Shaping the Experience of a Cognitive Investment Adviser

Heloisa Candello[1(✉)], Claudio Pinhanez[1], David Millen[2], and Bruna Daniele Andrade[1]

[1] IBM Research, São Paulo, Brazil
{heloisacandello, csantosp}@br.ibm.com, bandrade2@gmail.com
[2] IBM Watson, Cambridge, MA, USA
david_r_millen@us.ibm.com

Abstract. In this paper we describe the design process of a multi-bot conversational system to assist people to make more informed decisions about finance. Several user activities were held to understand the experience of investment decisions, the opportunities to design financial cognitive advisers, and the user perceptions of such systems. Valuable information was gathered from four user studies which assisted the project team to decide what would be the best approach to help people to make more informed decisions about investments using technology. The user studies findings highlighted that financial decisions are made based on information people receive from friends, news, and social networks, which led us to explore intelligent systems that would gather such information and play the role of financial advisers in a multiparty conversational system. We discuss the main design implications of our studies in the context of a prototype called *CognIA* and conclude discussing several challenges of designing conversational systems.

Keywords: Conversational interfaces · Dialogue systems · Multiparty dialogue · User experience

1 Introduction

Our main research focus is the often-blurred connections between human and machines which result in better decision-making. Every day people make decisions based on information they receive from social networks, news, and friends. Nowadays intelligent machines are starting to take the space of advisers informing or (not) informing us to make better or worse decisions. Intelligent advisers are known by various names such as virtual personal assistants, intelligent assistants, or cognitive advisers. They are present in several areas, such as health, well-being, finance, commerce, and education to mention some. In certain contexts, such as finance, the challenge to design for those systems is even bigger since private information is shared by humans with machines and trust is essential in those contexts.

Many technological advances in the early 2010s in natural language processing (spearheaded by the *IBM Watson*'s victory in *Jeopardy*) spurred the availability in the early 2010s of text-based chatbots in websites and apps (notably in China [1]) and

A. Marcus and W. Wang (Eds.): DUXU 2017, Part III, LNCS 10290, pp. 594–613, 2017.
DOI: 10.1007/978-3-319-58640-3_43

spoken speech interfaces such as *Siri* and *Cortana*. However, the absolute majority of those chatbot deployments were in contexts of *dyadic dialog*, that is, a conversation between a single conversational agent with a single user.

Human-computer interaction, in practice, has also been mostly about dyadic interaction since the dawn of computer systems in 1950s. The two dominant interaction paradigms, *command-line* and *point-and-click*, are both not well suited for multi-user interaction (one application with more than one user engaged in the same activity) or multi-app interaction (one user interacting seamlessly with more than one application), and even more for generic multiparty applications (many users and many bots simultaneously). Notably exceptions are surface interaction and multi-user games but the mainstream of human interaction with computers remains one-to-one.

In this sense, conversational interfaces powered by chatbots are an important breakthrough from the past of computer interaction because they naturally enable multiparty applications. By exploiting the many social protocols human had developed for multi-person conversations since the advent of language, conversation-based interfaces may finally break from the dyadic paradigm in computer interaction.

In this paper, we consider a scenario of wealth management where advice is provided by multiple chatbots. We highlight our design process and several design activities undertaken with potential users with limited financial knowledge to understand their rational when making investment decisions. The information gathered from user studies assisted us to define and refine the concept and graphical user interface of a financial adviser. We then explore opportunities to use this knowledge in the context of a multiparty dialogue system called *CognIA* (Cognitive Investment Adviser) which is a chat system aiming to help users with low knowledge of finances to take more informed decisions about their investments.

2 The Design Process and the Role of User Studies

The specific methods and tools used in this research were detailed and organized according to the main stages of the *Design Research* methodology [2, 3]. An overview of the design process is depicted in Fig. 1. As we see, the beginning of the design process was guided by user experience studies supported by theories, competitors' analysis of financial advisers, and technology availability and its constraints. We explain in detail the different methodologies, processes, and findings in the remainder of the paper.

Notice that at the beginning of the process there were no pre-selected form for the finance advising system: the decision to use a conversational system was a result of the overall design process. Although many aspects and activities were performed, the key components of the design process were four user studies we conducted to explore the needs of our users and, later, the nuances of human-machine interaction which informed the design of our conversational system *CognIA*. The aim of the user studies was threefold: understanding the everyday practices which prevent people to make better investment decisions; exploring how a system could be to help people to make more conscious investment decisions; and designing the basic information for the future system to work (the language/knowledge corpus used in the conversational system).

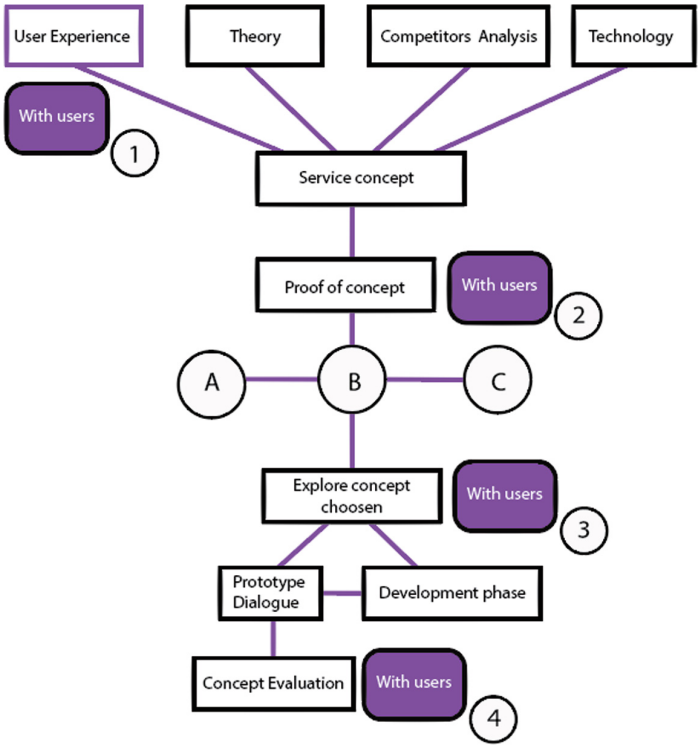

Fig. 1. Design process of CognIA

CognIA is targeted to investors who have limited knowledge of finance (and often not willingly to spend time learning it) but nevertheless would like to make good investments. We focused on people younger than 40 years old, well-educated, and with medium-high income. As reported by previous research [4] most Brazilians in this segment (about 70%) save money in basic, low-yield savings accounts, instead of investing in other financial products available.

Some of those products have returns much better than savings accounts, low risk, and reasonable liquidity, so many investors fail to make money simply due to lack of information. We recruited participants with those characteristics for our qualitative user studies, marked in Fig. 1 as "With users". The first user study aimed to understand financial practices, motivation stoppers to invest, and interaction channels to make investments (circle "1" in Fig. 1). The second user study focused on testing user preference concepts (circle "2" in Fig. 1). Users evaluated three service concepts envisioned by the project team. A concept based on an intelligent financial adviser was chosen by the design team based on the results of the second study. But further explorations were needed to design an intelligent financial adviser which led to the third user study (circle "3" in Fig. 1). Questions such as: How people would add data into the system? How people would interact with the system and receive feedback? Which kind of interaction modes are more suitable for our target audience?

Results of those preliminary activities guided the designers of the project team to design a graphical user interface for *CognIA*, envisioned as a multiparty dialogue system. An animated video demo which demonstrated how the application would work was created to guide our multi-disciplinary team to develop the UI. The video was a reference for discussions on interaction flow and visual design decisions in the team meetings. In parallel to system development, designers evaluated the first impressions of the system by potential users (circle "4" in Fig. 1). A video-card perception technique was employed to gather first impressions from users. This technique is a mix of using our video demo as scenario and the reaction card method.

In the next sessions, we explain with more details of our design research stages and how design activities were conducted and assisted us on project decisions and directions. But before going into the details of the design process, it is important to understand that our design challenges are bedded in the general context of how people make economical and finance decisions, an area often known by the term *Behavior Economics*.

3 Behavior Economics and Finance

For this project, we were informed by social science theories which could help us understand financial behavior and investment decision-making. In particular the work of [5] known as *Prospect Theory* was extremely helpful to us. Prospect Theory is based on four key elements: reference dependence, loss aversion, diminishing sensitivity, and probability weighting. Briefly, *reference dependence* is based on observations that people derive utility (value) from gains or losses relative to a specific reference point. Thus, individuals assess the outcome of a financial decision relative to a current value (e.g., current value of an investment). *Loss aversion* describes the asymmetry in sensitivity which people experience between loses and gains. Individuals are significantly more sensitive to a loss of a specific amount than to a gain of the same amount. *Diminishing sensitivity* and *probability weighting* describe how people behave unreasonably at the extremes of the value distribution.

There have been several attempts to exploit some of these principles in financial systems and interfaces. For example, [6] attempted, with mixed results, to overcome loss aversion by presenting more explicitly and detailed risk assessments and expected returns for financial investment decisions. [7] explored the influence of both few financial thought leaders and aggregated crowd choices on the investment decisions of older adults. They found that individuals who took aggregated crowd choices into account were able to make less risky investment decisions. [8] created an investment interface that highlighted potential losses (and gains) and motivated changes to investment allocations to minimize future losses. They effectively demonstrated that by appealing to loss aversion tendencies they could motivate investment changes.

For our purposes, we considered how to reasonably *frame* the investment decision and outcome so the user would have a frame of reference for the expected (or actual) gain or loss. We also considered design alternatives for presenting investment losses and gains to remedy the decision bias associated with loss aversion.

4 First User Study: Understanding Financial Decisions

Having explored the general decision framework of behavior economics, we start the description of our design process by describing the first user study which aimed to understand people's perceptions and their attitudes towards investment decisions and how they manifest themselves in practice. Additionally, the outcomes of this study oriented team discussions by considering real issues and situations reported by participants. The study consisted of a set of semi-structured interviews which were carried out to understand everyday practices, motivation stoppers, and interaction channels to make investments.

Participants. Twelve participants aged from 29 to 43 years old were interviewed. All the participants were Brazilians. Five were men and seven were women, and four participants had children. All the participants had a university degree and premium bank accounts in Brazil (clients with medium, high income).

Method. Twelve semi-structured interviews were conducted in May of 2015 taking approximately 20–30 min each. Participants were initially recruited by a snowball sample. A consent form was filled out by participants before the study started. The semi-structured interviews covered open-ended questions about the participants' past investments decisions, factors they consider when making investment choices, and sources they consult before investing. The interview also included a practical question about investments: *"If you had R$ 20.000,00 to invest, which type of investment you would do and why?"* All the sessions were audio-recorded and researchers took notes during the interviews.

Data Analysis. An exploratory and qualitative approach was undertaken to conduct the data analysis. The data was analyzed having three research questions as guidance: (1) What are the most common investments of participants? (2) What are their challenges when deciding where to invest? (3) Which kind of people and/or information sources they consult to know more about investments? The data was coded after semi-structured interviews and audio transcriptions. Categories emerged from the data and relationships emerged between the categories. This process was facilitated using *NVivo* software.

Findings. The most common investment types of our participants are savings accounts and fixed income investments (for example, treasury bonds known as *CDB* and *DI*). Ten participants considered themselves "conservative" investors, and not having enough knowledge of investing. Two participants, the ones who considered themselves as "moderate" investors, have used professional financial advisers in the past and had already taken elementary financial education. Two participants had lost money in the stock market previously and admitted they did not have enough knowledge and confidence to do this kind of investment. Most of our participants seem to overweight risk over returns. The ones who invested in fixed income investments followed advice from their bank managers or family and friends. Some of the challenges our participants faced when deciding to invest are described below with quotes from our participants:

Lack of Knowledge – Financial education is not a priority for our participants. Also, families do not pass information of how to invest to other family members, since money a private issue of which people do not talk about frequently. Bank managers are the main source of information but they are regarded by our participants as vendors. Participants sometimes rely on their advices in case of fixed income investment, or leave their money in less profitable and popular investments such as savings accounts. Bank managers also use terms and jargons which some of our participants are not familiar with ("risk protection", "liabilities", "liquidity"). In our participants' opinions, bank managers usually have a restrict portfolio of investments to offer and there is no *"transparency in relation to taxes and return" (P4).* They also mention that good bank managers do not show only the good investment options but also talk about the ones which are less profitable.

Lack of Time and Interest – Our participants do not feel attracted by financial information, most of them find boring to look for information in this area or to read it. Moreover, most of them think investing will be time consuming – *"I have friends who are slaves of their investments" (P6).* Occasionally, participants prefer to follow an advice of which they are not aware of all the constraints due lack of time to monitor investments, interest, and/or motivation. Financial information sounds complicated and difficult to them *"It is very difficult to know how do I make money investing." (P3).*

Emergency Funds and Liquidity – All the participants mentioned the necessity to have an emergency fund in case they need of cash. Single participants usually save money for travelling and leisure. Married participants with children usually save money for a future home purchase, eventualities, and family holidays.

Communication Channels – Bank managers usually contact our participants by phone, often in commercial hours. This is considered disrupting and annoying by our participants since talking about an investment portfolio is a private matter not to be discussed at work and also because they are not available to talk at the moment of the call. They prefer to receive e-mails than calls from the banks.

The most common information source participants consult before making an investment are family and friends with experience in investing. Trust was a big issue for our participants, and the lack of trust on bank managers sometimes results on choosing not so profitable investment choices. Four participants also rely on news and financial websites to validate options suggested by bank managers or family and friends. Even though, those four participants focus their reading on investments they were not confident to better compare with the other investments.

5 Designing the Service Concept

Having determined the key financial needs, issues, and worries of people in the target segment and understood deeper the rich mental and behavior processes involved in financial decision making, we proceed with a design phase whose aim was to determine the best approaches to serve the needs of this population. It is important to notice that we are not simply interested in designing a computer system which could support

investors. Our approach was to look to the problem holistically and target the designing of a *service* to be provided by a financial institutional or similar, including how it would be created and maintained, its business model, the flow of information, and its support systems.

In other words, it was a problem of *service design*, albeit with a strong IT component. After looking into different methodologies used in service design as described [9–12], we decided to use the notion of *service concept* as the central structuring element of this design phase. Edvardsson describes the service concept as *"a detailed description of the customer needs to be satisfied, how they are to be satisfied, what is to be done for the customer, and how this is to be achieved."* [13].

Based on this framework we researched and explored different innovative ideas and models which were being proposed in the finance landscape around 2015, as well as performed some ideation sessions, aiming to collect different service concepts (existing, planned, or futuristic) which could be a solution to our design goal. The result was a list of 10 service concepts which included services based on cognitive advising systems, social wisdom platforms, finance learning systems, etc. As a way to visualize and better compare the service concepts, we mapped them according to the source of the financial knowledge employed by the system, which could be the user, experts, wisdom-of-the-crowd, or from a machine. Figure 2 shows this mapping and the relative position of the 10 service concepts.

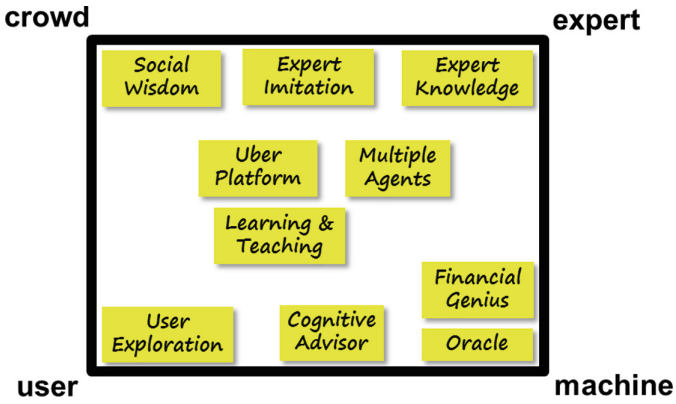

Fig. 2. The 10 service concepts mapped according to the source of finance knowledge used by the system.

To compare and select the most promising candidates for the next phase of the design process, we analyzed each service concept according to four dimensions: quality of *user experience*, matching to economic and finance *user behavior models*, level of maturity of the *supporting technology* required, and the number of players in the *competitive landscape*. This analysis was performed by four different individuals in the group who were experts in the corresponding area.

User Exploration

	Good	Bad	Ugly
User Experience	allows trial & error allows learning integration of information	interactive financial simulator can overwhelm people too much work for users	requires financial literacy personalization to literacy level lack of expert to trust
User Behavior Models		hard to overcome cognitive bias	
Technology	Watson Explorer APIs journalistic data published	financial data search engine financial reasoning models temporal reasoning	
Competitive Landscape	not really something as nicely aggregated out there	lots of similar stuff which people do not use	

Fig. 3. Table showing the result of the *Good/Bad/Ugly analysis* for the service concept *User Exploration*.

To collect the analysis in a comparable format, we employed a technique we call the *Good/Bad/Ugly* analysis[1]. For each service concept and dimension we listed its positive aspects (*The Good*), negative aspects (*The Bad*), and possible road blocks (*The Ugly*). Figure 3 shows the resulting table for a service concept where the user can explore freely financial terms, concepts, and products. Using the Good/Bad/Ugly tables created for each service concept, the four experts discussed and compared the different concepts and selected three of them as the most promising:

- **User Exploration:** a finance information aggregation system where all kinds of financial information (markets, social media, simulators) could be searched and explored freely by the user.
- **Expert Knowledge:** a human advisers-based system where the user could watch videos and read opinions from experts and celebrity investors, combined with some level of automatic personality and profile matching.
- **Cognitive Adviser:** an intelligent system able to match investment needs to products with support for user exploration through question-and-answer and provision of evidences and opinions to support financial decisions.

Having found three service concepts with enough potential and feasibility, it became clear that to go further in the design process we would need to have user feedback not on the service concepts but on actual realizations of them.

6 Second User Study: Exploring Service Concept Mockups

We then quickly proceeded to develop basic *mockups* of the service concepts, wireframe prototypes with sample interactions for each of the three service concepts selected in the previous design phase. Figure 4 shows sample screens from each of them.

[1] Inspired by 1966 Sergio Leone's classic western *The Good, The Bad, and The Ugly*.

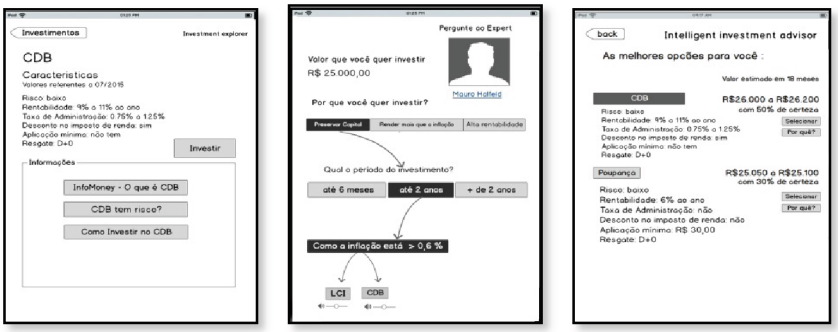

Fig. 4. Sample screens of the mockups built for the three service concepts: user exploration (left), expert knowledge (middle), and cognitive adviser (right).

To perform a user evaluation of the three mockups, a qualitative, inductive and interpretative approach was taken in which we employed the foundational theoretical ideas about how people make investment decisions and access financial information sources emerged from preliminary user studies and the previous design research activities. In other words, as well as delivering requirements for such a system, these activities generated issues about the different design parameters – interaction mode and flow, layout, navigation system, activity type, information design and many more – for which recommendations were generated. The main goal of the second design study was to gather user's impressions of three different mockups and to identify which aspects people expect from financial advisers supported by technology.

Participants. Fifteen participants aged from 26 to 40 years old participated in the study. All the participants were Brazilians. Seven were men and eight were women, and only one participant had children. All the participants have a university degree and premium bank accounts in Brazil (clients with medium, high-income). Participants were from diverse backgrounds and most of them had savings accounts. According to them savings accounts have a high liquidity and for this reason was the preferred investment choice. The majority preferred to use Internet and mobile banking and rarely would go to a bank branch. The ones who preferred to go to the branch bank would do so to withdraw money or to make financial transactions. Those considered going to the bank safer than doing online transactions.

Methods. In the lab sessions participants interacted with the three mockups guided by a scenario and answered questions with answers ranked by a *Likert scale*. The study took approximately 45 min for each participant. The moderator made clear that researchers were interested to know about the concepts and not keen on evaluating graphic design features. First, participants answered demographic questions and commented their previous investments experiences. Participants read aloud the scenario and then started the interaction with the mockups:

Scenario: *Consider you have an income of R$10.000,00 and you just received R $25.000,00 as bonus. You want to buy a new house (or car) to replace your older one in about 2 years. Your plan is not to use this money before 2 years. You need some help*

deciding which kind of investment is the best for you. Therefore, you are testing some investment apps.

Each interactive mockup illustrated one service concept and it was used as a *prop* to foster discussions with participants. In the first mockup users explored a tag cloud containing investment names with links to financial websites. In the second, an intelligent assistant based on the user's profile provided best investment choices with levels of confidence. The third was an interactive knowledge map which users could explore by interacting with topics and listening to audio samples from a celebrity human financial adviser. Participants were exposed to the three mockups counterbalanced. Participants were asked to think aloud and report their choice of investment after experiencing each mockup. For each mockup interaction, participants were asked to describe how they thought that mockup works and their expectation for the service offered by it. Additionally, they answered a Likert 5-point scale regarding 7 statements. They also were encouraged to give their rational choices while filling out the Likert scale. At the end of the study, participants were encouraged to compare the concepts and report their preferences.

Data Analysis. A qualitative approach was undertaken to conduct the data analysis. The data was analyzed with the aim to elucidate user preferences of design concepts and key facts that might affect financial decisions. The Likert scale was analyzed and served as a source to understand participants rational and concept choices. The data was coded after the user sessions and audio transcriptions. Categories emerged from the data and relationships emerged between the categories. This process was facilitated using *NVivo* software.

Findings. The main categories affecting financial decisions were: family opinions, lack of trust in bank managers and government, debts, and having or not children. Three of our participants were planning to have children, therefore their opinions tended to be more conservative. Other factors also affected how people share their service concept preferences: lack of experience with investments, previous experiences with investments, knowing the source of information, and expectation to validate information offered by our service concepts. Regarding each mockup, the following summarizes the findings:

User Exploration – Participants with more motivation and experience with previous investments preferred this option. Others with low knowledge of investments evaluated this option overwhelming, due to information overload, and likely time consumed. Participants also found this option an incentive to learn about investments while accessing the financial webpages links. This option also inspired credibility because of the variety of information sources.

Expert Knowledge – People not keen on investments do not like to spend time reading and investigating about investments. In our participant words *"It's straightforward, does not make me go around and search like the other one (UE)"*. Also, most of the participants considered this option narrow and biased in one expert opinion. All the participants said that they would validate investment options suggested in this concept with friends and family. The ones who know the expert valued this option with higher

degree. It was also evident the importance of having up-to-date information in this option more than on the others.

Cognitive Adviser – Participants with no previous experience with investments, except savings accounts, preferred this option. It was considered more straightforward than other options. They appreciated building their profile in collaboration with the system by choosing pictures which would match their profiles. For some participants, it is not transparent knowing the reasons why they are categorized by the banks as conservative, moderate, or aggressive investors. Although, this option was the preferred of our target audience, it was also considered less trustful than other options because it lacked data source transparency.

In our participant's views, all the service concepts should have detailed information of output values, including tax deductions. They also wanted more types of investments to help comparison. Moreover, information sources should be up-to-date and this should be more explicit in the service concept interfaces (e.g., showing the update date). Information curation and ownership was also highlighted as paramount to have knowledge of and as an issue to affect investment decisions. Overall, people with low financial knowledge preferred the intelligent assistant concept while people with some financial knowledge preferred exploring web sources. Most of our participants found narrow and not appealing the expert knowledge option, they argued the need for more than one adviser perspective. Most of participants preferred not having this kind of service being directly offered by banks due to information bias. Based on the results of this user study the project team decided that the most promising direction was the Cognitive Adviser, augmented by some ideas from the other service concepts which have received high praise from the participants of the study.

7 *CognIA*: A Cognitive Investment Adviser

Having converged the design process towards a conversational advising system, we start exploring the advantages and challenges of the approach. A key concern of the design team was how to build a system which inspired trust from its users. We had identified in the user studies the propensity of participants to see the advice of their bank managers and financial advisers as biased by a selling proposition. Concerned that the same problem would affect the perception of the advice provided by a machine, and inspired by some appearing automatic debating systems, we decided to take an approach where the user would converse not with a single chatbot but with multiple advisers, each advocating for a particular financial product, in a dialogue moderated by a trusted, non-partisan financial adviser.

Following this concept, we developed *CognIA* which is a multi-bot conversational system which helps people to make better investment decisions. In the current version three bots participate in the same dialogue with the user. *Cognia* is also the name of the agent which moderates the conversation; *SavingsGuru* is an agent which answers questions about savings accounts; and *CDBGuru* is an agent which answers questions related to CDB investments (a kind of treasury bond). A multi-bot platform, called

Sabia, is used to define the entities, relationships, and behaviors needed for the creation of coordinated chatbots which react or pro-actively act using natural dialogue.

The interaction starts with the *Cognia* agent which asks basic questions about the user needs. Then the *Cognia* agent can invite one or more chatbots in the chat into the conversation considering appropriate investments for the user. *Cognia* is able to redirect the topics based on the user's utterances and to enforce that the chatbots only send allowed messages. The *CognIA* visual/interaction design process has been informed by the design activities described in the last section. The questions embodied in the system and their intents were gathered from the results of the *Wizard of Oz* study described in the next section. The answers were composed based on financial websites and financial experts' posts. Then a concept video was produced to guide the design and deployment of *CognIA*.

8 Third User Study: Discovering Conversation Patterns

This experiment was designed to examine user dialogues mediated by a cognitive investment advisor using the *Wizard of Oz* technique [14, 15]. Participants believed they were interacting with a functional system. Fifteen user sessions were conducted in October 2015. Participants were invited to test a "the first version" of an intelligent financial adviser dialogue that could answer questions related to two kinds of investments: Savings accounts and a fixed-income investment called CDB (Bank Deposit Certificate). Participants followed a Wizard of OZ protocol. The user tests were remote and took approximately 30–40 min. The main data gathered were notes, audio and video recordings (screen captures). Participants were young adults (26 to 43 years old), highly educated and high-income bracket. All the participants described themselves as not interested or not keen on finances, particularly investments. All the participants answered positively to the consent form document, allowing us to use the data gathered.

Procedure. Participants were recruited by a snowball sampling [16] and invited to be part of the remote study. The sessions started with demographic questions and questions of their financial investment experiences. Following that, they shared their screen with the researcher and started interacting with the chat mock up, the supposed Intelligent Financial adviser. A human operator, that was not the same as the researcher facilitating the user session, answered their questions using a protocol. The human operator used a small table of content to answer the questions. The table was composed of 36 small paragraphs extracted from popular financial websites. The content relied on investment definitions, pros (return) and cons (risk) of two types of investments. Every table cell had a label (e.g. interest, safety, minimum value) to help the operator find the questions quickly during the sessions. The human operator could use sample answers in case she did not have an answer (1.I don't know; 2. Ask again please; 3. I don't have enough information). In the end of the session, the facilitator asked the participants to give their impressions about the system and disclosure the identity of the intelligent system.

Data Gathering and Analysis. Lightweight and heavyweight analyses were the approaches to analyze the data [17]. The lightweight analysis consisted in an affinity cluster extracted from notes and offered guidelines for the main categories to look for in the audio transcriptions. In the heavyweight analysis, the Nivivo software was used to analyze the data. Notes, Chat transcriptions and videos were analyzed. Categories from the affinity cluster phase were used as a base to analyze Chat transcriptions and video transcriptions. Videos were mainly a source for understanding why people wrote some questions for the Financial advisor. For example, sometimes people repeated a question, or rephrase a question before write the question and not always they typed what they wanted to know. Some reactions and contextual information were only possible to gather from watching some sessions again.

Findings. The main results were a categorization of questions, a set of questions of each investment (Savings and CDBs) and 18 design recommendations. Overall, 125 questions were gathered: 86 questions about CDB investments and 39 questions about Savings. The main categories of information for both investments were: definitions; advantages and disadvantages to invest; simulation of values and profitability; liquidity, fees; and risk (Table 1). Questions were asked in an informal language, although participants usually reframed their question thinking aloud before writing. Some participants were not sure if the system would understand them, for example: punctuation marks. Others expected the system would know Acronyms and main jargon from investments. In occasions, they received a negative answer from the system they were usually upset but forgiving. They also repeated variations of the same question when they did not get a satisfactory answer from the system (human operator). Additionally, results provided insights and recommendations for designing our prototypes of chat-based cognitive investment advisors. It also provided the first corpus for *CognIA*.

Table 1. Examples of categories and questions extracted from the study

Categories	Sample question	
Definition	What are the types of CDB?	Quais são as modalidades de investimento CDB?
Advantages and disadvantages	What are the advantages to invest on CDB?	Quais são os pontos contra para investir em cdb?

9 Visual Design and Graphical User Interface

The graphical interface design paradigm was created based on the familiarity of our participants with conversational apps such as *WhatsApp* and *Skype* and behavioral theories. The visual identity was designed to reach the audience of our project. A mood board, a semantic panel with graphical elements (colors, typefaces, shapes) with perceptual aspects gathered from the user studies, were created to guide the visual design. Graphical elements were added to the chat interface to create the sense of multiparty dialogue, such as: icons of the agents, location of icons on the screen, distinct colors for

each agent. Visual comparison and calculation features were considered essential features by design activities participants and were added to the system. Several interaction modes were included into GUI inspired by behavioral theories.

Mood boards are collections of images to represent concepts, sensations in response to a product or service. It is an instrument for designers to communicate with each other and with the clients and the project team [18, 19]. Designers working as part of the project team chose five words to serve as the base for the visual language of *CognIA*: finance; trust; motivation; technology and intelligence. Those words were inspired by user study findings. Those words referred to characteristics people value when making decisions about investments and characteristics the project team wanted to transmit with *CognIA*. Based on those words, designers searched for images using search websites (*Google images*, *Instagram*, *Pinterest*). The images were selected and grouped by the chosen words, Afterwards, common colors to compose the color scheme of *CognIA* were extracted from each group of images (see Fig. 5).

Fig. 5. Mood board – pictures related with concept areas (Color figure online)

Based on the observation of shapes and colors found in the mood board, designers created a *semantic panel* with three principal columns: *concept*, *perceptual aspects* and *results* (see Fig. 6). The concept column included rows for finance, intelligence, motivation, trust, and technology (Fig. 6 shows only the *finance* row). The *perceptual*

Concept	Perceptual Aspects		Results		
			Typography	Shape	Color
Finance Choose the best investment option.	Economy Control Investment Rate	Rational Graphics Measurement	Helvetica Neue Regular Sans Serif	Arrow Lines Right angles Scale	Yellow Blue Dark Blue Dark Green

Fig. 6. Part of semantic panel shows the *Finance* row (Color figure online)

aspects column contains terms which describe the characteristics which can be observed in the mood board, economy, investment, rational, graphics, etc. The *results* column was divided into three visual elements subgroups: typography, shape, and color. The *typography* column contains choices of the best typefaces for the project.

The *shape* column shows choices of shapes which fit to the concept. The *color* column shows colors options related to the concept which fit the guidelines of the sponsoring organization. The semantic panel was used as a resource for designers to create the visual identity and graphical user interface of *CognIA*.

Designers then explored several versions for the graphical user interface applying colors, shapes, and typefaces available in the semantic panel described above (see Fig. 6). One version was chosen in a shade of green, since this version was visually distinct from current financial apps in the marketplace (Fig. 7).

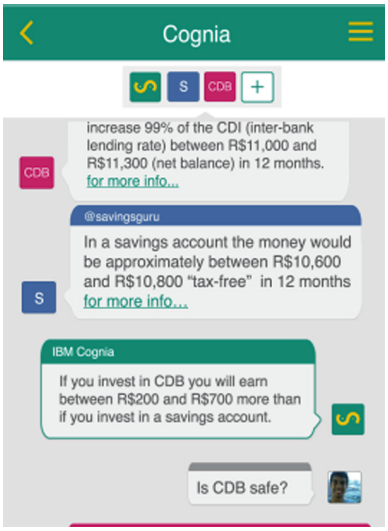

Fig. 7. Cognia's conversational user interface (Color figure online)

A user scenario was developed to illustrate the interaction flow of the system. This scenario was used to create the basic interaction flow of the systems and the expected behaviors of each chatbot. Since the project team is multidisciplinary, composed of computer scientists, designers and engineers we used a video demo as tool to guide the team through the development process of *CognIA*. The team would have an example of how the envisioned concept should work and based on it the functions illustrated in the video were developed. The video case scenario was inspired by previous design activities to understand investment decisions with real users. It was made by the designer members of the team. The video length was about three minutes and showed how the multiparty cognitive dialogue helped a user to make an investment decision.

The first part of the video consisted in a short introduction with an initial phrase *"Cognia has detected that you have considerable money in your bank account"* to

introduce the case scenario. Next, a user receives a push notification from the system with a message *"Your account in Banco Blue has R$ 16.000 that is not receiving any interest. Would you like to invest some of this money?"* This message acts like a trigger for opening the application. After that, *Cognia*, the moderator agent, displays three options with different values to be invested. The user selects one of those options and *Cognia* replies: *"There may be a penalty for early withdrawal for your investment. How long do you plan to leave the money in this investment?"*and shows three options of investment horizon. The user chooses one and *Cognia* invites two more expert agents to participate in the dialogue, one is *SavingsGuru* and the other is *CDBGuru*. *CDBGuru* provides a simulation of the value selected by the user and after *SavingsGuru* does the same. The user asks a question a clarifying question, *"What is CDB?"* and *CDBGuru* answers it. *Cognia* then compares the answers from the product gurus showing the user which investment is more profitable. *CDBGuru* and *SavingsGuru* show then some web articles with links which can be useful for the user. An interactive visualization comparing values is also shown for the user in the video. *Cognia* then asks for a decision and the user chooses the CDB option. *Cognia* redirect the user to a fictitious *Bank Blue* page where he can complete the transaction.

10 Fourth User Study: Evaluating the *CognIA* Prototype

As part of the design and development process we performed an evaluation study to understand the desirability and issues of the prototype of our multi-bot chat advising system, *CognIA*.

Participants. Ten Brazilian participants were recruited by a snow-ball sample. The background of our participants was varied (Linguistic, Design, Computer science, Anthropology, Computer science, Tourism, and Engineering). All the participants worked in the same technology company. Five participants were female and five were male, with an average age of 35 years. All of them had previous experiences with savings account and four of them with other types of investments (including CDB). It was the first-time participants have contact with a multiparty dialogue system. All the participants we recruited were familiar with chatrooms.

Method. In the first part of the study, participants were asked to watch a video demo and make any comments they judged necessary (see Figs. 8, 9, and 10). The instructor asked them first to select 10 adjectives of 118 *product reaction cards adjectives* [20] written in individual paper cards and to think aloud while making their choices [21, 22]. Afterwards, participants were asked to choose 5 adjectives among the 10 adjectives previously chosen and to give reasons for the choice to the instructor. Participants were also requested to add any words they did not find in the stack of 118 cards and to explain why. Following the reaction cards activity, participants were asked to watch again a specific fraction of the video in which *CognIA* invites other agents to participate in the conversation. Then, the instructor questioned the participant: *"Who are the participants of this conversation? What is your impression of this concept?"* Participants then shared their thoughts of multiparty dialogues with the instructor.

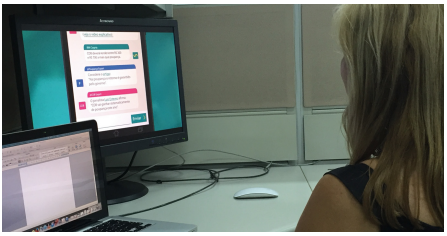

Fig. 8. Participant watches the video.

Fig. 9. Participant choosing cards.

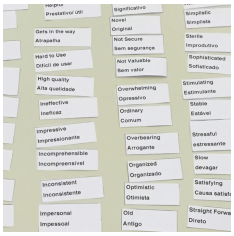

Fig. 10. Cards used in the study.

At the end of the session participants answered a semi-structured interview with demographic questions and their previous experience with investments. A consent form was filled out by participants before the study. The experiment was conducted in a lab and in some cases remotely via video conference. For the remote participants, the paper cards were substituted by a table with the adjectives. Remote participants were asked to highlight 10 words in red and the 5 words in bold. The length of the session was on average 15–25 min. All the sessions were audio- and video-recorded. The observation data was analyzed with the support of the notes which researchers took during the sessions. The use of a notepad was vital to gather information in case any problems might happen with the video recording. Participants were rewarded with a small gift.

Data Analysis. The data was analyzed using descriptive statistical methods and qualitative methods. Basic statistical analysis was carried out to analyze the data from the questionnaires: demographic data, semantic scales, and design preferences. Tables and cross tabulation were applied to compare the results among participants and the use of the system. The restricted number of participants in the study was not enough to ensure the validity of the statistical analysis. Research questions were kept in mind while the data was classified and codified. Relevant issues were classified into sets of codes. The principal set of issues emerged from the data were: reactions of multi-agent concept and reactions to interface and information design (positive and negative adjectives). The transcriptions of the videos, observation analysis and coding were assisted by a qualitative software.

Findings. Participants chose 35 words from 118 reaction cards. Thirty-two words selected were positive and three negative (Fig. 11). The positive words more frequently chosen by participants were: *Easy to use, Sophisticated, Friendly, Straightforward, Helpful,* and *Connected*. The three negative words chosen by participants were: *Scary, Insecure*, and *Intimidating*.

The interface was considered familiar and easy to use, and participants identified similar tools they already use similar to *CognIA* concept. The dialogue styles were recognized as useful since participants perceive this mode as allowing them to type and ask what they desire. Overall, participants found useful the separation of investments in agents. Some of them considered this approach useful for information design since

Fig. 11. Tag cloud results

each agent has a different color which helps in organizing and distinguishing information in the chatroom screen. Others did not notice the agents' separation. *"Funny, I found interesting this approach, but for me I was talking to CognIA all the time. It was what I felt. Even though we had characters (CDB and Savings) for me the interface is the app, the system."* Some participants emphasized the importance of the quality of the information and not the way it was delivered (single bots or multi-bots).

Additionally, participants highlighted that having different agents for different investments helped them to compare investments and decide which investment would be more suitable for them. A better job could have been done to distinguish the differences between agents in the system. Even though they liked the separation it was not possible to be sure they understood the system was multi-agent. It was also clear the option of having multi-agents in the system did not affect negatively the participants' perceptions. For more details of this study please see [23].

11 Conclusion and Further Research

We described in this paper a design process of a financial adviser system which led to the development of a prototype of a multi-bot conversational system. The design process went through multiple steps and used four different user studies to guide key design choices and to validate the multi-bot approach. Multiple design and user study methodologies were combined to address the complexity of the challenge and needs of the multi-disciplinary design team. The resulting prototype of the CognIA system was highly evaluated by potential users. At the moment of the writing of this paper the prototype is being refined and improved to be deployed in a financial information website in Brazil.

In today's point-and-click interfaces, it is very hard to two or more people to explore together the options in an investment website and even harder in a smartphone app. In most cases, mouse control is appropriated by one of the users, creating a natural dominance which is not inductive to collective decision-making. This can be addressed in our proposed system, a multiparty chat where investment chatbots and humans can

talk to each other in the same dialogue. In the next version of the system, it will be possible to bring to the conversation their spouses, family, or other people they trust such as a human investment adviser. Similarly, multiparty chats naturally allow competitive behavior and its effective management. In the investment scenario, the users could bring banks with competitive products to a single conversation, allowing easier comparison and even auction-like competition between the bots.

Challenges to design multi-bot systems are to understand and implement humans' protocols, managing and monitoring turn-taking in a dialogue [24], track of threads and topic changes, and to design to support multiple roles (for example bank managers, family) [25]. Those are key technical challenges to be overcome to effectively deploy multiparty conversational systems.

References

1. Olson, P.: Get Ready for the Chat Bot Revolution: They're Simple, Cheap and About to be Everywhere. Forbes (2016). http://www.forbes.com/sites/parmyolson/2016/02/23/chat-bots-facebook-telegram-wechat/. Accessed 19 Aug 2016
2. Nieveen, N., Mckenney, S., Van Den Akker, J.: Educational design research: the value of variety. In: Van Den Akker, J., et al. (eds.) Educational Design Research: The Design, Development and Evaluation of Programs, Processes and Products. Paperback, New York (2006)
3. Zimmerman, J., Forlizzi, J., Evenson, S.: Research through design as a method for interaction design research in HCI. ACM (2007)
4. SPC (2015). https://www.spcbrasil.org.br/imprensa/pesquisas
5. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica: J. Econom. Soc. 263–291 (1979)
6. Zhang, Y., Bellamy, R.K.E., Kellogg, W.A.: Designing information for remediating cognitive biases in decision-making. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM (2015)
7. Zhao, J.C.: To risk or not to risk? Improving financial risk taking of older adults by online social information. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. ACM (2015)
8. Gunaratne, J., Nov, O.: Informing and improving retirement saving performance using behavioral economics theory-driven user interfaces. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM (2015)
9. Chase, R.B.: It's time to get to first principles in service design. Manag. Serv. Qual. **14**(2/3), 126 (2004)
10. Holmlid, S.: Service design methods and UCD practice. In Proceedings of the INTERACT 2005 Workshop on User Involvement in e-Government Development Projects, Rome, Italy (2005)
11. Mager, B., Gais, M.: Service Design. Fink, Paderborn (2009)
12. Curedale, R.: Service Design: 250 Essential Methods. DCC Press, Topanga (2013)
13. Goldstein, S.M., Johnston, R., Duffy, J., Rao, J.: The service concept: the missing link in service design research? J. Oper. Manag. **20**(2), 121–134 (2002). doi:10.1016/S0272-6963(01)00090-0

14. Steinfeld, A., Odest, C.J., Scassellati, B.: The Oz of wizard: simulating the human for interaction research. In Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, HRI 2009, pp. 101–108. ACM, New York (2009). doi:10.1145/1514095.1514115
15. Grill, T., Tscheligi, M.: The ConWIZ protocol: a generic protocol for wizard of Oz simulations. In: Moreno-Díaz, R., Pichler, F., Quesada-Arencibia, A. (eds.) EUROCAST 2013. LNCS, vol. 8112, pp. 434–441. Springer, Heidelberg (2013). doi:10.1007/978-3-642-53862-9_55
16. Bryman, A.: Social Research Methods. Oxford University Press, Great Britain (2008)
17. Goodman, E., Kuniavsky, M., Moed, M.: Observing the User Experience: A Practitioner's Guide to User Research. Elsevier, Amsterdam (2012)
18. Chang, H.M., Díaz, M., Català, A., Chen, W., Rauterberg, M.: Mood boards as a universal tool for investigating emotional experience. In: Marcus, A. (ed.) DUXU 2014. LNCS, vol. 8520, pp. 220–231. Springer, Cham (2014). doi:10.1007/978-3-319-07638-6_22
19. Mcdonagh, D., Storer, I.: Mood boards as a design catalyst and resource: researching an under-researched area. Des. J. **7**(3), 16–31 (2004)
20. Benedek, J., Miner, T.: Measuring desirability: New methods for evaluating desirability in a usability lab setting. Proc. Usability Professionals Assoc. **2003**, 8–12 (2003)
21. Love, S.: Understanding Mobile Human-Computer Interaction. Elsevier, Oxford (2005)
22. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. Wiley, New York (2007)
23. Candello, H., Andrade, B.D.: Evaluating multi-agent conversational interfaces in the early stages of the design process. Revista de Design, Tecnologia e Sociedade **3**(1), 1–15 (2016)
24. Traum, D.: Issues in multiparty dialogues. In: Dignum, F. (ed.) ACL 2003. LNCS, vol. 2922, pp. 201–211. Springer, Heidelberg (2004). doi:10.1007/978-3-540-24608-4_12
25. Dignum, F.P.M., Vreeswijk, G.A.W.: Towards a testbed for multi-party dialogues. In: Dignum, F. (ed.) ACL 2003. LNCS, vol. 2922, pp. 212–230. Springer, Heidelberg (2004). doi:10.1007/978-3-540-24608-4_13