his or her medication is insufficient to guarantee that they will follow the medical prescription instructions. A chatbot could improve this support by providing key information about the products and encouraging patients to follow their treatments carefully.

The start-up Smart Health UG has developed a chatbot called Florence\textsuperscript{17}, which allows patients to have a personalized follow-up for taking their treatments, with motivational messages as well as information on the drugs prescribed for the treatment.

\textsuperscript{13} - TeenChat, a Chatterbot System for Sensing and Releasing Adolescents’ Stress, 2015, Xi’an Jiaotong University
\textsuperscript{14} - A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods, Karolinska Institutet, Stockholm, Mittuniversitetet, Östersund, Linköping University, Linköping, 2017
\textsuperscript{15} - Designing a Chat-chatbot for Non-Verbal Children on the Autism Spectrum, The Australian E-Health Research Centre, CSIRO, Australia, 2018
\textsuperscript{16} - Chat-chatbots for People with Parkinson’s Disease: Science Fiction or Reality?, The Australian E-Health Research Centre, CSIRO, University of Queensland, 2015
\textsuperscript{17} - https://www.florence.chat
This company has also developed a free chatbot called «Izzyperiod» on Facebook Messenger, allowing women to obtain information on menstrual cycles, sexuality and contraception. It also offers personalized pain management and follow-up, as well as raising awareness of endometriosis. Its objective is to make it easier for patients during the menstruation period by providing support, particularly on contraceptive use\(^\text{18}\).

Chatbots can also provide patients with more information about their treatments or the products they use. This is the case, for example, with «Annabelle»\(^\text{19}\), the chatbot for Naturactive products, a brand of phytotherapy, aromatherapy and food supplements from the Pierre Fabre laboratory. This chatbot was developed by the Naturactive team and the BotDesign start-up and more than 400 questions and answers were integrated.

At Sanofi, for the Novanuit range, the «Nina»\(^\text{20}\) chatbot is the first chatbot on sleep disorders: at all times, it provides users with tips, advice and playlists to help them sleep better. It has the advantage of being available 7 days a week and 24 hours a day via Facebook Messenger.

In the case of restricted and/or complex prescription drugs, a chatbot would be a relevant tool for physicians to support them with their prescriptions, as shown in the perception study presented in the section «Evaluation of healthcare chatbots».

**CONCLUSION**

There are many examples of healthcare chatbots, and a wide range of applications. Many countries and health stakeholders have understood the importance of this technology in meeting the challenges of the sector in the years to come. Fears have been expressed that chatbots could replace the physician or interfere with the patient-physician relationship, but this has not occurred in the experiments conducted to date. As we have seen above, these tools can, on the contrary, help improve healthcare, such as by supporting diagnosis or patient referral, which would leave the physician more time to attend to the patient. The future of healthcare chatbots will depend on the perception of health professionals and patients and the extent to which they adopt them.

---

\(^{18}\) - http://www.superizzy.ai/
\(^{19}\) - https://www.mindhealth.fr/article/11633/pierre-fabre-lance-un-chatbot-pour-sa-marque-naturactive/
\(^{20}\) - https://www.facebook.com/teamdenuitparNovanuit/
In your opinion, what would be the role for chatbots in the e-health ecosystem?

We live in an increasingly connected, digitized world … in which we want fast, if not instantaneous, access to information, and a common question is whether this information is reliable. Health is not immune to this transformational change as doctors and patients need simple, fast, reliable and secure solutions. Healthcare chatbots are conversational assistants developed to provide a service to patients or healthcare professionals within a specific area. They need to be positioned as a trusted solution, which users can rely on as and when needed. If chatbot uses and associated technologies develop as we expect them to today, chatbots are set to become an essential e-health solution.

How would they fit in with general medical practice today?

Even though their technologies and uses have not yet fully matured, chatbots could quickly become helpful to healthcare professionals, for example by providing them with information on diseases or the proper use of medicines, in a manner appropriate for their needs. Since chatbots provide fast, reliable answers based on up-to-date scientific and/or regulatory data and would be accessible 24/7, they could be a meaningful tool to guide the proper use of healthcare products.
From a pharmaceutical company’s perspective, why are chatbots relevant?

As a drug manufacturer, we are in the business of developing, producing and supplying the treatments that individuals need to stay healthy, while paying particular attention to the safety of those treatments.

The use of chatbots could bring many benefits to pharmaceutical companies, particularly in terms of providing information about the drugs they sell, since chatbots are:

- Available all the time (and everywhere)
- Fast
- Time-saving (a unique answer to every question)
- An evolving system (based on data maintained in accordance with state-of-the-art principles and the latest scientific data available)

Chatbots could provide first-line information on drugs, helping ensure that they are used properly.

What are the areas to be attentive about and the key success factors for healthcare chatbots?

As a pharmaceutical company, we are responsible for the information we provide on our medicines.

In addition to the innovative aspect, developing such a tool could turn out to be very useful, but this requires ensuring that the answer provided is:

- Reliable (the right answer to the question)
- Referenced (supported by a scientific source or backed by the health authorities)
- Updated (kept up to date at all times)
- Relevant (in line with user expectations)
- Clear and easily understandable

The key success factors are thus as follows:

- Chatbots must be developed in collaboration with end users, be they patients or healthcare professionals, to ensure that the tool is relevant to them and endorsed by them.
- Chatbots must be reliable.
- Their usefulness must have been assessed and validated in user practice.
SPECIFICITIES OF HEALTHCARE CHATBOTS

In healthcare, more than in any other field, a chatbot must meet certain essential criteria in terms of form, content, safety or technology.

The notion of risk when the chatbot is used by health professionals or patients must be integrated into its design, and, depending on the level of risk, appropriate regulations must be applied. Quality, relevance, efficiency, safety and data processing are key criteria to be taken into account at each phase of a chatbot project.

The sector is governed by many standards and regulations, such as ISO standards for medical devices or FDA (the United States Food and Drug Administration) certifications, and each project team will have to consider whether or not it is necessary to meet these different criteria.

THE DATA

As seen above, the chatbot is part of an ecosystem composed of multiple e-health solutions.

If it is to bring value to its users, patients or healthcare professionals, the chatbot must be a reliable solution. For example, it will be necessary to indicate at the beginning of the conversation the sources of the data used by the chatbot or how the information has been scientifically validated by health professionals, expert groups or authorities. Studies or documents can also be cited throughout the conversation without adding to the dialogue, for example with links that go directly to the sources. The chatbot dialogue will have to be validated by the users. For a pharmaceutical company, medical, regulatory and pharmacovigilance validation of answers is essential. Beyond this point, the whole project will have to comply with the GxP process (Good Practices - the «x» refers to the particular domain) corresponding to the chatbot. This process is determined by the FDA to ensure that the tool is safe, usable and meets strict quality standards throughout the production process.

Apart from drawing on validated sources, in the health field the chatbot needs to be updated almost in «real time». For example, if the chatbot provides information about a drug and the drug’s indications or dosage change, this will have an impact on answers. The update should be done as soon as possible in
order to avoid an incorrect response that could have consequences. To this end, besides the implementation of the chatbot solution, related use cases and dialogue, it is important to anticipate the impact on processes and businesses. As with the answers to questions sent to laboratories through traditional channels, a well-organized update process will help taking into account as quickly as possible any change in product information related to the chatbot theme that could influence the content of the answers provided.

The question of data security must be addressed when setting up a chatbot.

As it is possible that users may communicate personal health data in their questions, whether or not it is the subject of the chatbot, the management and storage of the data must be studied. There are several possible approaches, among which:

- an information message at the beginning of the conversation that asks the user not to provide personal health information
- keeping all data received by the chatbot anonymous, using a tool that filters, for example people’s names, birth dates...
- storing of the conversation history with a certified host that complies with the regulations in force in the country where the chatbot is used.

Data security is also a key element for building user confidence in the chatbot. In a study on e-health conducted by Pfizer in November 2018\(^1\), only 38% of patient associations believe that data is sufficiently secure in e-health applications and 40% of them believe that it is no more secure than on social networks. These perceptions can put people off using them even though there are regulations to protect users. The chatbot must therefore succinctly express how it secures the data and provide users with the legal information necessary to reassure them.

\(^1\) Etude Pfizer / Cercle P - L’e-santé vue par les patients : risque ou opportunité ? – Novembre 2018
THE PERSONALITY

Personality has now been established as an essential principle when designing chatbots to interact with users in everyday life.

Most often, computer modelling of personality consists in reproducing characteristics of a human personality. However, personality is a complex concept whose most commonly used model in artificial intelligence is that of «traits», considered as a set of fixed and stable psychological characteristics that differentiate individuals. Consequently, the design work is carried out by varying different parameters whose combination is supposed to «embody» personalities (for example, interaction style, linguistic expression, speech characteristics when using a vocal chatbot, etc.). Their designers sometimes give them first names, a sense of humor, cultural references, etc.

While noting that this idea of personality concerns all chatbots whatever they may be used for, we are entitled to consider that the medical field is a separate field, requiring a thorough study of the personality of the conversational agent during the design phase. Indeed, the relationship between the patient and the physician is the most important aspect of medical practice. The ability to listen, understand and be compassionate towards the patient is fundamental in building a relationship of trust. A patient chatbot will therefore need to show empathy to develop patient trust and potentially bolster this key patient-physician relationship.

Thus, in healthcare, it is important to remain very vigilant about the defined personality, as it will make the conversation and the information provided both credible and audible to the user. Depending on the target users, patients or health professionals, the chatbot must adapt the tone and vocabulary used in the dialogue, especially for understanding sometimes complex medical terms. For example, a reassuring and courteous tone, without condescension or judgement, seems to be preferable when the use is intended for patients. As for the physicians, they recommend a sober and concise tone.

However, very few studies have focused on understanding how users perceive personality possibilities, which depends on the proposed use in the medical field. Consequently, we have also chosen to mention studies on the use of robots that have a vocal feature. In this context, a study conducted by Tay, Jung and Park22, consisted in testing different personalities with users for a «nursing robot»: an «extraverted» personality characterized by a robot that speaks loudly and quickly and initiates conversations and, an «introverted» personality characterized by a robot that is more reserved, speaks slower, more quietly and waits for the user to initiate interaction. In addition to these traits, each personality was tested with a «gender» based on the voice (female or male). According to the authors, the study showed the effects of stereotypes related to social role and gender: participants judged the «nursing robot» more positively (for example in terms of trust or acceptance) when it was «embodied» by a female voice and

---

when it appears to have an extraverted personality’. A second study by Tapus and Mataric23 using robots designed to help people undergoing physical rehabilitation after a stroke, was conducted as part of exercises such as «turning the pages of a newspaper» or «placing books on a shelf». They observed that robots that produce compassionate messages such as «I hope it’s not too difficult» or «I’m here for you» are more suitable for introverted people, while robots that produce directive messages such as «you can do more than that» are more effective and appreciated by extroverted people. It should be noted, however, that this last study was conducted with «healthy» participants who were asked to use their non-dominant upper limbs (for example, left hand for a right-handed person) to simulate the damage usually caused by a stroke.

In conclusion, the personality of healthcare chatbots is a real issue that still needs to be explored given the limited literature available on chatbots used in real-life conditions. In addition, the personality of the chatbot should be determined according to the user and the use case (examples: support, medical information, therapeutic follow-up, etc.).

THE TECHNOLOGIES

Today, chatbots, and in particular the connected speakers developed by the GAFAs (Google, Apple, Facebook, Amazon and Microsoft...), such as Amazon’s Alexa, are marketed without any specific use in mind. They can answer questions about knowledge stored in their data base, place orders, play music for you, manage appointments or even make a compliment...

The GAFA model aims to optimize its solution through use, by learning questions that consumers will ask. Chatbots thus increase their ability to respond and enrich use cases as needed.

This is not the case for a healthcare chatbot. Before a chatbot is made available, the topics to which it must respond have to be validated, for example through workshops with future
It is unthinkable that a healthcare chatbot would be deployed for its users without being capable of delivering a very high rate of relevant answers. The best way to train the chatbot is through the learning phase, of fundamental importance, and the dataset must be, first and foremost, of high quality and also comprehensive. The critical point of a healthcare chatbot will be its ability to respond with relevant information. If a user asks a question A, then answer A must be given regardless of how question A is expressed. The risk of answer B being given to question A must be minimized.

NLU engine will be indispensable for achieving an acceptable rate of relevant answers for healthcare chatbots. In a written exchange between two people, there may be spelling and grammatical errors and a wide range of synonyms for expressing a given word. A chatbot needs training (or learning) to understand the user’s natural language. Based on artificial intelligence, NLU is the first step in natural language processing. It is necessary to train a chatbot from a training dataset, then test it with a new dataset to evaluate the performance of its Natural Language Understanding algorithm and then improve it.

While there are many generalist NLU engines, the major health challenge is to have access to a medical corpus that is rich enough to improve them. This is vital because once the chatbot is able to understand the question, it will be able to generate the appropriate answer as configured.

Today we can rely on certified and very extensive medical databases that are available on the Web:

- **One of the best-known databases is Medline**, which offers abstracts of scientific articles.

- **The National Library of Medicine (NLM) offers the Unified Medical Language System (UMLS)**, which is certainly the most extensive terminology resource currently available in the healthcare field.

- **SNOMED CT® (Clinical Terms)** is an international system of clinical terminology. It includes a list of verified, detailed and medically validated clinical terms with their synonyms.

- **The CISMEF (Catalogue et Index des Sites Médicaux de langue Française)** is a French database dedicated to teaching and research.

- **Bioportal** lists more than 770 medical ontologies for 9 million concepts. The lexical field is wide and highly technical. For example, in the context of diagnosis support, the ontology of clinical phenotypes includes more than 16,000 concepts for genetic diseases and 350,000 concepts for the structured nomenclature of clinical terms.

However, analytical work is required to use them. After downloading the files and cleaning the data, it will be necessary to identify the desired terms and then label them with concepts. This work will produce a corpus on the defined subject. This is important for integrating the range of synonyms. In addition, the chatbot will have to process the:

- **Homonymy.** Example «I performed a medical procedure» versus «he came to the medical center».

- **Polysemy.** Example «I put in a central venous line» versus «his state of health
is in line with what it should be». However, given the richness of medical semantics and the technical complexity of the vocabulary used, artificial intelligence cannot take over completely at the current time. Human expertise can overcome some comprehension difficulties by improving the design of the conversation, for example with buttons. The user can also be asked to clarify, in order to complement the NLU, as can be seen in the example below.
Furthermore, if the chatbot is developed by a pharmaceutical company, particular attention must be paid to pharmacovigilance aspects. Indeed, a question/request expressed by a health professional or a patient may contain a pharmacovigilance case. The chatbot solution must therefore be able to detect any cases of pharmacovigilance, inform users and give them the option of reporting this case to the health authorities in accordance with the regulations. There are many types of pharmacovigilance cases. The adverse reaction is the best known but there are many others: misuse (or inappropriate use), lack or loss of efficacy, overdose, use without marketing authorization (off-label use), exposure during pregnancy or breast-feeding, etc. If the chatbot is not itself able to identify and analyze these different situations, it will have to be interfaced with another technical solution that can detect pharmacovigilance situations and send a signal to the chatbot so that it can propose a link to the user to report and qualify the case.

In fact, developing a healthcare chatbot involves additional technological complexity that as to be addressed.

EVALUATION OF HEALTHCARE CHATBOTS

The use of e-health solutions is booming in the health sector. These innovations raise the question of how the medical service provided to patients can be evaluated. What are the real impacts of these solutions on the healthcare system? What level of acceptability can be attributed to these real-life solutions for patients or healthcare professionals? How can the relevance of the information disseminated through these tools be validated?

These questions are now pushing health stakeholders to take a position on these issues: in France, the Haute Autorité de Santé (HAS) [the French Health Authority], for example, proposed in November 2016 a manual including 101 good practices to be complied with when developing connected objects and mobile health applications. Similarly, the Conseil National de l’Ordre des Médecins (CNOM) [National Medical Association] had already taken a position in 2014 on the need to have a reference framework for evaluating digital health tools. The private sector is no exception to the rule regarding this desire to support the evaluation of digital tools: for example, a start-up, Medappcare, is offering to certify mobile applications and websites in order to validate data protection, the quality of medical

24 - https://www.has-sante.fr/portail/jcms/c_2682685/fr/applis-sante-la-has-ebatit-101-regles-de-bonne-pratique
information and compliance with regulatory provisions. Even if these initiatives mainly concern mobile applications and connected objects, Artificial Intelligence and chatbots will need evaluating as these innovations are currently a major concern of health stakeholders.

However, to our knowledge, it has not been possible to follow up these initiatives with real-life scientific assessments and a common reference framework, which would enable a standardized and reproducible assessment, has not been established.

ASSESSMENT UPDATE

As seen above, there are many uses for chatbots in healthcare sector. They cover many areas, from prevention to diagnosis including prescription compliance and medical information.

These tools are an emerging phenomenon in the health field and their evaluation has yet to be explored in the scientific literature.

A systematic review of the literature on healthcare chatbots highlighted the lack of their scientific evaluation\(^\text{25}\). In this review, more than 1,500 scientific articles worldwide on chatbots for patients or healthcare professionals were analyzed. To refine this analysis, chatbots whose use consists in clicking on predefined answers were excluded, as well as those that do not answer the question asked (through pre-programmed messages) or that respond later by email. After excluding articles that do not meet these criteria, it appears that only 14 chatbots were evaluated in their complete final version with Artificial Intelligence using natural language. Several evaluation schemes were used, including randomized controlled trials but also cross-sectional studies and quasi-experimental models. Most of these 14 evaluations were carried out between 2014 and 2018, although the oldest ones date back to 2003. The chatbots studied, although innovative, often remain at the experimental stage and are hardly used in everyday situations. Finally, their effectiveness has rarely been evaluated.

Only one study evaluated the effectiveness of a chatbot\(^\text{26}\) on clinical benefits, highlighting a significant impact of chatbot use in reducing the symptoms of depression. The remaining evaluations may be divided into three categories: evaluation of the technical performance

---

25 - Conversational agents in healthcare: a systematic review, Laranjo et al., 2018
26 - Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial, Fitzpatrick KK, Darcy A, Vierhile M, 2017
(question recognition, response precision); evaluation of user experiences (satisfaction, usability, technical problems encountered by the user); and health indicators measurements.

The dearth of evaluations of healthcare chatbots and the heterogeneity of evaluation methods show that the field still needs to be explored, in particular by standardizing the characteristics of chatbots but also by developing evaluation methods to measure their effectiveness.

Following this observation, Sanofi and Kap Code conducted a study to evaluate physicians’ perceptions of a healthcare chatbot and identify initial criteria for evaluating this type of solution.

**METHODOLOGICAL APPROACH**

**ILLUSTRATION BY A STUDY OF ACCEPTABILITY AND HOW THE INNOVATION IS PERCEIVED**

Sanofi, in partnership with Orange, has developed the prototype of a conversational agent to support general practitioners for the proper use of restricted and/or complex prescription drugs through a first use case. Sanofi, taking an agile approach, wanted to validate the usefulness of this conversational agent and identify areas for improvement to enable new development iterations. Kap Code performed this evaluation on a sample of ten physicians. The prototype of this conversational agent was presented to the selected physicians in order to take their feedback into account.

There were three objectives for this evaluation:

- **Assess the degree** to which physicians accept such a tool.
- **Highlight the expected effects** of the use by physicians of digital health tools in general, and chatbots in particular.

In general, the aim was to evaluate the difference in relevance expected and perceived by physicians of a conversational agent in different situations in their practice, as well as the effectiveness of the chatbot, through feedback from users.

This evaluation was based on a qualitative study. The qualitative study was chosen to enable a detailed analysis of the effectiveness, acceptability and expected effects of the use of a conversational agent by a small sample of users. This type of study should be complemented by quantitative studies to improve the representativeness of the results.

To carry out the evaluation, interviews of approximately one hour were conducted with the physicians. Each interview was
conducted in three stages:

• **A quick introduction** on the context of the proper use of complex prescription drugs as well as the perception of digital health tools and chatbots by physicians.

• This was followed by a **presentation of the test version** of the conversational agent with a first demonstration by the investigator, then role-playing by the physician who was invited to use the agent as if he/she was conducting a normal medical consultation.

• **The third stage** involved discussion and feedback on the conversational agent.

Individual interviews were conducted to explore different predefined themes while allowing participants to introduce new themes through an open discussion.

Following the interviews, a horizontal analysis by theme yielded several results. It appeared that physicians, although familiar with the digital health tools they use (telemedicine, connected measuring equipment, consultation station, etc.), have not yet experimented with healthcare chatbots in their practices.

Several positive points were noted regarding the contribution of a chatbot to medical practice. First, physicians find the very concept of chatbot to be interesting. On the one hand, they can provide precise, scientifically validated information in a practical, user-friendly format. On the other hand, they save physicians time when searching for medical information by getting an immediate answer to their questions. This time saving is perceived positively by physicians as it would allow them to spend more time with patients during consultations and thus improve their healthcare. The chatbots were also appreciated for their format since most of the physicians felt that this type of tool could easily be integrated into their current practices, during consultations or outside of them.

While healthcare chatbots are generally perceived positively, some reservations were nevertheless expressed by the physicians interviewed during the study. The first concerned the effectiveness of chatbots. In order for the chatbots to add value, it is important for physicians that their content be exhaustive and cover broad health topics in order to centralize medical information. They expect such a tool to assist them in several areas, whether for indicating drug interactions that might involve a risk for patients, informing them about prescribing protocols and adverse drug reactions (as well as the possibility of reporting these adverse reactions) or more generally helping them make decisions by providing guidance on how to proceed, while leaving the final choice to the physician when faced with a particular situation. The second reservation expressed by the physicians concerns the possibilities of dialogue between the physician and the tool. They believe that the chatbot should be able to understand the physician’s natural
PHYSICIANS’ TESTIMONIES ON CHATBOTS

«It shouldn’t be fun, it should be interesting, it should be interactive, and it should be relevant. First, I want validated information, which I haven’t read, I want an answer that I haven’t read.»
woman, 8 years of practice

«To have an answer, to be more efficient in your research. Sometimes we have the impression that the answer is not collegial, on enzyme inductions or the pill for example. Now I have the answer in two seconds. That’s what makes the tool so practical.»
woman, 8 years of practice

«We have to think carefully about the content, it has to be professional content, it has to be precise, in a telegraphic style because we don’t have time to read 3 pages. Shorter answers would be better even if it means having to search more or having references to more precise stuff. Something like that would be more relevant. In my opinion, this would be the right thing to expand and develop for this concept. But the concept itself is exciting, I find it very interesting.»
man, 18 years of practice

«It shouldn’t be fun, it should be interesting, it should be interactive, and it should be relevant. First, I want validated information, which I haven’t read, I want an answer that I haven’t read.»
man, 30 years of practice
language while also formulating answers in natural language, in order to facilitate interactions.

Healthcare chatbots are thus positively perceived by interviewed physicians both with respect to the format and their potential contribution. As with all digital solutions or services, however, they should be adapted to physicians’ current practices by providing a fluid and comprehensive dialogue.

KEY FACTORS IN THE EVALUATION

In order to meet these expectations, several criteria can be taken into account when evaluating a healthcare chatbot.

In the context of the study mentioned above, the following criteria were taken into account.

Perceived relevance, contribution and usefulness

- **Confidence in the answers**, the completeness of the answers, the scope covered,
- **Relevance to the topic addressed by the chatbot**, accessibility to information, speed of access, clarity and tone used by the chatbot,
- **Relevance**: for decision-making, patient support,
- **The contribution in practice and criteria**, areas in which the tool can make a benefit,
- **Impact**: better assessment of the benefit-risk trade-off, awareness of the issues addressed.

Level of satisfaction/acceptability during demonstration/use

- **Level of satisfaction** rated on a scale from 0 to 10 (zero: not satisfied / 10: very satisfied),
- **Dissatisfaction and satisfaction criteria,**
- **Reuse of the tool when available**, reasons for reuse and non-reuse,
- **Recommending the tool to other professionals.**

Evaluation appears to be an essential step in ensuring the rigor and relevance of a healthcare chatbot if it is to be adopted by practitioners in their day-to-day practice. As with medical devices, the answers provided by a chatbot should be validated scientifically and medically by a health authority. This would both build trust among users and certify the medical quality of such tools. In the following paragraphs, we would like to propose several key parameters that we believe are essential for effective evaluation of a chatbot.

First, the reliability of the chatbot’s answers should be verified. This should be done by checking the sources on which its answers are based and whether there is a scientific consensus about them. These might include, for example, answers that reflect official recommendations, validated by the health authorities. With regard to information on medicinal products, the answers to questions about a product must be based first and foremost on its MA (Marketing Authorization) and on the official documents of the health authorities. The reliability of the answers can be assessed with the natural or legal person producing the chatbot before it is commissioned. This first part of the evaluation would be indispensable and crucial for building trust in the chatbot.

The second key point of the evaluation concerns the use of the chatbot. This would involve checking the accessibility of the tool, but also its ease of use, the fluidity and user-friendliness of the interface and the average time taken to get an answer. It would also be essential to assess the chatbot’s ability to
understand the user. This should be done by measuring the rate of relevant answers given by the chatbot, i.e. the ability of the chatbot to understand the user’s intention and provide the corresponding answer. All this would make it possible to measure the user experience of a chatbot. This user experience can be evaluated during the development of the chatbot in order to integrate feedback well in advance and at the time of commissioning.

In addition to the two points mentioned above, a third factor could be taken into account to evaluate the perceived contribution, or the medical service provided by a chatbot. The aim would be to evaluate the contribution of the chatbot in terms of improving public health and according to the use case in question. It would also be possible to measure the impact of the use of a chatbot on the practices of health professionals or on patients, taking into account different criteria such as the frequency of use of the chatbot (occasional or regular), integration into current practice or satisfaction after use.

As seen above, innovative health solutions raise the issue of their evaluation. To support the development of chatbots, health authorities could define a standardized evaluation scale that would take into account all the criteria mentioned. This would reinforce the trust of users, patients or healthcare professionals, in this technology.
What are the key success factors for healthcare chatbots?

What seems essential is the medical and scientific validation of a chatbot’s answers. How relevant are the answers to the questions asked? Have these answers been validated by the appropriate health authorities or by what is known as evidence-based medicine? In my view, that’s the first point. Secondly, for a chatbot to be successful, I think you really need to work on its ease of use. Is the chatbot easy to connect to? How long does it take to get an answer? Is the answer clear, easily understandable and appropriate for the person asking the question? The answer should be different depending on whether it is meant for a patient, a doctor or another healthcare professional. The user experience with the chatbot interface should be as smooth and intuitive as possible. There may be a third point, which concerns the integration of chatbots into the tools that doctors use daily in their consultations with patients. This must be extremely smooth to help doctors save time.

Who requests the evaluation? For what purpose?

This could be a bit like the validation of medical software, which can be considered to be medical devices. In so far as diagnosis or treatment information is being provided, assessment is almost mandatory. An evaluation could thus be requested by the health authorities, and this would mean the use of chatbots has been validated.

At what stage in its development should a chatbot be evaluated?

For the first evaluation requirement concerning answer reliability, the evaluation should be carried out before the chatbot is commissioned by the person or organisation that produced it. Before making such a tool available to doctors or patients, the reliability of answers should be assessed, together with how the answers have been structured and which sources have
been used. Without such an evaluation, chatbots should not be allowed to be published online. In addition, as far as user experience is concerned, once a chatbot has been published online and its answers have been verified as being reliable, its ease of use and the user experience can clearly be assessed through the doctors and patients using it.

**What is the current status concerning the evaluation of healthcare chatbots?**

There is very little out there in terms of evaluation. To my knowledge, there are no specific examples of chatbots being used extensively following careful testing and publishing of the results. This is because it is a brand-new area, which is still being shaped. It can be assumed that evaluation results and chatbot solutions will begin to appear in specific medical fields in the months to come.

**What would you say as a concluding remark?**

The key word is "demonstrate" in order to build confidence in the tool. People may be wary of talking to an artificial intelligence device that provides answers to their questions, this is only natural. Therefore, we need to demonstrate, through evaluation, that the results produced can be trusted. This is why evaluation is absolutely key and vital.
CONCLUSION

WHAT WILL HEALTHCARE CHATBOTS BE USED FOR IN THE FUTURE AND WHAT CHALLENGES WILL THEY MEET?

Our study of uses shows that healthcare chatbots have not yet reached maturity yet. They are only in the early stages. However, it seems certain that their use will be generalized in the near future. Chatbots are being increasingly integrated into homes, especially via connected speakers that become a daily companion and begin to address healthcare uses. Tomorrow, parents may turn to a vocal chatbot for advice when their child shows minor clinical signs. Patients will carry out a pre-consultation via a questionnaire on a chatbot at the time they make their appointment. Doctors will interview a chatbot to obtain specific information about a drug or pathology.

There are three main reasons why chatbots could improve and guarantee access to patients’ health information:
• The use in the chatbot’s answers of vocabulary adapted to the user
• The non-stigmatizing quality of the exchanges for the user (no questions are “too stupid” for a machine, there are no taboo subjects)
• The availability of a voice version, for example for visually impaired or illiterate people

Chatbots could also help adapt the habitat for home-based healthcare through oral communication:
• By promoting access to home automation
• By enriching current remote assistance offers with an alert system based on user responses
• By providing a socializing activity (triggering a phone call or video call, discussion)

However, all new technologies, including chatbots, could be integrated into uses only on the condition that health ethics are respected. For example, confidentiality is important, as is data security.

While many simple or repetitive tasks are and will increasingly be handled by chatbots, robots will never replace humans until technologies are able to understand emotions and unexpressed, implicit information. This is even truer in the health field.

If you have questions after reading this, please contact the authors!

49
Who helped draft the white paper at Sanofi?

- Mercedes Courtois
- Victor Dussart
- Kristina Fauvelle
- Joël Le Pape
- Brigitte Leymarie
- Denis Vigouroux

What about Orange Business Services?

- Pauline Derrien
- Fanny Jardinaud
- Eric Taranson

What about Kap Code?

- Jason Koman
- ADEL Mebarki

Our thanks to all!