

# Chatbots in Professional Contexts: Building on Real Uses for Chatbot Design

Empirical studies on the implementation and adoption of three chatbots in professional contexts

**Marion Gras Gentiletti<sub>1, 2</sub>, Gaëtan Bourmaud<sub>1</sub>, Myriam Fréjus<sub>2</sub>, Françoise Decortis<sub>1</sub>**

*1) C3U-Paragraphe Laboratory, Paris 8 University, 2 rue de la Liberté, 93200 Saint-Denis, FRANCE*

*2) EDF R&D Boulevard Gaspard Monge, 91120 Palaiseau, FRANCE*

How do chatbots transform various work-related activities within a mid-size company? How can chatbots make our lives easier at the workplace? What are the conditions for their successful implementation? To answer these questions, this paper investigates how three chatbots were implemented and adopted in a professional context based on real user activity. First, we show how chatbot design processes centered on “innovation” are built on a fragmented vision of human activity. Second, we illustrate how the evolutive capacity of chatbots is inhibited by the compartmentalization of sociotechnical system actors. Next, we explore the emergence of potential forms of usage relevant to the aforesaid actors. To conclude, we discuss the potential advantages of designing chatbots based on actual user activity. The main objective of our research is to demonstrate the relevance of chatbot design processes focused on user groups, AI pertinence, and the evolutive nature of human-chatbot interactions.

**CSC CONCEPTS** • Empirical studies in HCI • Artificial intelligence • Design

**Key words and additional phrases:** chatbots, real user experience, user-centered design

**ACM Reference Format:**

## 1 INTRODUCTION

Advances in Artificial Intelligence (AI) since 2010 have seen a growing number of AI-based interactive services emerge, the most successful of which is the chatbot [43]. Using automatic natural language processing algorithms, chatbots have the potential to communicate autonomously with humans in various interactive modes (textual, vocal) [7, 8]. Furthermore, they are theoretically capable of learning and evolving through experience, for increasingly contextualized and personalized human-chatbot interactions [18, 42]. The chatbot, or text-based conversational agent (CA), is widely used in professional settings [9, 35, 41]. Available 24/7, chatbots provide businesses with fast, convenient, and cost-

effective solutions in line with today's customer relations culture increasingly focused on the ongoing improvement of the customer experience [16, 28, 41, 46] or the employee experience [45, 3].

### ***The HCI revolution***

Appeared in 1966, with the Eliza program designed by J. Weizenbaum to simulate a Rogerian psychotherapist, chatbots have met with real success since the beginning of the 2010s especially for the new possibilities to link social interactions [37]. The technological advances of today's chatbots make it possible to assess the benefits of chatbot implementation by empirically investigating real user practices [9, 14, 22, 23, 24, 25, 44]. These efforts are part of a wider movement bringing together AI and HCI communities through a shared interest in human-centered AI [1, 23, 26, 36]. Yet prior works have shown a technocentric trend in AI-based system design, including chatbot design [17, 19]. From the perspective of the activities "replaced" by chatbots, human input is seen as a means for rectifying unforeseen events that machines are incapable of managing [17]. From the perspective of the end users, human interactions contribute to the evolutive capacity of the chatbot so that, by guiding chatbot interactions, they ultimately provide their own services [19]. Whether facilitator, trainer, supervisor, assistant, or self-producer, the human user is thought of as an aide to the chatbot in the design of interactive systems [14, 15, 21, 31]. The Human-Computer interaction (HCI) community aims to bring about a paradigm shift: to place humans (the users), and not technologies, at the center of sociotechnical system [31]. The emergence of chatbots – which Microsoft likens to the Internet revolution – has nevertheless opened the field of HCI research by letting humans interact with systems in novel ways [8]. For instance, advances in AI enable natural language user-chatbot communication, augmenting thus the chatbot's evolutive capacity [21]. HCI researchers therefore need to reevaluate user-centered design processes, notably by readjusting their methods to better understand and evaluate how humans utilize evolution-capable interactives devices [18, 25, 41].

Numerous works investigate these issues empirically, drawing on data from questionnaires [20] and usability tests [7, 25, 41] to offer recommendations aiming to authenticate the usability of chatbots [7, 8, 13], the quality of interactions [21], the motivations [7] and expectations of users [30]. Design recommendations are thus provided, like clearly tell people what tasks the chatbot can do and prevent false expectations or save information from one task to the next [9], make clear how well the system can do what it can do [1], use context awareness (weather and time of day for example) [13], or implement in the chatbots social knowledge (humor, intimacy) [32, 13, 22].

### ***Implementing an approach centered on real use situations***

Yet there are authors who believe that real uses and design processes need to be considered simultaneously [5, 39]. The underlying argument is that the user's activity is an ongoing process constantly shaped by the emerging situation [33, 38]. It must therefore be described in real situations viewed intrinsically from the user's perspective [38]. Such an approach can provide insights into what users in real-life situations consider a relevant interaction, while simultaneously making allowances for the time-based, context-specific, and singular nature of human activities [4, 27, 28, 33, 40]. This shift towards real user activity enables us to understand how chatbots can benefit and transform human activities, but also how they impact group interactions. How do chatbots transform various work-related activities within a mid-size company? How can chatbots make our lives easier at the workplace? What are the conditions for their successful implementation? The goals of this paper are to understand the issues around integrating chatbots into real-world work settings and to demonstrate the advantages of an analysis of real user activity for the design of chatbots. We have therefore examined the ways in which chatbots are implemented and utilized by exploring how different types of user groups interacted with three specific types of bots.

We are also interested in determining how success in the design and deployment of AI-based systems could be defined today. Following a presentation of the empirical findings of our study on the introduction of three chatbots in a professional context, we will discuss how the study's results could impact the design of interactive systems in the future.

## **2 CONTEXT**

The company in which the study is taking place has been betting for several years on the development of virtual assistants to help the businesses and the customers of the company. These assistants are embodied by chatbots, also called "conversational agents", which consist of a computer program capable of simulating a conversation with one or more humans by textual exchange. This study, requested by the Information Technology Services Department (ITSD), offers an in-depth analysis of the insertion of three chatbots into business practices and their reception by users. These chatbots are based on a natural language understanding algorithm (Natural Language Processing), a structured and scalable and evolving knowledge base, and a control interface for the people building the knowledge base. This interface also provides access to usage statistics.

### 3 METHODOLOGY

Our data collection is built with the aim of having data allowing to understand the context of introduction and the effects of the introduction of chatbots on the human activity. To do this, it seems relevant to us to access a variety of situations instrumented by chatbots, both in terms of their advancement and their field of application. With this in mind, three chatbots already deployed in the company have been selected:

- the “**virtual jurist**”,<sup>i</sup> designed to provide the company users<sup>ii</sup> with answers to frequently asked legal questions;
- the “**virtual researcher**,” designed to provide new researchers with domain-specific information through simulated exchanges with experienced retired colleagues;
- the “**virtual human relations consultant (HRC)**,” designed to provide users with answers to recurring HR questions.

The study therefore focused on the different groups of people revolving around these chatbots, which an exploratory study distinguished into four categories: **users**, the people who use chatbots; **business lines** (legal experts, human relations consultants, business travel consultants, researchers), the areas where conversational agents can replace human activity; **project managers**, the managers who initiated the request for the chatbot implementation project; and **chatbot supervisors**, the people who ensure long-term chatbot operations, specifically the analysis of agent-chatbot conversations.

We conducted 27 semi-structured interviews lasting about 60 minutes each (See Table 1) with participants from the above categories (business lines, chatbot supervisors, users, and project managers). On the other hand, an exchange of one hour and a half with all the chatbots designers (about fifteen people) was conducted.

A presentation of the study and the research objectives was made before starting the interview around a grid of themes to be addressed. These exchanges also offered the interviewees the possibility of moving the interview to other themes giving rise to discovery. The dimensions of the socio-technical context and the activity to be documented for each of the subject groups (users, business lines, project managers, chatbot supervisor) were defined upstream and dealt with the following themes: expectations related to the chatbot project and the tool, role of the participant in the chatbot design process, description of the activities related to this role, description of the uses around the chatbot and objects of the activity, positioning of the chatbot in relation to existing instruments systems.

Lived experience of activity interviews [10, 11] were conducted with users to collect information about real-life use situations. Moreover, four real use situations were filmed and discussed “on the spot” (See Table 2). In addition, a study was conducted drawing on a corpus of documents, including the company’s user forums, its literature supporting the

design process, its internal and external publications presenting the chatbots, and its chatbot back office (corpus of around thirty documents relating to chatbots).

Table 1 semi-structured interviews: profiles of the participants met

	Virtual jurist	Virtual HRC	Virtual researcher	Total
<b>Business lines</b>	-1 jurist (m, 45) who participated in creating the chatbot knowledge base -1 jurist (m, 36) who participated in creating the chatbot knowledge base	-1 HRC (f, 29) who participated in creating the chatbot knowledge base	-1 retiree (m, 64): researcher whose experience was integrated into the chatbot -1 retiree (m, 65): researcher whose experience was integrated into the chatbot	<i>5 Business lines</i>
<b>Chatbot supervisors</b>	-The "virtual jurist" supervisor (m, 25)	The "virtual HRC" supervisor (f, 28)	The "virtual researcher" supervisor (f, 56)	<i>3 chatbot supervisors</i>
<b>Users</b>	-1 user (f, 36): regular use -1 user (f, 32): regular use -1 user (m, 54): only one use -1 user (m, 29): only one use -1 user (f, 28): only one use -1 user (m, 43): only one use	-1 user (f, 32): less than 5 uses -1 user (m, 52): only one use -1 user (m, 29): only one use -1 user (f, 28): only one use	-1 user (m, 36): less than 5 uses -1 user (m, 28): less than 5 uses -1 user (f, 41): only one use -1 user (m, 29): only one use -1 user (m, 28): only one use	<i>15 users</i>
<b>Project managers</b>	-The "virtual jurist" project manager (f, 34)	-The "virtual HRC" project manager (f, 32)	-The "virtual researcher" project manager (f, 48) --The "virtual researcher" project manager (m, 51)	<i>4 projects managers</i>
<b>Total</b>	<i>10 subjects</i>	<i>7 subjects</i>	<i>10 subjects</i>	<i>27 subjects</i>

Table 2 : real-life use situations observations: Profiles of the participants met

	Virtual jurist	Virtual HRC	Virtual researcher	Total
<b>Users</b>	-1 user (f, 32): first utilization <i>Is searching for trademark registration information</i>  -1 user (f, 28): less than 5 uses <i>Is looking for information to carry out a legal contract</i>	-1 user (m, 32): second use <i>Is looking for information on saving for vacations</i>	-1 user (f, 33): only one use <i>Is searching for information on a nuclear component</i>	<i>4 users</i>

We conducted a qualitative data analysis [12] starting with a floating attention. The aim of such floating attention is to allow treatment categories to emerge which were not initially identified. Several themes thus emerged, in addition to those we wanted to document. They result from a work of resemblance / differences which consists of coding the units of meaning in multithematics. Each unit thus refers to several thematic series. Using a hypothetico-deductive model, chatbot usage themes and categories were determined, making it possible to classify the findings in terms of design processes, obstacles, and implementation conditions. A back and forth between facts and hypotheses finally brought out themes and categories of treatments, in order to highlight and organize the elements.

The first readings of the data and their hypothetico-deductive analysis led to the development of a systematic analysis grid, composed of the following elements:

- The chatbots design process;
- Understanding of situations instrumented by chatbots: components and determinants of the activity; user intentions; systems of instruments mobilized by people;
- The constraints encountered by people in the use of chatbots and the strategies implemented to overcome these difficulties;
- The reorganization of work activities after the insertion of chatbots;
- The functions of the chatbots planned by the designers and mobilized by the users and the functions created in the use, the forms of appropriation or abandonment of the chatbots by the users.

Finally, the documentary corpus was the subject of a discourse analysis, in particular through concepts derived from the sociology of expectations [2, 6, 29]. Themes thus emerged, from a floating attention then from a hypothetico-deductive analysis, allowing in particular to identify the following semantic categories: Culture of innovation (users implication, motivations...) and Characterization of technoscientific promises (users expectations...).

## **Empirical findings**

### **3.1 An approach to design centered on “innovation”**

*Using AI to innovate – come what may.* Firstly, our study of CA design processes reveals how chatbot implementation conditions are tied to aspirations of innovation. For instance, the introduction of the “virtual jurist” is part of a department-wide strategy promoting legal staff digital acculturation (*project managers interview*). The “virtual jurist”, inducted by the Legal Department in 2017, is a good illustration of the conditions for introducing this type of innovative device, by being part of a project initiated to strengthen “*the acculturation of the Group’s lawyers to digital technology*” (*discursive analysis of the corpus*). The outline of the project emerges during a “Digital DJ” day organized in 2016 to “*reflect on digital tools and innovative methods essential to the transformation of the legal function*” The flagship tool that is imagined is a chatbot presented as a “*conversational robot responding to thousands of recurring requests made daily and helping lawyers draft contracts*” (*discursive analysis of the corpus*).

Relieved of their tasks, the lawyers now focused “*on complex issues that are at stake for the Group*”. Beyond its role in the organization of work, this new tool is being extended, characterized, in internal and external communications, by elements of language around its innovative character (*discursive analysis of the corpus*).

Winner of two innovation prizes, the “virtual jurist” is also presented in the company as a benchmark for the development of internal (HR, purchasing, etc.) and external (customer relations) chatbots: “*And then if the chatbot has won a Pulse Prize, it's also because even within the group, it was something recognized as innovative*” (project manager interview).

The implementation of the “virtual researcher” attests to a similar innovation-focused design, the R&D department being mainly concerned with the transfer of knowledge between researchers nearing retirement and those recently recruited. Though an existing method for knowledge capitalization, which includes a series of digital factsheets, is perceived by the staff in a very positive light (*business lines interview*), the top management is critical of the method, which it sees as not being innovative (*project manager interview*). The team therefore opted for a technological solution – the chatbot – in line with its “*forward-looking vision*” (*project managers interviews, discursive analysis of the corpus*).

**A failed user-centered design process.** The “end user” is generally represented by “typical users”, personas (archetype of a group of people) or even design actors who put themselves in the shoes of “end users”. In addition, the “end user” is often called in during the downstream phase of the design to assess the acceptability of the service. As the company’s business lines were not initially considered in the design process, it was hard for the lines of business (LOB) managers to imagine the future benefits of using chatbots in the completion of day-to-day tasks, as one jurist attests, for example, who tells us: “*I knew that no one would ever ask these questions [to the chatbot] because it is too specific an area but we had to ask as many questions as possible quickly. So I listed some definitions*” (*business lines interviews*).

By the same token, the users were overlooked in the design process: it was left up to the LOB managers to determine how the chatbots could support the agents by “*playing the part of the employee interacting with the chatbot*” (*discursive analysis of the corpus*). Such an approach, although grounded in a user-centered vision, entails a number of risks. The jurists, for example, may find it difficult to put themselves in the shoes of a user with little legal knowledge, owing to their own legal expertise. In addition, the chatbots are not deemed by the users to be more effective than existing work methods, despite the touting of their innovative capabilities (*users interviews*), as a researcher illustrates: “*In our department, when we need help, there are the [digital] factsheets. The chatbot refers us to them, so why shouldn't we consult them directly?*” (*user interview*).

**Unsatisfactory encounters owing to the gap between the user’s expectation and the system’s performance.** The emphasis is placed on chatbot’s capacity for “human-like”

communication, instead of the actual services it provides. The three chatbots are thus described as very efficient devices. Lastly, there is the issue of social presence (humor, empathy, etc.), the data integrated into the chatbots' knowledge base to create an impression of warmth. For example, the "HR chatbot" is featured as a video thus leading the agents to expect seamless human-chatbot interactions similar to natural human-human exchanges, whereas this is not the case. The queries recorded in the "virtual HRC" back office testify to these expectations. Entries include such polite requests as "*Hi, could you please tell me [...]?*". One of them is even featured in videos while the others are showcased in prolific external communication (*discursive analysis of the corpus*).

The group's chatbot valorization plan also focuses on performance improvement, on the optimization of business line activities and users satisfaction:

*"AI capabilities enable the chatbot to understand the user's query and to retrieve the right answer from the legal knowledge base compiled by our jurists. If the user's request requires further precision, the chatbot asks questions to clarify it. [...] The chatbot therefore allows users to concentrate on the cases that require value-added analysis. The time saved makes the staff more efficient and productive, while contributing to employee well-being,"* a legal department manager commented on the company's legal intranet (*discursive analysis of the corpus*).

Consequently, the users are inclined to compare the company chatbot to others CAs capable of simulating conversations, which may, in the case of unsatisfactory earlier experiences, make them less inclined to comply with requests made by the chatbot. By contrast, satisfactory interactions with other chat interfaces tend to lead to higher users expectations. One user we interviewed mentioned a disappointing experience with the "HR chatbot" The user, a regular user of a mass-market CA deemed highly effective, was disappointed with the business travel chatbot's functions.

The chatbot's usefulness proved disappointing from the business line perspective, however. The departments were confronted with a technical device partly requiring staff support: "*We were expecting too much in terms of what the tool can actually provide*" (*business lines interview*).

### **3.2 The design process built on a fragmented vision of human activity**

***The social and material reality of today's world ignored.*** The users use a range of resources in their day-to-day activities, notably computer-based resources. Consequently, they expect new devices to be designed according to standard interaction principles. The current approach to chatbot design which ignores prior user experiences and expectations results in a lack of usability. In one real-use situation we observed, a "virtual jurist" user



attempted to click on a highlighted, underlined word but no link opened. The user had interpreted the underlined word as a potential meaningful event (a hyperlink click) (*real-life use situations observations*). Observations as well as data from the chatbots' back-office show that, although the designers recommend that queries be formulated in full sentences, most agents use key words when entering requests (*discursive analysis of the corpus*). A lack of visibility among existing resources may also cause users to be less inclined to use the device. For example, accessing the law department bot is particularly difficult and therefore, discouraging. "*I went there on purpose [the company website] and typed 'chatbot' in the search bar. Instead of opening the app, it pulled up articles linked to empty pages. [...] I have better things to do with my time...*" (user interview).

A comment left on the company portal gives another glimpse of the impression left by the integration of the chatbot into the existing instrument system: "*Hello, I've been trying to access the chatbot for ten minutes. Legal Department: a great innovation put forward by EDF... but not by [the company portal] unfortunately ;-)* You seem annoyed with the links to useful applications... Am I the only one wandering from page to page without finding of link to our applications, because the search engine links to articles on the subject but not to useful links? Thanks for your support because maybe I'm using [the company portal] the wrong way (try typing "virtual lawyer" into the search engine and you'll see what I mean). Kind regards. " (*discursive analysis of the corpus*).

On the business line end, coordinating the chatbots with the existing tools is also problematic: "*We work in business travel, so it's a bit tricky for us, using a tool that can only handle requests, and another tool to deal with complaints.*" (*business lines interview*). Here, the support team expresses the difficulty of users in distinguishing the tool on which they can express complaints and therefore a personalized treatment of a problematic file from the chatbot tool dedicated to so-called general questions. This difficulty rebounds and complicates the work of the business lines team, which finds itself dealing with unhappy users because of their unprocessed cases. However, the chatbot, by anonymizing the conversations, does not allow the support team to process a request that has been made by a user. This case once again underlines the failure to take into account existing instruments (here, the online complaint tool) in order to adjust the role, form or even the very relevance of the chatbot. This point also underlines the limits of a typical replicated chatbot even though there are a variety of situations.

***Some tasks are eliminated and others... created.*** Our study shows that the chatbots were introduced by the projects managers to help with recurring tasks with limited added value (*project managers interviews, discursive analysis of the corpus*). This task-oriented vision points to a lack of business-line-activity systemic analysis and, as such, makes it

difficult to assess the ways in which the introduction of chatbots may affect day-to-day activities, namely through the creation of new tasks. The users, for example, turn to the group's jurists to validate information retrieved from the virtual jurist. One jurist received an e-mail saying: *"I entered the following into the chatbot, this is the answer I got. Could you please confirm the response, namely that if I do decide to make a movie tomorrow [...] I need to ask for an authorization?"* (user interview). In this case, the first level of information is no longer communicated by the jurists, who control it, confirm it, and track it instead. Though a quantitative analysis of the position's task-load show that the jurists have fewer recurring questions, the introduction of the chatbot has created a new verification task for the department (*business lines interviews*).

### **3.3 The evolutive capacity of the chatbots inhibited by a lack of interpersonal interaction**

#### ***A flawed vision of the so-called "successful conversation."***

chatbots are described by the company as autonomous conversational agents (*project managers interviews, discursive analysis of the corpus*), whereas they actually require a significant amount of human intervention. First, there is the knowledge base, which requires a cross-department collaboration between the business lines and the chatbot supervisors. Secondly, once the chatbots are up and running, the knowledge bases need to be gradually expanded through the analysis of failed requests. The evolutive capacity of the chatbots is therefore directly tied to content monitoring and to the compiling of new knowledge. Specifically, an analysis is made of the conversations the machine considers "unsuccessful," i.e. the requests the chatbot is unable to assign answers to. Conversations characterized as "successful", and again this is from the perspective of the machine, are the interactions that enable the chatbot to relate a question to one or more pieces of information. However, from the users's point of view, the chatbot's ability to connect a query with a response does not necessarily constitute a satisfactory conversation. One user we observed in a real-work situation commented on her actions and her feelings as she enquired about the company's trademark registration. After first consulting other resources (Internet search, call to a colleague...), she asked the "virtual jurist" for assistance. The chatbot provided her with five answers, constituting, from the perspective of the machine, a successful conversation (*real-life use situations observations*). However, from the user's perspective, the answers were not entirely satisfactory: they are related to her question but do not provide her with the information she is looking for. In one answer, for example, she was prompted to contact the legal department: *"It's asking me to contact the legal department. [...] That's what I expected. It is the legal chatbot, after all..."* she remarks. Her assessment of the experience is therefore mixed: the chatbot provided answers related to

her query but the answers did not contain the requested trademark registration information. But again, from the perspective of the machine, it was a successful conversation.

At the end of the exchange, the user said she would like to give feedback, to contextualize her query and express how she feels about the interaction. The chatbot responded by asking, "Are you satisfied?" followed by "YES or NO." Since the user's experience is nuanced – she is neither entirely satisfied nor entirely dissatisfied – she decided to exit the page without leaving a comment. If no information on either the interaction context or the user's impressions is provided, the chatbot supervisor will classify the conversation as successful and it will not be analyzed (see Figure 1<sup>1</sup>). This absence of actor-to-actor communication is a failed opportunity to develop the system to better meet the end users' needs.

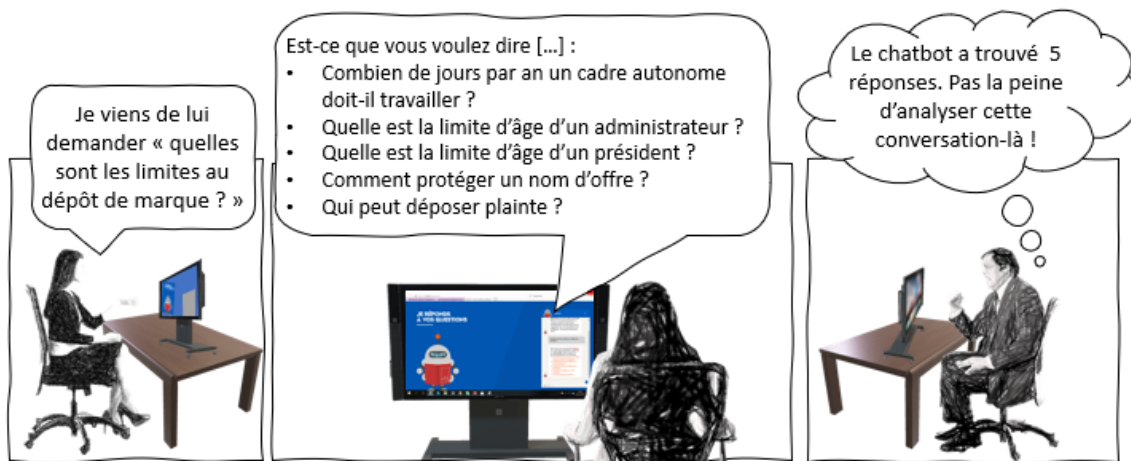


Figure 1 "Successful" conversations set aside for analysis

**Requests that supervisors have a hard time interpreting.** Ultimately, the users are not given the chance to provide chatbot performance feedback, and the satisfaction questionnaire often leads to supervisor confusion: "Do they [the users] say 'no, not satisfied' because the right answer was not provided or because they dislike the answer they were given?" (*chatbot Supervisor interview*). This lack of a dedicated space for user feedback further limits the chatbot's evolutive capacity in that it influences whether or not the chatbot's responses are considered relevant. Lastly, this absence of interpersonal communication

<sup>1</sup>Translation :

Speech bubble 1: I just asked him "what are the limits on trademark registration?"

Speech bubble 2: [the chatbot answers] Do you mean: How many days per year must an autonomous manager work? What is the age limit for an administrator? What is the age limit for a president? How to protect an offer name? Who can file a complaint?

Speech bubble 3: The chatbot has found answers. No need to analyze this conversation! It is successful!

complicates the work of the supervisors who often find themselves obliged to analyze conversations without any context whatsoever provided: “*This is the request I am supposed to interpret: ‘Subcontractor.’ It’s impossible*” (chatbot Supervisor interview).

**The end users want to be part of the chatbot improvement process.** Even so, the end users are interested in taking part in the chatbot’s evolution: “*I’d like to give it some feedback*” (real-life use situations observations). Numerous comments on the company portal attest to the end users’ willingness to contribute to the chatbot’s evolution, beyond expressing their dissatisfaction with the lack of a dedicated quality assessment space (discursive analysis of the corpus). For example, one user wrote on a dedicated chatbot forum: “*Hello, I just asked [the chatbot] a question about illegal subcontracting and received the following answer, ‘illegal favoritism’ [...] Perhaps it’s some sort of a misunderstanding?*” When asked about the comment, the author said it was a message meant to help the supervisors improve the chatbot’s performance: “*I wanted to give the programmers and jurists feedback to help improve [the chatbot]. That’s why I posted the comment*” (user interview).

### 3.4 Unforeseen forms of use and promises of added value

Back office data show that chatbot use is often quickly abandoned (discursive analysis of the corpus). For instance, at the time of its organization-wide rollout in January 2019, the law department chatbot recorded 600 interactions. By March that same year, the number had dwindled to 200 interactions a month.<sup>iii</sup> Nonetheless, the interviews and observations show the emerging uses had not been anticipated during the design phase.

For example, the chatbot is mainly used by chatbot supervisors to analyze conversations, but it can be used as a monitoring tool to **predict needs** as well: “*Once a month, it allows me to check my back office for conversations I may not be aware of, to see whether any employee comments come up when I enter keywords like ‘retirement,’ ‘reform,’ and so on*”<sup>iv</sup> (chatbot Supervisor interview).

On the business line end, the chatbot can be used to **store and centralize information**. The jurists, for example, use it to retrieve contract outlines (requests included in the 200 interactions a month mentioned above) The business travel support team gradually started doing the same. Furthermore, the chatbot makes it possible to **harmonize the answers provided to the users**: “*Legal matters entail a certain degree of interpretation, practices may vary from one site to another, as they do in every company. It’s [the chatbot’s harmonization capability] very reassuring for HR*” (business lines interview).

From the users’ perspective, the chatbot is seen as a tool allowing users to bypass existing channels to **make anonymous requests**. During one observation session, the user mentioned being uncomfortable with the idea of directly contacting a coworker, whose

contact information had just been pulled up, to request a piece of legal information: *“I’m not very keen on the idea of contacting this particular person since we’ve been out of touch for two or three years. [...] So, here’s her file. Well, apparently, she’s a department head now, she’s a manager which makes it even more awkward. [...] Well, I’d heard talk about the chatbot, so I think, “Why not! Maybe it will be able to help me get the information I need” (user interview).* The chatbot is also perceived as a tool for storing and **centralizing updated group-level information**: *“The problem, as we know, with [digital] records is that we each have a different version in six months’ time... [...] If the records are not regularly updated, we stop using them in the end” (user interview).* In addition, the chatbot is viewed as a good way of **gaining greater autonomy**, and especially as a way not to bother colleagues with too many questions: *“I enjoy discussing things with my colleagues – it’s good way to learn – but I felt like I was taking up too much of their time, if I consulted them too often, they’d grow tired of all my questions, if I asked too many questions, they’d be less inclined to answer... I didn’t want to be a nuisance” (user interview).* Lastly, the chatbot’s ability to expand the user’s initial query is another highly appreciated feature: *“I entered ‘image authorization.’ The chatbot [...] reprocessed my request with two follow-up suggestions – under 18 or adult. I found this very useful. It pulled up two documents: one for people under 18 and one for adults. I was pleased with this because when you work in the industrial tourism sector, you tend to forget that there are people under 18, families with underage children. The authorization specifically concerns the parents and not the child, so I was satisfied with the exchange because it brought to my attention the existence of additional content” (user interview).*

## DESIGN CONTRIBUTIONS

Our findings show that the chatbot design phase was mainly concerned with meeting the challenges of innovating and modernizing business entities. We believe that insights can be gained by creating a best practices framework for chatbot design in which the chatbot is considered not as an end but as a harnessed or potentially harnessable resource for actor activities across the sociotechnical system. It is particularly interesting to note that our ecological and developmental approach makes it possible to confirm recommendations identified in an experimental situation and specified in the Literature [3, 9, 18, 20, 22, 34, 35], such as:

- Clarify to user that the chatbot is software (not human);
- Be clear about the chatbot’s capabilities and disabilities so as not to create false expectations;
- Allow user to interact with a human if necessary by offering, for example, contacts of available and competent people;
- Help the user understand the origin of chatbot mistakes.

The contribution of our study is ultimately to show that a developmental and ecological approach makes it possible to provide results of a different nature, such as the creation of new tasks as a result of the chatbots and the ongoing management work involved, alongside the usefulness of ongoing user feedback. This is a novel area, HCI research often only considers the initial implementation. It is important to consider the ongoing management, maintenance and iterative improvement challenges that are faced when chatbots (and indeed other systems) are deployed in work environments.

On the other hand, observing real situations leads to the following recommendations:

- Allow the user a plurality of interaction methods (in particular the reproducibility of practices such as keyword requests);
- Make the chatbot easily accessible in the existing workspace.

Moreover, the diverse profiles of the members of the sociotechnical system orchestrated by the chatbot need to be examined, as do the interpersonal transactions that occur within the system. Indeed, an unusually wide variety of users are drawn into the chatbot's orbit, owing especially to the large amounts of data required to power AI, which still partially depend on the intervention of human agents. A user group-centered approach is therefore recommended to allow the bolstering of interpersonal transactions among the various sociotechnical system actors by identifying, from the user's perspective, what is considered a relevant response. Such an approach allowed us to show the existence of the gulf between the human perspective and that of the machine, in terms of what constitutes a "successful conversation." Furthermore, it reveals the diverse identities of the users covered by the blanket term "user," who need to be considered during the design phase. Supporting spaces of expression between the different types of users (chatbot users and supervisors for example) will allow to contextualize the requests (even the requests considered today as "successful") in order to make the chatbot always more relevant by in relation to user needs.

The need to implement solutions centered on the evolutive nature of human-chatbot interaction. Chatbot implementation projects present a dilemma: should chatbots be introduced as finished systems capable of providing innovative, high-performance solutions, or as systems "in-the-making?" In the first case, the designers and project managers wait to deploy the system at a very advanced stage. But to be relevant from the agents' perspective and to hone the chatbot's knowledge, the chatbot needs to learn from stored interaction data. We therefore recommend the second option: to test the prototype early on, so that it can accumulate data quickly and be developed to best meet the agents' needs. The drawback here is that it might discourage the agents, whose initial experience with a system "under development" may lead them to quickly abandon the technology and

not to return to it at a later date. In the second case, the chatbot's technical capabilities, and in particular its evolutive capacity, need to be clearly presented so that users know what it can (and cannot) do. Particular emphasis should be placed on the user's contribution as an actor in the design process. This is particularly true for the agents, who have little to do with the design process and who may be unaware of the system's operating logic. To fully include them in the design process loop, a space is required for providing feedback and to contextualize their requests after each interaction.

It is true that some of these recommendations may seem trivial to the HCI community. Nevertheless, this empirical study shows how important it is to make them explicit in the industrial world in which they still struggle to be considered.

## **DRAWBACKS AND PERSPECTIVES**

The need to implement solutions centered on AI relevance. Different forms of chatbot use raise the question of whether it is relevant to integrate AI into chatbots or not. Some emerging practices (storing information and documentation, processing anonymous requests) do not in fact require the integration of learning algorithms. Nevertheless, other practices could be improved through AI technologies, collecting agent feedback, for example. The advantage in this case would be to have a tool capable of progressively finetuning its own assessments so as to help the supervisors select the content to be created and stored for the agents. We believe that approaching design as a usage-based enhancement loop allows AI to be considered as a useful systems addition, and not as an end goal.

For businesses, the financial and strategic stakes of a chatbot implementation project are extremely high. Not only are chatbots a time- and money-saving business solution, they are indicative of a forward-looking company driven by innovation. Design processes implemented in a short-term timeframe but periodically involving the business lines (in choosing the chatbot interface, for example) give the illusion of a design process built around a user-centered project, making the intervention of UX professionals appear unnecessary. Our study shows that factors like these make it difficult to consider the various actors of the sociotechnical system. We therefore recommend the integration of design processes early on for future chatbot implementation projects.

The results of this study have been made available to all parties concerned. Meetings will be held in the coming months to further document the implementation of the recommendations, their impact on the design process, and the ways in which the chatbots are used by different user categories.

## CONCLUSION

The results of our empirical study show that chatbots do not replace human activity but rather disrupt it and transform it. It is urgent today that data based on user activity be put back into the design process, as organizations driven by a distorted vision of task automation continue to cut back on human resources. It is therefore necessary to reconsider chatbots as a means for supporting human activity (rather than an end), namely by basing design on real user experiences, by including all members of the sociotechnical system in the design process, and by integrating emerging forms of use from the start of the chatbot implementation project. Such a shift will enable a better understanding of the relevance of AI, by framing it as a positive contribution to the activities of sociotechnical system actors: agents, supervisors, lines of business. This same shift moreover will make it possible to break with the trend of *innovation-driven chatbot design*.

## References

- [1] S. Amershi *et al.*, « Guidelines for human-AI interaction », in *Proceedings of the 2019 chi conference on human factors in computing systems*, 2019, p. 1–13, doi: <https://doi.org/10.1145/3290605.3300233>.
- [2] M. Audétat *et al.*, *Sciences et technologies émergentes: pourquoi tant de promesses?* Hermann, 2015.
- [3] N. Asher, *A warmer welcome: Application of a chatbot as a facilitator for new hires onboarding*, 2017.
- [4] P. Béguin et P. Rabardel, « Designing for instrument-mediated activity », *Scandinavian Journal of Information Systems*, vol. 12, n° 1, p. 1, 2000.
- [5] S. Bødker et C. N. Klokmoose, « The human–artifact model: An activity theoretical approach to artifact ecologies », *Human–Computer Interaction*, vol. 26, n° 4, p. 315–371, 2011.
- [6] M. Borup, N. Brown, K. Konrad, et H. Van Lente, « The sociology of expectations in science and technology », *Technology analysis & strategic management*, vol. 18, n° 3-4, p. 285–298, 2006.
- [7] P. B. Brandtzaeg et A. Følstad, « Chatbots: changing user needs and motivations », *Interactions*, vol. 25, n° 5, p. 38–43, 2018, doi: <https://doi.org/10.1145/3236669>.
- [8] P. B. Brandtzaeg et A. Følstad, « Why people use chatbots », in *International Conference on Internet Science*, 2017, p. 377–392, doi: [https://doi.org/10.1007/978-3-319-70284-1\\_30](https://doi.org/10.1007/978-3-319-70284-1_30).
- [9] R. Budi, « The user experience of chatbots », *Retrieved December*, vol. 13, p. 2018, 2018.
- [10] B. Cahour et C. Licoppe, « Confrontations with traces of one’s own activity », *Revue d’anthropologie des connaissances*, vol. 4, n° 2, p. a–k, 2010.



- [11] B. Cahour et P. Salembier, « The user phenomenological experience: Evoking the lived activity with “re-situating” interviews », in *Workshop on Theories Behind UX Research and How They are Used in Practice, CHI*, 2012, p. 5–10.
- [12] J. M. Corbin et A. Strauss, « Grounded theory research: Procedures, canons, and evaluative criteria », *Qualitative sociology*, vol. 13, n° 1, p. 3–21, 1990.
- [13] G. Cameron *et al.*, « Best practices for designing chatbots in mental healthcare – A case study on iHelpr », 2018, p. 1-5, doi: 10.14236/ewic/HCI2018.129.
- [14] A. Casilli, « Antonio Casilli: "Workers are the Heart of the Algorithm" », 2017.
- [15] J. Cassell, « The ties that bind », *Reseaux*, n° 2, p. 21–45, 2020.
- [16] R. Dale, « The return of the chatbots », *Natural Language Engineering*, vol. 22, n° 5, p. 811–817, 2016.
- [17] C. Dubois, J. M. Salotti, D. Seminel, et N. Simonazzi, « Le chatbot: un outil de la relation aux clients », *Hermes, La Revue*, n° 2, p. 95–97, 2019.
- [18] D. Duijst, « Can we improve the User Experience of Chatbots with Personalisation? », University of Amsterdam, Amsterdam, 2017.
- [19] M.-A. Dujarier, « The three sociological types of consumer work », *Journal of Consumer Culture*, vol. 16, n° 2, p. 555–571, 2016, doi: <https://doi.org/10.1177/1469540514528198>.
- [20] A. Følstad et P. B. Brandtzaeg, « Users’ experiences with chatbots: findings from a questionnaire study », *Quality and User Experience*, vol. 5, p. 1–14, 2020, doi: <https://doi.org/10.1007/s41233-020-00033-2>.
- [21] J. Hill, W. R. Ford, et I. G. Farreras, « Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations », *Computers in human behavior*, vol. 49, p. 245–250, 2015, doi: <https://doi.org/10.1016/j.chb.2015.02.026>.
- [22] S. Holmes, A. Moorhead, R. Bond, H. Zheng, V. Coates, et M. McTear, « WeightMentor: A New Automated Chatbot for Weight Loss Maintenance », 2018, p. 1-5, doi: 10.14236/ewic/HCI2018.103.
- [23] K. Inkpen, « Does My AI Help or Hurt? Exploring Human-AI Complementarity », in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, 2020, p. 2–2, doi: <https://doi.org/10.1145/3340631.3395384>.
- [24] M. Jain, P. Kumar, R. Kota, et S. N. Patel, « Evaluating and informing the design of chatbots », in *Proceedings of the 2018 Designing Interactive Systems Conference*, 2018, p. 895–906, doi: <https://doi.org/10.1145/3196709.3196735>.
- [25] M. H. Jarrahi, « Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making », *Business Horizons*, vol. 61, n° 4, p. 577–586, 2018, doi: <https://doi.org/10.1016/j.bushor.2018.03.007>.
- [26] E. Kamar, « Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence. », in *IJCAI*, 2016, p. 4070–4073.
- [27] V. Kaptelinin, K. Kuutti, et L. Bannon, « Activity theory: Basic concepts and applications », in *Human-Computer Interaction*, Berlin: Springer, 1995, p. 189-201.

- [28] D. Lahoual et M. Fréjus, « When Users Assist the Voice Assistants: From Supervision to Failure Resolution », p. 8, 2019.
- [29] R. K. Merton, « Éléments de théorie et de méthode sociologique », 1965.
- [30] D. A. Norman, « How might people interact with agents », *Communications of the ACM*, vol. 37, n° 7, p. 68–71, 1994, doi: <https://doi.org/10.1145/176789.176796>.
- [31] D. A. Norman et S. W. Draper, *User centered system design; new perspectives on human-computer interaction*. L. Erlbaum Associates Inc., 1986.
- [32] D. Potdevin, C. Clavel, et N. Sabouret, « Virtual Intimacy, this little something between us: A study about Human perception of intimate behaviors in Embodied Conversational Agents », in *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, 2018, p. 165–172.
- [33] P. Rabardel, *People and technology*. 2002.
- [34] R. Ren, J. W. Castro, S. T. Acuña, et J. de Lara, « Usability of Chatbots: A Systematic Mapping Study. », in *SEKE*, 2019, p. 479–617.
- [35] M. Skjuve et P. B. Brandzaeg, « Measuring User Experience in Chatbots: An Approach to Interpersonal Communication Competence », in *International Conference on Internet Science*, 2018, p. 113–120.
- [36] L. Suchman et J. Weber, « Human-machine autonomies », *Autonomous weapons systems: Law, ethics, policy*, p. 75–102, 2016.
- [37] E. Stoekli, F. Uebnickel, et W. Brenner, « Exploring affordances of slack integrations and their actualization within enterprises-towards an understanding of how chatbots create value », 2018.
- [38] J. Theureau, « Course-of-action analysis and course-of-action centered design », *Handbook of cognitive task design*, p. 55–81, 2003.
- [39] P. Tolmie, « Orienting to the Wild », in *Into the Wild: Beyond the Design Research Lab*, Springer, 2020, p. 195–235.
- [40] P. Tolmie et A. Crabtree, « Deploying research technology in the home », in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, 2008, p. 639–648, doi: <https://doi.org/10.1145/1460563.1460662>.
- [41] C. Toxtli, A. Monroy-Hernández, et J. Cranshaw, « Understanding chatbot-mediated task management », in *Proceedings of the 2018 CHI conference on human factors in computing systems*, 2018, p. 1–6.
- [42] Transparency Market Research, « Chatbot Market », *Transparency Market Research*, févr. 10, 2020. <https://www.transparencymarketresearch.com/chatbot-market.html> (consulté le août 06, 2020).
- [43] J. Velkovska et M. Zouinar, « The illusion of natural conversation: interacting with smart assistants in home settings », 2018, doi: <https://doi.org/10.1145/3170427.3170619>.
- [44] J. Zamora, « Rise of the chatbots: Finding a place for artificial intelligence in India and US », in *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*, mars 2017, p. 109-112, doi: <https://doi.org/10.1145/3030024.3040201>.

[45] S. Zel, C. Bridgeport, et E. Kongar, « Transforming Digital Employee Experience with Artificial Intelligence », 2017.

[46] A. Zerfass, Á. Moreno, R. Tench, D. Verčič, P. Verhoeven, et others, *European Communication Monitor 2017: How strategic communication deals with the challenges of visualisation, social bots and hypermodernity: Result of a survey in 50 countries*. BrusselsEACD/EUPRERA, 2017.

---

<sup>i</sup> The names of the three chatbots were changed here for reasons of confidentiality.

ii Chatbots are designed for internal use. “Company agents” therefore refer to staff members who use chatbots. For greater readability, the term “agent” will be used throughout.

iii We must keep in mind that it is a group with nearly 150 000 employees.

iv In 2019, at the time of the study, the government was working on a pension reform proposal.