

UNIVERSIDADE REGIONAL DE BLUMENAU
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
CURSO DE SISTEMAS DE INFORMAÇÃO – BACHARELADO

CHATDUCA

ANDRÉ EDUARDO DA SILVA

BLUMENAU
2019

ANDRÉ EDUARDO DA SILVA

CHATDUCA

Trabalho de Conclusão de Curso apresentado ao curso de graduação em Sistemas de Informação do Centro de Ciências Exatas e Naturais da Universidade Regional de Blumenau como requisito parcial para a obtenção do grau de Bacharel em Sistemas de Informação.

Prof(a). Joyce Martins, Orientadora

**BLUMENAU
2019**

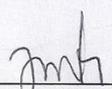
CHATDUCA

Por

ANDRÉ EDUARDO DA SILVA

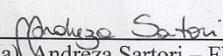
Trabalho de Conclusão de Curso aprovado
para obtenção dos créditos na disciplina de
Trabalho de Conclusão de Curso II pela banca
examinadora formada por:

Presidente:



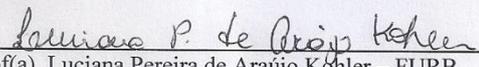
Prof(a). Joyce Martins – Orientador, FURB

Membro:



Prof(a) Andreza Sartori – FURB

Membro:



Prof(a). Luciana Pereira de Araújo Köhler – FURB

Blumenau, 09 de julho de 2019

Dedico este trabalho a todas as pessoas da minha família que sempre me deram todo o apoio necessário para a conclusão dele.

AGRADECIMENTOS

Agradeço aos meus pais, que sempre me apoiaram e incentivaram meus estudos, à minha namorada Manuela, por todo apoio e carinho nas horas mais difíceis, e à minha orientadora Joyce, que com sua paciência, competência e dedicação, me deu um norte desde o início deste trabalho.

No que diz respeito ao empenho, ao compromisso, ao esforço, à dedicação, não existe meio termo. Ou você faz uma coisa bem feita ou não faz.

Ayrton Senna

RESUMO

A pesquisa apresentada neste trabalho descreve o processo de especificação e de implementação de um *chatterbot* capaz de responder perguntas sobre notas, frequência e boletos de um software de gestão educacional. Para se comunicar com o *chatterbot* através de uma página web, o usuário deve ter um cadastro previamente já efetivado. O *chatterbot* mantém a comunicação com o usuário utilizando a língua portuguesa. As perguntas que o *chatterbot* consegue responder são especificadas em bases de conhecimento em AIML. O interpretador, responsável pela manipulação das bases de conhecimento, busca nas bases as respostas para as perguntas que são feitas. As bases de conhecimento foram criadas manualmente e, com isso, se fez necessário definir a estrutura, os assuntos, as perguntas e as respostas a serem tratadas. Uma dificuldade encontrada são erros de digitação, como troca de letras em palavras, os quais não são tratados pelo *chatterbot*. Apesar disso, os objetivos foram alcançados, uma vez que foi possível a interação do usuário com o *chatterbot* para sanar dúvidas sobre notas, frequência e boletos de um software de gestão educacional.

Palavras-chave: *Chatterbot*. AIML. Bases de conhecimento. Gestão Educacional.

ABSTRACT

The research presented in this paper describes the process of specifying and implementing a chatterbot capable of answering questions about the grade, frequency and ticket of an educational management software. In order to communicate with the chatterbot through a web page, the user must have an account. The Chatterbot maintains communication with the user by using the Portuguese language. The questions that chatterbot can respond are specified in AIML knowledge bases. The interpreter, responsible for manipulating the knowledge bases, searches the bases for the answers to the asked questions. The knowledge bases were created manually and, therefore, it was necessary to define the structure, the subjects, the questions and the answers to be treated. One difficulty encountered is typographical errors, such as the exchange of letters in words, which are not handled by the chatterbot. In spite of this, the aims were achieved, since it was possible to interact with the chatterbot to resolve doubts about notes, frequency and tickets of an educational management software.

Key-words: *Chatterbot*. AIML. Knowledge bases. Educational management

LISTA DE FIGURAS

Figura 1 – Inserção de texto no ChatterEDU	17
Figura 2 – Interação com o ChatterEDU: usuário faz pergunta ao robô	17
Figura 3 – Interação com o ChatterEDU: usuário responde uma pergunta de forma correta (a) de forma incorreta (b)	18
Figura 4 – Interação entre um usuário e o <i>chatterbot</i>	19
Figura 5 – Diálogo entre o <i>chatterbot</i> e o usuário utilizando <i>hiperlinks</i>	20
Figura 6 – Exemplo de tag <hiperlink>	20
Figura 7 – Resultado do uso da tag <hiperlink>	21
Figura 8 – Diagrama de casos de uso	23
Figura 9 – Diagrama de atividades	24
Figura 10 – Pergunta com e sem vínculo de tema	29
Figura 11 – Tela de login	29
Figura 12 – Tela de cadastro de usuário	30
Figura 13 – Saudação inicial para um estudante	30
Figura 14 – Saudação inicial para um responsável	31
Figura 15 – Pergunta sobre um assunto conhecido	31
Figura 16 – Pergunta sobre assuntos desconhecidos	32
Figura 17 – Lista de pendências do administrador	33
Figura 18 – Pergunta sobre novos assuntos	33
Figura 19 – Conversação do usuário com o ChatDuca	42

LISTA DE QUADROS

Quadro 1 – Classificação evolutiva dos <i>chatterbots</i>	14
Quadro 2 – Trecho de arquivo AIML.....	16
Quadro 3 – Assunto e resposta da base de conhecimento	26
Quadro 4 – Assunto e tema da base de conhecimento	26
Quadro 5 – Rotina que retorna assuntos conhecidos	27
Quadro 6 – Código para incluir assuntos ou temas nas bases de conhecimento	27
Quadro 7 – Comparativo entre os trabalhos correlatos e o ChatDuca.....	33
Quadro 8 – UC01: Cadastrar usuário	40
Quadro 9 – UC02: Fazer login	40
Quadro 10 – UC03: Dialogar com o ChatDuca.....	40
Quadro 11 – UC04: Manter as bases de conhecimento.....	41
Quadro 12 – Avaliação do ChatDuca pelo usuário	42

LISTA DE ABREVIATURAS E SIGLAS

AIML – Artificial Intelligence Markup Language

BIA – Bradesco Inteligência Artificial

FAQ – Frequently Ask Questions

FURB – Fundação Universidade Regional de Blumenau

IBM – International Business Machines

RF – Requisitos Funcionais

RNF - Requisitos Não Funcionais

UC – Use Case

UML – Unified Modeling Language

XML – eXtensible Markup Language

SUMÁRIO

1 INTRODUÇÃO.....	12
1.1 OBJETIVOS.....	13
1.2 ESTRUTURA.....	13
2 FUNDAMENTAÇÃO TEÓRICA	14
2.1 CHATTERBOTS	14
2.2 AIML.....	15
2.3 TRABALHOS CORRELATOS	16
2.3.1 ChatterEDU.....	16
2.3.2 Robô Andrew	18
2.3.3 <i>Chatterbot</i> sobre o processo de atualização de software	19
3 DESENVOLVIMENTO.....	22
3.1 LEVANTAMENTO DE INFORMAÇÕES	22
3.2 ESPECIFICAÇÃO	23
3.2.1 Casos de uso.....	23
3.2.2 Diagrama de atividades	23
3.3 IMPLEMENTAÇÃO	24
3.3.1 Técnicas e ferramentas utilizadas.....	24
3.3.2 Organização e manipulação das bases de conhecimento AIML.....	25
3.3.3 Organização e manipulação do banco de dados.....	28
3.3.4 Operacionalidade da implementação	29
3.4 RESULTADOS E DISCUSSÕES.....	33
4 CONCLUSÕES.....	36
4.1 EXTENSÕES	37
REFERÊNCIAS	38
APÊNDICE A – DESCRIÇÃO DOS CASOS DE USO	40
APÊNDICE B – TESTE DO <i>CHATTERBOT</i>	42

1 INTRODUÇÃO

Desde o surgimento da internet no início dos anos 90, o acesso à informação ficou mais rápido e fácil.

Estão disponíveis na Web milhares de páginas cobrindo os mais variados assuntos e interesses. Estimativas [...] de 2004] afirmam existir cerca de 2,5 bilhões de documentos, com uma taxa de crescimento de 7,5 milhões ao dia. Mas, diferentemente das bibliotecas, os documentos da Internet não estão classificados segundo um padrão determinado. Portanto, o usuário precisa localizar informações de um grande volume de páginas disponíveis, sem qualquer organização. (BRANSKI, 2004, p. 71).

Dez anos depois, em 2014,

O universo digital já é tão vasto quanto o real e continua se expandindo. Estudo da EMC, empresa líder do mercado internacional de armazenamento de dados, [...] revela que já existem disponíveis hoje no mundo quase 1 septilhão de bits de informação — ou o número 1 seguido de 24 zeros, [...]. A estimativa é que, até 2020, o número de dados armazenados em computadores, servidores, celulares, smartphones e tablets seja, no mínimo, multiplicado por seis. (MACHADO, 2014, p. 1).

Partindo da premissa que o crescimento da utilização da internet é constante, conclui-se que também é crescente o uso de sistemas e, conseqüentemente, o número de dúvidas de usuários na utilização de sistemas e de seus processos. Por isso, as empresas buscam alternativas para levar a informação mais rápida ao usuário, tais como: Frequently Ask Questions (FAQ), manual de usuário, páginas de edição colaborativas (*wiki*), Intranet, entre outras. Contudo, em geral, estas alternativas são estáticas e são construídas por uma equipe de pessoas da própria empresa. Logo, tornam-se desatualizadas muito facilmente. Com a evolução das tecnologias e a necessidade cada vez maior de acesso a informações confiáveis, os usuários necessitam de meios que possam lhes auxiliar de forma rápida.

Como opção para suporte técnico, verifica-se que algumas empresas ainda hoje investem em atendentes que tiram dúvidas pelo telefone. Para uma empresa, ter vários atendentes disponíveis representa um custo elevado além desse tipo de suporte nem sempre agradar a todos os usuários devido ao custo das ligações telefônicas. Em função disso, certas empresas começaram a investir no desenvolvimento de *chatbots*. Como exemplos, cita-se a Sara, a assistente virtual da Senior Sistemas (SENIOR SISTEMAS, 2017), e a BIA, a inteligência artificial do Bradesco (SILVA, 2018). Segundo Senior Sistemas (2017), a Sara é uma assistente virtual baseada em uma plataforma de computação cognitiva em sistemas para gestão empresarial. Os autores afirmam que Sara aumenta a produtividade dos profissionais e o acesso mais rápido e eficiente a informações para tomada de decisão, podendo responder a comandos de voz, apresentar o status dos clientes cadastrados, atualizar dados e conferir a agenda de reuniões. Já a BIA é um *chatbot* desenvolvido a partir da plataforma de computação cognitiva Watson em parceria com a IBM que funciona com base no comportamento do usuário

com o objetivo de “solucionar dúvidas e problemas comuns dos clientes da forma mais natural possível, facilitando a experiência de uso, 24 horas por dia, 7 dias por semana” (SILVA, 2018).

Segundo Teixeira e Menezes (2003), *chatbot* é um programa de computador que tenta simular um ser humano na conversação com as pessoas. O objetivo é responder às perguntas de tal forma que os usuários tenham a impressão de estar conversando com outra pessoa e não com um programa de computador. Como base para funcionamento de um *chatbot* é utilizada uma base de conhecimento para realizar a conversação com o usuário. Essa base de conhecimento é o local onde todo o conhecimento do *chatbot* fica adicionado e é com ela que o *chatbot* consegue ter perguntas e respostas e assim, manter uma conversação de forma natural com o usuário.

Diante do exposto, este trabalho apresenta o desenvolvimento de um *chatbot* capaz de tirar dúvidas de usuários em relação a um software de gestão educacional, para auxiliar na resolução de questões relacionadas a boletos, notas e frequências, buscando diminuir a demanda de suporte de uma organização.

1.1 OBJETIVOS

O objetivo desse trabalho é desenvolver um *chatbot* para tirar dúvidas de usuários sobre um software de gestão educacional.

Tem-se como objetivos específicos:

- a) disponibilizar uma interface web para interação entre o usuário e o *chatbot*;
- b) complementar a base de conhecimento do *chatbot* a partir da interação com o usuário.

1.2 ESTRUTURA

O trabalho realizado está dividido em capítulos. O segundo capítulo apresenta os conceitos pertinentes ao tema e descreve três trabalhos correlatos ao projeto desenvolvido. No terceiro capítulo são especificados os requisitos funcionais e não funcionais do software, relacionadas as ferramentas que foram utilizadas no desenvolvimento, como foram organizadas as bases de conhecimento e descrita a operacionalidade da implementação. No quarto e último capítulo são descritas a conclusão do trabalho e algumas ideias propostas para desenvolvimento futuro.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo tem como objetivo explorar os assuntos que foram necessários para o entendimento e a realização do trabalho. Desta forma, foi subdividido em três seções, sendo que a seção 2.1 conceitua *chatterbots* e mostra sua evolução com base nas suas três mais importantes gerações, além de descrever sua arquitetura, enquanto a seção 2.2 apresenta a linguagem Artificial Intelligence Markup Language (AIML) e, por fim, a seção 2.3 descrever três trabalhos correlatos a esse projeto.

2.1 CHATTERBOTS

Segundo Comarella e Café (2008), *chatterbots* são programas que simulam uma conversa, como as estabelecidas entre seres humanos, sendo utilizados com os mais diversos propósitos, tais como um “amigo virtual” ou para uso comercial.

O objetivo é responder às perguntas de tal forma que as pessoas tenham a impressão de estar conversando com outra pessoa e não com um programa de computador. Após o envio de perguntas em linguagem natural, o programa consulta uma base de conhecimento e em seguida fornece uma resposta que tenta imitar o comportamento humano (TRENTIN, 2016, p. 11).

Moura (2003) classifica os *chatterbots* seguindo a evolução das tecnologias. O Quadro 1 apresenta as gerações e algumas de suas principais características.

Quadro 1 – Classificação evolutiva dos *chatterbots*

geração	característica
1 ^a	casamento de padrão e uso de regras gramaticais
2 ^a	técnicas de inteligência artificial
3 ^a	linguagens de marcação

Fonte: Trentin (2016, p. 14).

A 1^a geração teve como *chatterbot* principal Eliza, desenvolvida por Weizenbaum (1966), que com um pouco mais de 200 linhas de código, comparava o dado informado pelo usuário com os armazenados em uma base de conhecimento, chamada de lista de padrões. O *chatterbot* buscava o padrão mais semelhante, quebrava o dado de entrada com base em regras gramaticais e então dava a resposta ao usuário.

A 2^a geração teve como *chatterbot* principal Júlia, que foi desenvolvida por Mauldin (1994) e tinha como objetivo auxiliar um usuário durante um jogo. Nessa fase, começou-se a integrar o conceito de inteligência artificial aos *chatterbots*. No caso da Júlia, dependendo das jogadas feitas pelo usuário, ela conseguia se adaptar ao novo contexto gerado e dar uma resposta mais assertiva. A integração com inteligência artificial também ficou marcada por dar mais realismo à conversação com o *chatterbot*.

A 3ª geração apresenta Alice, desenvolvida por Wallace (2018), como principal *chatbot*. Ela possuía vários módulos de conversação que eram divididos por idade, gênero, profissão e até localização geográfica. Alice também tinha a capacidade de armazenar com quem (usuário) e sobre o que (assunto) estavam conversando. Isso fazia com que o usuário não pudesse alterar de assunto de forma brusca e que o *chatbot* tivesse mais assertividade em suas respostas.

Para o desenvolvimento de um *chatbot*, deve-se elaborar a base de conhecimento, que armazena todas as perguntas e respostas sobre o(s) assunto(s) que o *chatbot* sabe conversar. É necessário que exista também um interpretador da base de conhecimento para processar os dados fornecidos pelo usuário e para buscar as respectivas informações na base de conhecimento, fornecendo-as ao usuário. Essa estrutura é básica em qualquer *chatbot* e ela cria a arquitetura base para o funcionamento do mesmo.

2.2 AIML

A linguagem AIML nasceu no âmbito do projeto de desenvolvimento do *chatbot* Alice. Para os *chatbots*, a AIML é a linguagem que é utilizada para criar as bases de conhecimento. A linguagem trabalha com base na eXtensible Markup Language (XML) e com a utilização das *tags* é possível criar e organizar uma árvore de perguntas e respostas, permitindo também que o *chatbot* encontre a resposta mais facilmente.

Uma base de conhecimento AIML é formada, segundo Marietto et al. (2013), por blocos básicos de diálogos, chamados de categorias, onde cada categoria contém um padrão que representa uma sentença do usuário e uma resposta corresponde [sic] que o *chatbot* dará ao usuário. (OLIVEIRA, 2015, p. 15).

No Quadro 2, tem-se um exemplo de uma base de conhecimento em AIML. Observa-se no Quadro 2 que:

- a) na linha 1 tem-se a *tag* <aiml>, que é usada para delimitar o início e o fim do documento AIML;
- b) da linha 2 até a linha 11 é definida uma categoria com a *tag* <category>, que engloba uma possível sentença/pergunta do usuário (*tag* <pattern>) e a respectiva resposta que o *chatbot* dará (*tag* <template>). Na *tag* <pattern> é utilizado o asterisco para indicar que podem existir uma ou mais palavras antes ou depois da sentença TUDO BEM, como em “Olá, TUDO BEM com você?”. Na resposta do *chatbot* é utilizada a *tag* <random> para randomicamente escolher uma das respostas que estão nas *tags* .
- c) da linha 23 em diante é definida uma nova categoria, mas dentro dela utiliza-se a *tag*

<srail>. Esta *tag* faz um link com outra pergunta da base de conhecimento. Neste exemplo, a *tag* <pattern> *NOTA* faz referência à *tag* <pattern> NOTA e, desta forma, se a pergunta for, por exemplo, “Gostaria de saber como ver as NOTAS do meu filho.”, será apresentada a resposta associada à *tag* <pattern> NOTA (da linha 15 a linha 20).

Quadro 2 – Trecho de arquivo AIML

01	<aiml>
02	<category>
03	<pattern>*TUDO BEM*</pattern>
04	<template>
05	<random>
06	Tudo em ordem :D, no que posso ser útil hoje?
07	Muito bem, obrigado :D, no que posso ser útil hoje?
08	Eu estou bem :D, no que posso ser útil hoje?
09	</random>
10	</template>
11	</category>
12	
13	<category>
14	<pattern>NOTA</pattern>
15	<template> Para você conferir as suas notas e as suas médias, acesse a
16	opção "Notas e Frequências" (à esquerda). Caso você queira conferir suas
17	Notas Parciais em uma disciplina, você pode clicar em cima da média da
18	disciplina desejada. Se a sua média não estiver disponível, você pode clicar
19	em cima da sigla NP.
20	</template>
21	</category>
22	
23	<category>
24	<pattern>*NOTA*</pattern>
25	<template>
26	<srail>NOTA</srail>

Fonte: elaborado pelo autor.

2.3 TRABALHOS CORRELATOS

São apresentados trabalhos com características semelhantes aos principais objetivos do estudo desenvolvido. O primeiro é o ChatterEDU, um *chatbot* de suporte a alunos que pretendem estudar Geografia (MARTINS, 2016). O segundo é um *chatbot* para dar suporte a possíveis calouros da FURB, chamado Robô Andrew (OLIVEIRA, 2015). O terceiro é um *chatbot* de suporte ao usuário no processo de atualização de um software (TRENTIN, 2016).

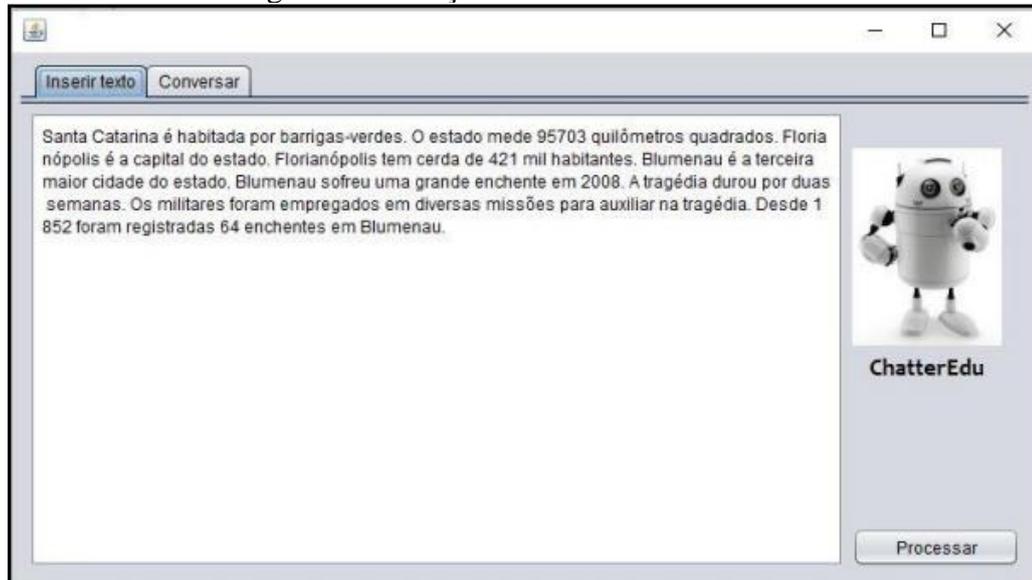
2.3.1 ChatterEDU

Martins (2016) desenvolveu um *chatbot* para auxílio acadêmico que tem a capacidade de conversar em língua portuguesa, respondendo ou até mesmo fazendo perguntas sobre um determinado texto inserido pelo usuário. Observa-se que a base de conhecimento do *chatbot* é construída a partir do texto inserido. A proposta inicial era tratar textos da educação básica da área de Geografia. No entanto, da forma como foi implementado, o ChatterEDU trata de textos

de qualquer área de conhecimento desde que não possuam informações que necessitem de interpretação de contexto, sendo que cada oração deve ser composta por sujeito, verbo e objeto.

O processo de conversação do ChatterEDU começa com a inserção do texto para geração da base de conhecimento, conforme pode ser visto na Figura 1.

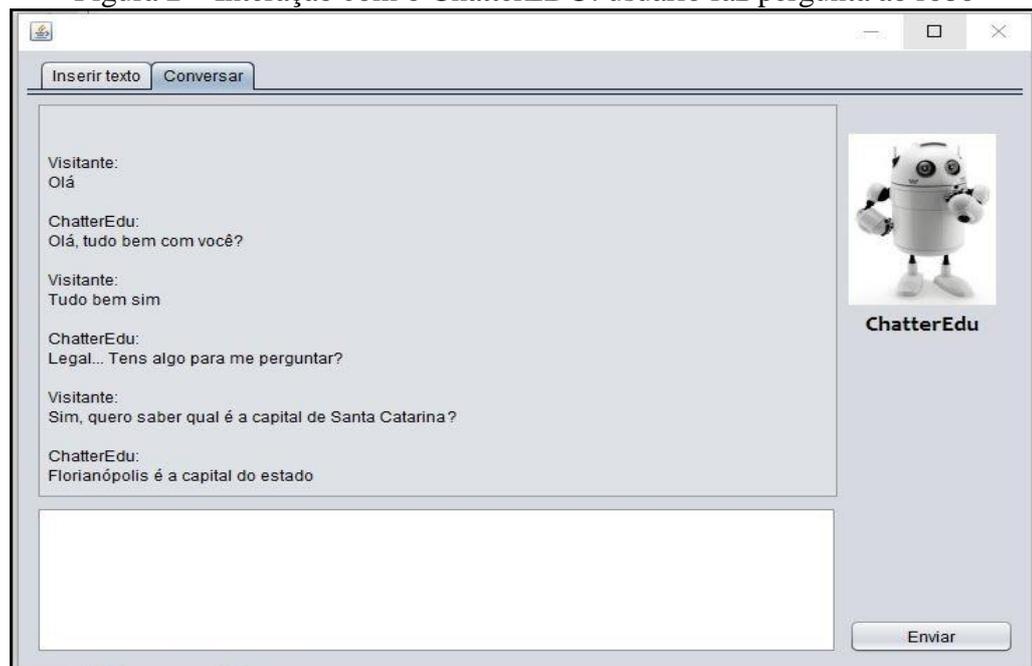
Figura 1 – Inserção de texto no ChatterEDU



Fonte: Martins (2016, p. 37).

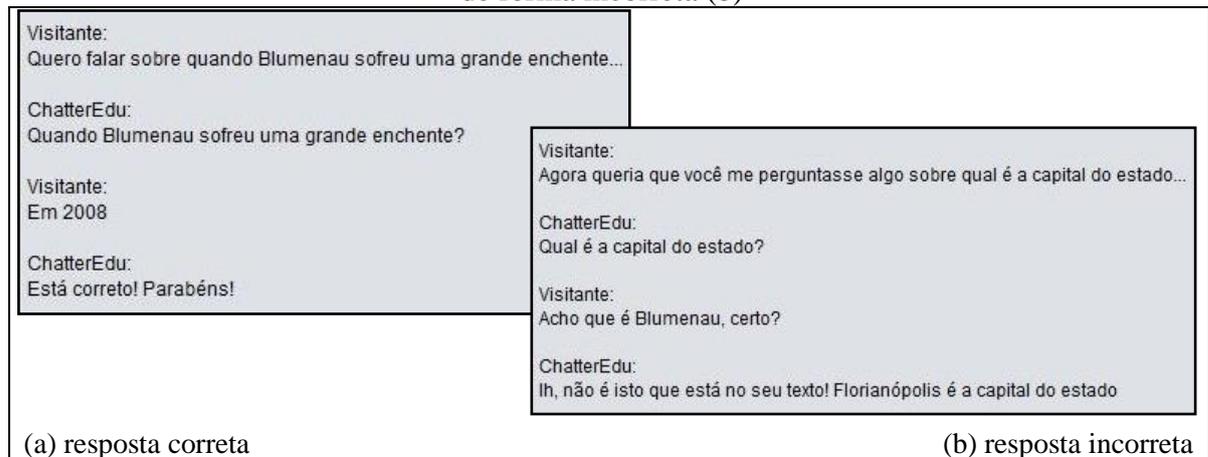
A partir disso, são geradas perguntas e respostas que ficam armazenadas em uma base de conhecimento em formato AIML e o usuário é redirecionado para a guia de conversação. A Figura 2 e a Figura 3 ilustram o ChatterEDU interagindo com o usuário.

Figura 2 – Interação com o ChatterEDU: usuário faz pergunta ao robô



Fonte: Martins (2016, p. 38).

Figura 3 – Interação com o ChatterEDU: usuário responde uma pergunta de forma correta (a) de forma incorreta (b)



Fonte: adaptado de Martins (2016, p. 38-39).

Deve-se dizer que o ponto chave para a geração das perguntas e respostas que constituem a base de conhecimento são os papéis semânticos, como explicado pela autora do trabalho.

O ponto principal para a geração das perguntas é o papel semântico, a partir do qual é possível definir o pronome ou o advérbio interrogativo, tais como quem, quando, quanto, qual e onde, que deve ser usado para elaborar a pergunta. Além disso, são feitos ajustes para possibilitar a conjugação verbal em determinados casos e criar perguntas mais concisas. (MARTINS, 2016, p. 7).

Para exemplificar a utilização e funcionamento dos papéis semânticos, tem-se que na frase “Florianópolis é a capital do estado.” a palavra “Florianópolis” possui o papel semântico TH, ou seja, o tema principal da oração. Com base no papel semântico identificado, são geradas as perguntas “Qual é a capital do estado?” ou “O que é Florianópolis?” e as respectivas respostas “Florianópolis” e “Florianópolis é a capital do estado”.

Vale salientar que, segundo Martins (2016), a base de conhecimento do ChatterEDU não é incrementada, toda a base de conhecimento é substituída com informações contidas no texto de entrada processado. No entanto, o ChatterEDU é uma ferramenta que gera perguntas e respostas automaticamente. Observa-se que a maioria dos objetivos do trabalho foram atingidos, exceto pela plataforma. O sistema foi desenvolvido para desktop, porém esperava-se que o ChatterEDU fosse desenvolvido como uma aplicação web, para que ficasse disponível a todas as pessoas.

2.3.2 Robô Andrew

O *chatbot* desenvolvido por Oliveira (2015) tem a proposta de auxiliar pessoas em dúvidas sobre as maneiras de ingressar em cursos na FURB através de uma página web. Para que o *chatbot* possa responder as perguntas do usuário, existem bases de conhecimento em

AIML que armazenam essas informações. Na Figura 4, pode-se ver uma sessão de conversação do Robô Andrew com o usuário.

Figura 4 – Interação entre um usuário e o *chatterbot*



Fonte: Oliveira (2015, p. 50).

Cumprе observar preliminarmente que, conforme descrito por Oliveira (2015), a forma de interpretação das bases de conhecimento em AIML faz uso de *tags* específicas para verificar uma ontologia sobre possíveis dúvidas no ingresso em cursos da FURB, expressas de formas diferentes e desta forma aprimorando as respostas. A ontologia foi modelada com as informações específicas das formas de ingresso para calouros, tais como cursos e vagas ofertados, semestre de oferta, data de divulgação dos resultados, entre outras. Para extrair informações das frases que o usuário escreve, é utilizada a ferramenta SPARQL para identificar o sujeito, o verbo e o objeto das frases. As informações da ontologia podem ser alteradas por meio de uma área administrativa do sistema. Dessa forma é possível que o conhecimento do *chatterbot* seja mantido e acrescido.

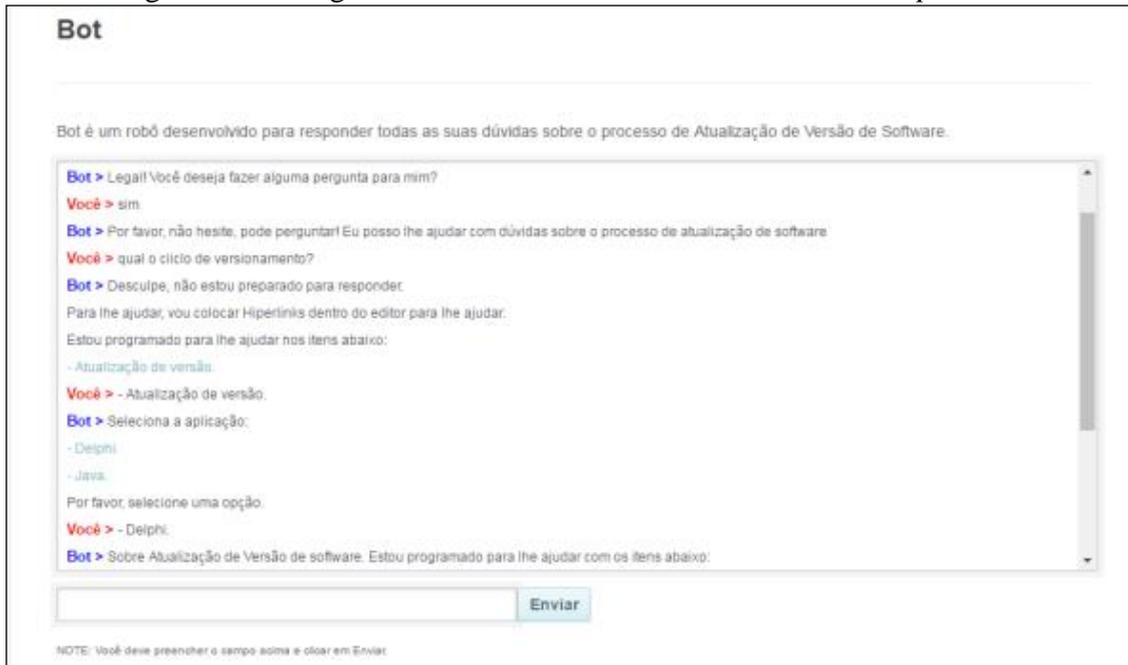
Oliveira (2015) relatou que a criação manual de bases de conhecimento é trabalhosa e pode haver dificuldade de identificar a palavra se possuir erros ortográficos. No entanto, mesmo assim, o resultado do trabalho foi satisfatório e o *chatterbot* conseguiu dialogar com os usuários.

2.3.3 *Chatterbot* sobre o processo de atualização de software

No caso em análise, Trentin (2016) apresentou o desenvolvimento de um *chatterbot* que tem como principal objetivo a interação com o usuário para auxílio ao processo de atualização de versão de um software para gestão hospitalar, a partir de uma página web. Tendo em vista a

intenção de facilitar a interação com o *chatterbot*, caso o usuário questione algo que o *chatterbot* não tenha conhecimento necessário para sanar sua dúvida, é iniciado um diálogo com *hiperlinks*. Salienta-se que para que haja a criação dos *hiperlinks*, consulta-se a base de conhecimento desenvolvida utilizando AIML e usam-se *tags* específicas para formar as respostas. A Figura 5 ilustra um diálogo com o uso de *hiperlinks*.

Figura 5 – Diálogo entre o *chatterbot* e o usuário utilizando *hiperlinks*



Fonte: Trentin (2016, p. 47).

O uso de *hiperlinks* é feito através de um conjunto de *tags* no arquivo AIML. Tem-se que a *tag* <hiperlink> é composta por uma *tag* <header> e uma *tag* <option>. A *tag* <header> possui o texto que será mostrado ao usuário na conversação e a *tag* <option> armazena a informação que será consultada na base de conhecimento. Na Figura 6 é mostrada a *tag* <hiperlink> e a Figura 7 ilustra como as informações são mostradas para o usuário.

Figura 6 – Exemplo de *tag* <hiperlink>

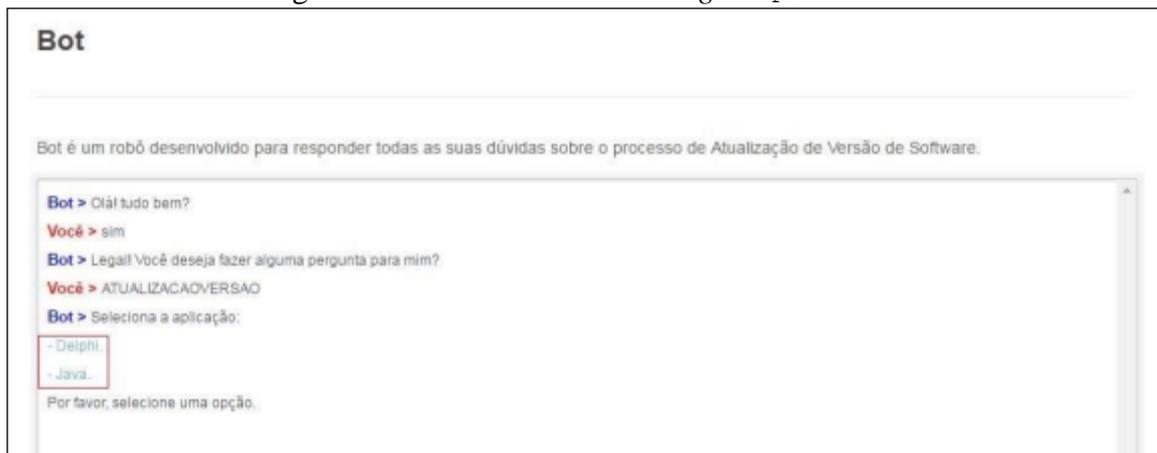
```

1  <?xml version="1.0" encoding="UTF-8"?>
2
3  <aiml>
4
5      <category>
6          <pattern>ATUALIZACAOVERSAO</pattern>
7          <template>Seleciona a aplicação:
8          <hiperlink><header>- Delphi. </header><option>atualizaVersaoDelphi</option></hiperlink>
9          <hiperlink><header>- Java. </header><option>atualizaVersaoJava</option></hiperlink>
10         <srai>SELECIONAOPCAO</srai></template>
11     </category>
12
13 </aiml>

```

Fonte: Trentin (2016, p. 28).

Figura 7 – Resultado do uso da tag <hiperlink>



Fonte: Trentin (2016, p. 28).

Trentin (2016) diz que foi perceptível a dificuldade de responder sentenças com erros de digitação, necessitando a criação de vários padrões de domínios semelhantes para as respostas. Além disso, Trentin (2016) relatou que o tempo necessário para cadastrar bases de conhecimento é elevado, visto que não existe uma ferramenta que auxilie ou automatize esse processo e todo o cadastro da base de conhecimento deve ser feito de forma manual. É importante ressaltar que o autor utilizou um banco de dados MySQL no auxílio do desenvolvimento do *chatterbot*, contendo informações do software de gestão hospitalar, que a cada nova versão do software é atualizado, dispensando a manutenção da base de conhecimento AIML. Assim, o *chatterbot* também consulta o banco de dados para responder as dúvidas sobre o processo de atualização do software.

3 DESENVOLVIMENTO

Este capítulo apresenta o processo de desenvolvimento do software desenvolvido no trabalho. A seção 3.1 descreve os requisitos funcionais e não funcionais identificados. Na seção 3.2 são mostrados os artefatos gerados durante o processo de modelagem do sistema. A seção 3.3 trata da implementação do sistema. Por fim, na seção 3.4 são discutidos os resultados alcançados.

3.1 LEVANTAMENTO DE INFORMAÇÕES

O *chatterbot* desenvolvido, denominado ChatDuca, deverá atender os seguintes Requisitos Funcionais (RF) e Requisitos Não Funcionais (RNF):

- a) RF01: permitir que um usuário crie um login para acessar o *chatterbot*, informando nome, senha e papel, podendo optar entre ser um responsável ou um estudante;
- b) RF02: possuir uma tela de login que permita somente usuários cadastrados acessarem o *chatterbot*;
- c) RF03: permitir que o usuário interaja com o *chatterbot* através de uma página web;
- d) RF04: permitir a interação em língua portuguesa;
- e) RF05: permitir que o *chatterbot* responda perguntas sobre notas, frequência e boletos;
- f) RF06: exibir o histórico da conversa durante o diálogo com o *chatterbot*;
- g) RF07: permitir que o usuário informe, durante o diálogo, que um tema, para o qual o *chatterbot* não encontra uma resposta, refere-se a um determinado assunto das bases de conhecimento do *chatterbot* ou é um assunto novo;
- h) RF08: ter uma tela de administração para que o usuário com acesso de administrador possa incrementar as bases de conhecimento em AIML a partir da interação com o usuário;
- i) RNF01: o *backend* deve ser desenvolvido em JavaScript usando NodeJS;
- j) RNF02: o *frontend* deve ser desenvolvido em JavaScript usando Angular junto com o *framework* Materialize e com a biblioteca jQuery;
- k) RNF03: o banco de dados utilizado para armazenar informações sobre assuntos desconhecidos do *chatterbot* e sobre os usuários deve ser o MySQL;
- l) RNF04: o conhecimento do *chatterbot* deve ser armazenado em bases de conhecimento AIML, sendo uma para o responsável e outra para o estudante;
- m) RNF05: a biblioteca usada para interpretar as bases de conhecimento deve ser a *AimlInterpreter*.

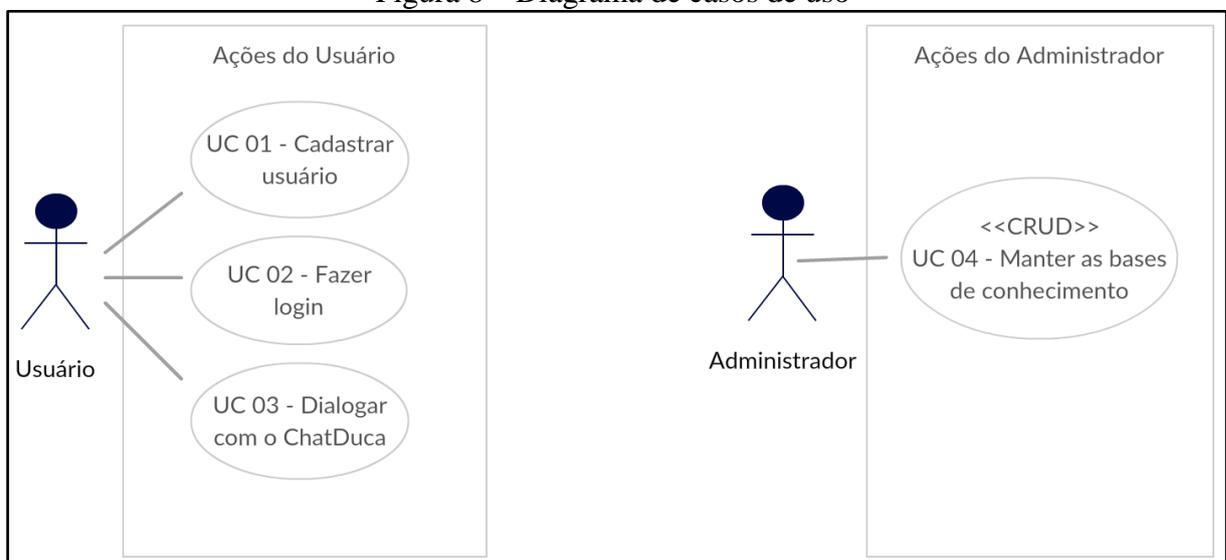
3.2 ESPECIFICAÇÃO

Nesta seção é apresentada a especificação da solução desenvolvida, através dos casos de uso e do diagrama de atividades da Unified Modeling Language (UML).

3.2.1 Casos de uso

Este projeto contém quatro casos de uso (em inglês Use Case - UC), representados na Figura 8 e detalhados no Apêndice A. Estes quatro casos de uso foram separados entre dois atores, um denominado Usuário e o outro denominado Administrador. O Usuário possui três casos de uso, sendo o primeiro deles “Cadastrar usuário”, que possibilita que seja feito um cadastro para acesso ao *chatterbot*. Os casos de uso “Fazer login” e “Dialogar com o ChatDuca” representam, respectivamente, que o usuário poderá acessar o *chatterbot* informando suas credenciais e, após isso, conversar com o ChatDuca. Essas três ações podem ser feitas por usuários do papel estudante ou responsável. Já o Administrador possui somente uma ação, que é manter as bases de conhecimento, permitindo que o ChatDuca aumente a quantidade e a variedade de assuntos sobre os quais pode responder.

Figura 8 – Diagrama de casos de uso



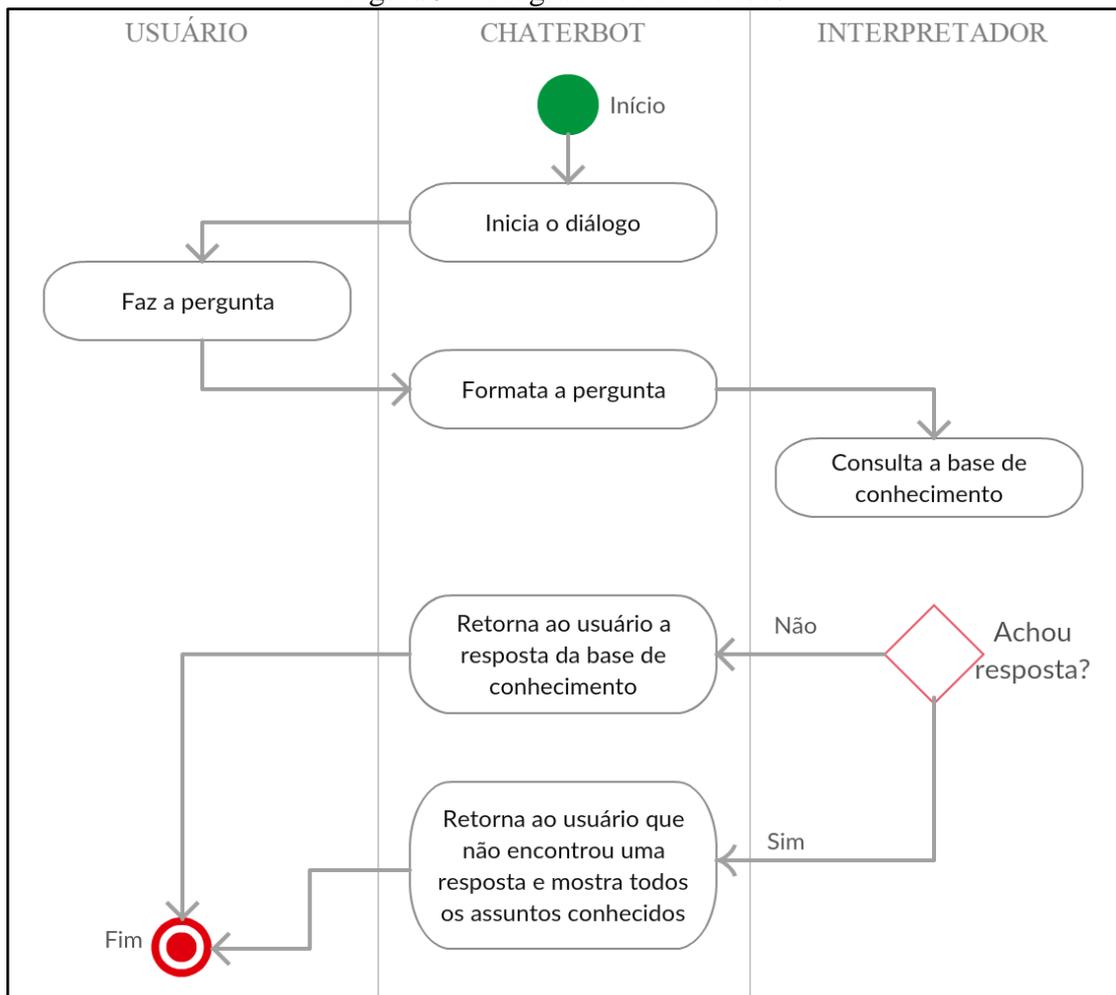
Fonte: elaborada pelo autor.

3.2.2 Diagrama de atividades

Esta seção apresenta o diagrama de atividades do ChatDuca. A Figura 9 mostra o processo de interação entre o usuário e o *chatterbot* e como o *chatterbot* responde uma pergunta do usuário. Após o usuário fazer corretamente o login na aplicação, é dado início ao diálogo a partir de uma mensagem de saudação do ChatDuca. Na sequência, o usuário pode fazer perguntas para o *chatterbot*. O ChatDuca recebe e formata a pergunta, para então enviá-la para

o interpretador das bases de conhecimento AILM. O interpretador recebe a pergunta e busca a respectiva resposta na base de conhecimento correspondente ao papel do usuário autenticado (estudante ou responsável). Caso encontre a resposta sobre o assunto perguntado, devolve a mesma ao *chatterbot*, caso contrário, informa sobre quais assuntos tem conhecimento. Após o *chatterbot* receber a resposta, ele mostra para o usuário e o processo de um questionamento é encerrado.

Figura 9 – Diagrama de atividades



Fonte: elaborada pelo autor.

3.3 IMPLEMENTAÇÃO

Nesta seção são apresentadas as ferramentas utilizadas para o desenvolvimento do projeto. É apresentada a organização e a manipulação das bases de conhecimento e do banco de dados, assim como a operacionalidade da implementação.

3.3.1 Técnicas e ferramentas utilizadas

No desenvolvimento do *chatterbot* foram usadas as seguintes ferramentas e bibliotecas:

- a) Visual Studio Code: ambiente *open source*, mantido pela Microsoft, utilizado no desenvolvimento da aplicação;
- b) NodeJS 8.10.0: plataforma de programação utilizada para o desenvolvimento do servidor *backend* da aplicação;
- c) Angular 6.0.3: *framework* JavaScript de código aberto, mantido pelo Google, que permite manipular os dados informados no *frontend* da aplicação;
- d) JavaScript: linguagem de programação interpretada que foi utilizada em conjunto com a biblioteca jQuery na versão 3.4.1;
- e) AimlInterpreter: biblioteca utilizada para pesquisar nas bases de conhecimento AIML do *chatbot* a pergunta feita pelo usuário;
- f) InsertLine: biblioteca utilizada para ler, excluir e escrever o texto das bases de conhecimento AIML;
- g) Materialize: *framework frontend* baseado em Material Design usado para desenvolver telas responsivas;
- h) MySQL versão 5.7.22: banco de dados utilizado para armazenar os dados dos usuários, tais como login, nome, senha e papel, podendo ser estudante, responsável ou administrador. Também é usado para armazenar as perguntas que o *chatbot* não conseguiu responder, fazendo o vínculo com o assunto que o usuário informou.

3.3.2 Organização e manipulação das bases de conhecimento AIML

O ChatDuca utiliza duas bases de conhecimento para conversar com os usuários. Uma das bases é carregada quando o usuário possui o papel de estudante e a outra é carregada quando o usuário possui o papel de responsável. As bases de conhecimento são salvas utilizando a extensão `.aiml.xml` e são lidas, interpretadas e manipuladas através da biblioteca AimlInterpreter do NodeJS. Foi feita a separação das bases de conhecimento visando melhorar a comunicação com estudantes e com responsáveis, pois, assim, é possível dialogar de forma diferente com cada um. Com estudantes pode-se utilizar uma linguagem mais informal, diferente da linguagem usada com responsáveis. Desta forma, também é possível que a base de conhecimento do estudante tenha assuntos que não existem na base do responsável e vice-versa.

As bases de conhecimento são estruturadas por assuntos e temas. Cada assunto possui uma resposta e pelo menos um tema vinculado a ele, qual seja, o assunto precedido e seguido por asteriscos. No Quadro 3 é mostrado um trecho da base de conhecimento do estudante onde existe o assunto `NOTA PARCIAL` (linha 2) e sua respectiva resposta (linhas 3 a 13). Já no Quadro 4 são mostrados dois temas vinculados ao assunto anterior, o tema `*NOTA PARCIAL*` e o tema

NP, nas linhas 2 e 8, respectivamente. No Quadro 4 pode-se perceber que as bases de conhecimento são estruturadas de tal forma que cada tema fica em um conjunto de *tags* separado. Com a definição de temas é possível fazer perguntas como “O que é uma nota parcial?” ou “Você poderia, por favor, explicar o que é NP e onde eu posso visualizar.”, sendo a resposta a mesma associada a `NOTA PARCIAL`, ou seja, evita-se duplicação de conteúdo.

Quadro 3 – Assunto e resposta da base de conhecimento

01	<category>
02	<pattern>NOTA PARCIAL</pattern>
03	<template>
04	As Notas Parciais (N.P.) são as notas que você obteve em cada atividade
05	realizada em uma disciplina durante o trimestre. Caso você queira conferir
06	suas Notas Parciais em uma disciplina, você pode clicar em cima da média, se
07	estiver disponível, ou em cima da sigla NP, caso contrário. As informações
08	serão apresentadas em outra janela, contendo, para cada avaliação, a data, o
09	tipo (prova, trabalho, recuperação, outra), o assunto e a nota obtida. Se já
10	estiver disponível, você poderá conferir também a fórmula (FM) utilizada para
11	calcular a média trimestral, a média na disciplina (MP) e a média após a
12	recuperação (MX).
13	</template>
14	</category>

Fonte: elaborado pelo autor.

Quadro 4 – Assunto e tema da base de conhecimento

01	<category>
02	<pattern>*NOTA PARCIAL*</pattern>
03	<template>
04	<srain>NOTA PARCIAL</srain>
05	</template>
06	</category>
07	<category>
08	<pattern>*NP*</pattern>
09	<template>
10	<srain>NOTA PARCIAL</srain>
11	</template>
12	</category>

Fonte: elaborado pelo autor.

A partir dessa estruturação das bases de conhecimento é possível montar a mensagem de saudação do ChatDuca, onde são listados os assuntos conhecidos. É considerado um assunto qualquer sequência de caracteres que estiver em uma *tag* <pattern> sem estar precedida e seguida por asteriscos. No *backend*, existe uma rotina que retorna os assuntos conhecidos da base de conhecimento do estudante e outra que retorna os assuntos do responsável. Conforme mostrado no Quadro 5, esta rotina: lê a base de conhecimento correspondente, do estudante (linha 2); coloca o conteúdo do arquivo em uma variável (linha 4); através de uma expressão regular, identificam todas as *tags* <pattern> que possuem somente caracteres entre A a Z (linhas 5 e 6); colocam cada *tag* em um *array* de assuntos (linha 7).

Quadro 5 – Rotina que retorna assuntos conhecidos

```

01 const route = router.get('/1', (req, res, next) => {
02   var caminhoEstudante = './src/aiml/base_estudante.aiml.xml';
03   fs.readFile(caminhoEstudante, function(err, buf) {
04     var conteudoArquivo = buf.toString();
05     var regex = new RegExp('<pattern>[A-Z]+(\\s*[A-Z]+)*</pattern>', 'g');
06     var assuntos = conteudoArquivo.match(regex);
07     res.status(200).send({assuntos: assuntos});
08   });
09 });

```

Fonte: elaborado pelo autor.

É de responsabilidade do administrador do *chatterbot* manter as bases de conhecimento atualizadas. Para tanto, o administrador pode editar manualmente as bases de conhecimento para incluir assuntos, respectivas respostas e temas, ou pode, usando a funcionalidade correspondente (UC04: Manter bases de conhecimento), adicionar novos assuntos ou associar temas a assuntos existentes a partir de sugestões feitas pelo usuário durante o diálogo com o ChatDuca. Para incrementar a base de conhecimento, o administrador pode realizar duas ações, incluir um assunto novo e sua respectiva resposta ou incluir um tema relacionado com um assunto já existente. Independente se o administrador incluir um assunto novo ou incluir um tema, o Angular irá redirecionar estas informações para uma única rotina no *backend*. No Quadro 6 pode-se ver um trecho desse código, onde é validado se está sendo incluído um assunto novo (linha 1) para, então, validar se o novo assunto deverá ser inserido nas duas bases de conhecimento (linha 2) ou em apenas uma delas (linha 9). Dependendo dessa validação, é feita a chamada dos métodos que fazem a escrita nas duas bases de conhecimento (entre as linhas 3 e 5), ou apenas em uma delas - somente na base do estudante (linha 10) ou somente na base do responsável (linha 12).

Quadro 6 – Código para incluir assuntos ou temas nas bases de conhecimento

```

01 if (req.body.novo !== undefined) {
02   if (req.body.inserirEstudante == true && req.body.inserirResponsavel == true) {
03     escreveEstudante(req.body.pergunta, req.body.novo).then(
04       () => {
05         escreveResponsavel(req.body.pergunta, req.body.novo).then(() => {});
06       }
07     );
08   } else {
09     if (req.body.inserirEstudante == true) {
10       escreveEstudante(req.body.pergunta, req.body.novo).then(() => {});
11     } else {
12       escreveResponsavel(req.body.pergunta, req.body.novo).then(() => {});
13     }
14   }

```

Fonte: elaborado pelo autor.

Para adicionar um novo assunto na base de conhecimento, inicialmente, a pergunta é formatada. Em uma variável do tipo *string*, é armazenada uma *tag* `<pattern>` com a pergunta, toda em letras maiúsculas, e é criada uma nova *tag* `<pattern>` que inclui um tema sobre a pergunta, ou seja, a própria pergunta precedida e sucedida por asteriscos, juntamente com a *tag*

`<srai>` que referencia o assunto sem os asteriscos. Essas *tags* sempre são inseridas no final das bases de conhecimento, baseando-se na *tag* `</aiml>`. Após ter definido as *tags* que serão inseridas, é feita a leitura da base de conhecimento e todo o conteúdo é alocado em uma variável chamada `conteudoArquivo`. Com essa variável, é feita a identificação da *tag* `</aiml>` e essa *tag* é substituída pelo conteúdo definido anteriormente. Após esse procedimento, tem-se na variável `conteudoArquivo` a base de conhecimento incrementada e então realiza-se a exclusão da base de conhecimento atual e a criação da nova base de conhecimento, com o novo conteúdo que foi definido. Após esse procedimento, é feita a leitura da base de conhecimento nova, para que o ChatDuca consiga responder sobre o novo assunto. O processo de inclusão de um novo tema é semelhante ao da inclusão de um novo assunto, o que os diferencia é que a variável que armazena o conteúdo das *tags* irá armazenar somente a *tag* `<pattern>` com o tema, precedido e sucedido por asteriscos, e a *tag* `<srai>`, referenciando o assunto correspondente.

3.3.3 Organização e manipulação do banco de dados

O banco de dados possui duas tabelas que não têm nenhum relacionamento entre si. Uma delas é a tabela `usuarios`, que armazena nome, login, senha, papel e um código sequencial dos usuários habilitados a utilizar o *chatbot*. O campo que armazena o login do usuário possui uma chave única, o que restringe ele a ser único no banco de dados. A tabela `usuarios` é populada com usuários estudantes ou responsáveis, a partir do cadastro de usuários feito no próprio *chatbot*. A inclusão de usuários administradores é feita de forma manual, diretamente no banco de dados.

A outra tabela é chamada de `questoes_validar` e armazena as perguntas que o ChatDuca não soube responder durante a interação com o usuário. Nesta tabela existe: (1) um campo chamado `pergunta`, que armazena exatamente a pergunta que o usuário fez ao Duca; (2) um campo chamado `resposta`, que armazena o assunto que o usuário vinculou à pergunta, sendo que se ele não vincular um assunto, este campo por padrão é `null`; (3) um campo chamado `sn_avalidado`, que armazena 0 ou 1 e indica se a pergunta já foi avaliada pelo administrador; (4) um campo com um código sequencial. A tabela `questoes_validar` é populada a partir da interação com o Duca. Quando o Duca não sabe responder sobre a pergunta que o usuário fez, ele envia uma mensagem ao usuário informando que não sabe responder o que foi perguntado e pergunta se o questionamento tem relação com notas, frequência ou boletos. Se o usuário informar que algum desses assuntos faz vínculo com o que ele perguntou, ele preenche essa tabela como exemplificado na Figura 10a, onde “Me fala sobre avaliações” tem, segundo o usuário, relação com “notas”. Caso o usuário informe ao Duca que a pergunta

dele não possui vínculo com nenhum assunto conhecido, ele preenche essa tabela como exemplificado na Figura 10b para a pergunta “O que é uma cadeira?”.

Figura 10 – Pergunta com e sem vínculo de tema

id	pergunta	resposta	sn_avalidado
11	ME FALE SOBRE AVALIAÇÕES	notas	0
15	O QUE É UMA CADEIRA?	(NULL)	0

(a)

id	pergunta	resposta	sn_avalidado
15	O QUE É UMA CADEIRA?	(NULL)	0

(b)

Fonte: elaborada pelo autor.

Ao serem inseridas no banco de dados, essas questões pendentes já são automaticamente mostradas para o administrador caso ele esteja logado no *chatterbot*. Quando o administrador exclui ou inclui uma das perguntas, o sistema altera o campo *sn_avalidado* para 1 e deixa de mostrá-la ao administrador.

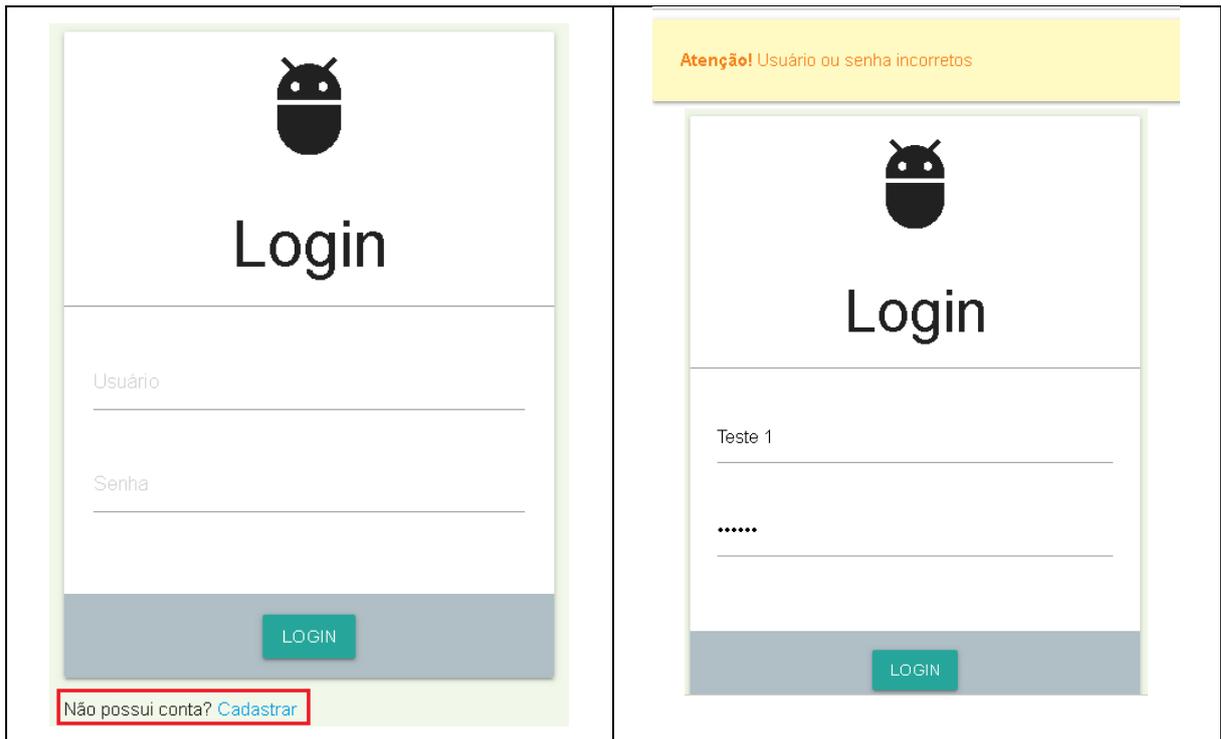
3.3.4 Operacionalidade da implementação

O ChatDuca foi desenvolvido como uma aplicação web e, por conta disso, caso esteja hospedado em algum servidor, basta apenas acessá-lo pelo navegador, não sendo necessário instalá-lo para ser utilizado. Inicialmente, a partir da tela mostrada na

Figura 11 (à esquerda), o usuário pode informar o seu login e senha já previamente cadastrados. Caso o usuário informe o login ou a senha de forma incorreta, é mostrada uma mensagem de erro (

Figura 11 à direita).

Figura 11 – Tela de login



Fonte: elaborada pelo autor.

Caso o usuário não tenha um login cadastrado no ChatDuca, pode clicar no link **Cadastrar** (em destaque na

Figura 11 à esquerda), para ser redirecionado para a tela de cadastro de usuários (Figura 12). Para efetuar o cadastro é necessário informar: o nome, o login (usuário), a senha e selecionar o papel (estudante ou responsável). Ao finalizar o cadastro, o usuário fica habilitado para utilizar o ChatDuca.

Figura 12 – Tela de cadastro de usuário


Cadastro de usuário

Nome

Usuário

Senha

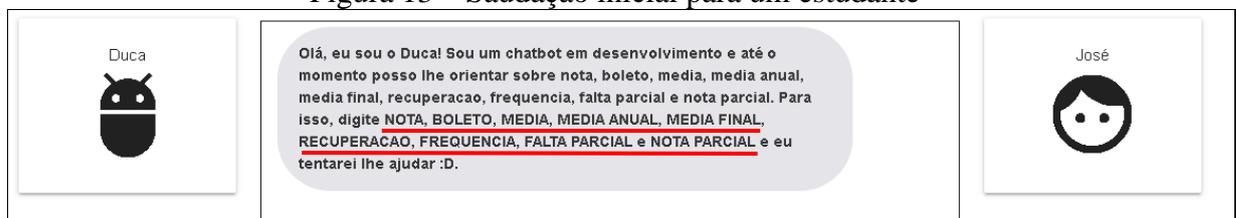
Estudante
 Responsável

CADASTRAR

Fonte: elaborada pelo autor.

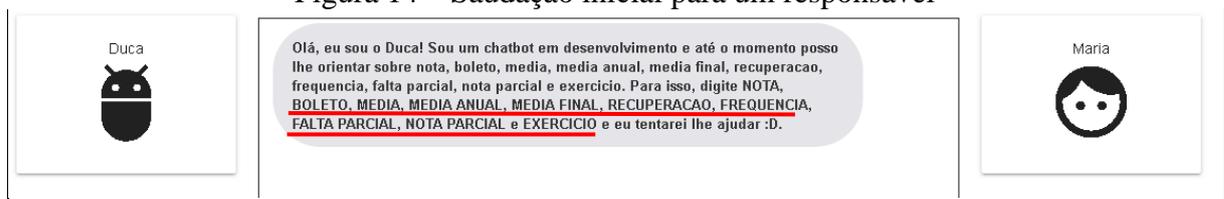
Quando o usuário informa um login e uma senha válidos, vinculados ao papel de estudante ou responsável, ele é redirecionado para a tela de conversação e inicia-se a interação com o ChatDuca (Figura 13 e Figura 14). Pode-se observar que, além da saudação inicial, o *chatbot* apresenta todos os assuntos sobre os quais ele sabe conversar, de acordo com o papel do usuário que está logado (na Figura 13, um estudante; na Figura 14, um responsável).

Figura 13 – Saudação inicial para um estudante



Fonte: elaborada pelo autor.

Figura 14 – Saudação inicial para um responsável

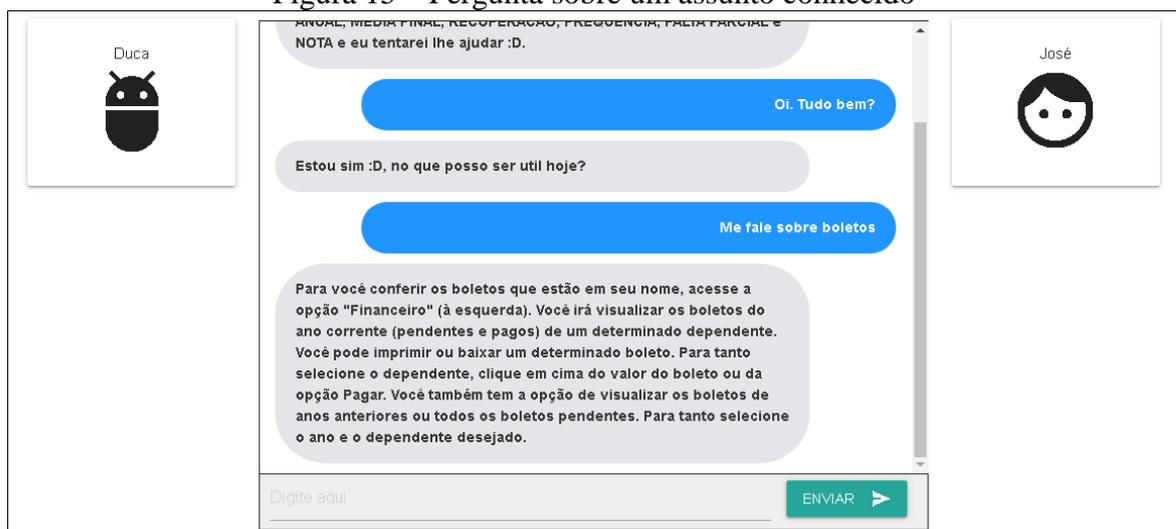


Fonte: elaborada pelo autor.

Na Figura 15, é mostrado o Duca respondendo à pergunta de um assunto que ele já conhece. Já na Figura 16 pode-se ver que foi questionado ao Duca sobre assuntos que são desconhecidos e ele reagiu questionando o usuário se o que ele quis dizer tem relação com notas, frequência ou boletos, ou seja, assuntos sobre os quais ele tem informação.

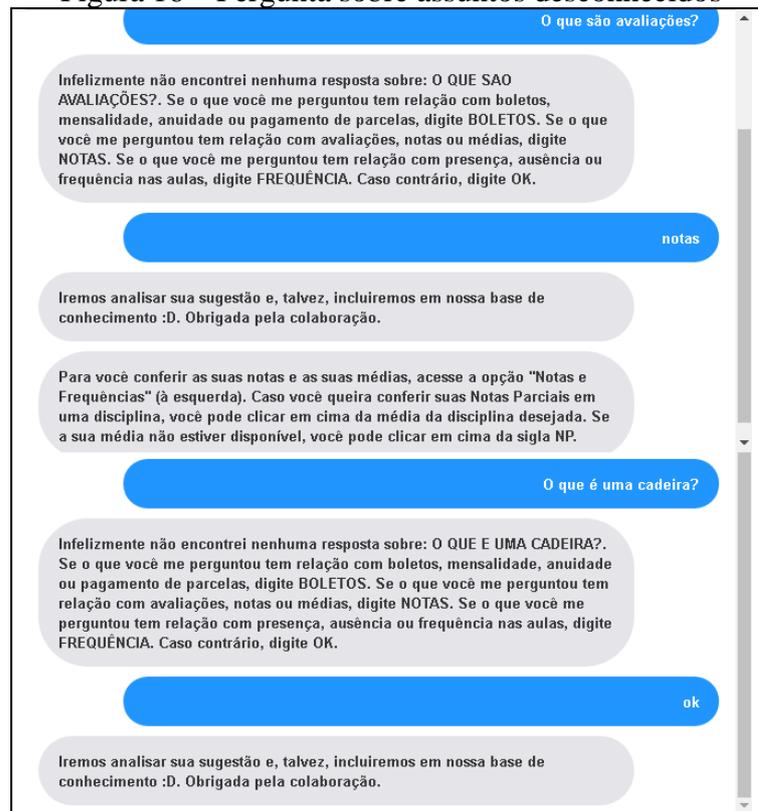
O Duca realiza ações diferentes, dependendo de como o usuário interage com ele. No primeiro exemplo da Figura 16, foi feito o relacionamento da pergunta “O que são avaliações?” com o assunto NOTAS, já que o usuário informou que existe relação entre esses assuntos, sendo mostrada a resposta associada ao assunto NOTAS. Já no segundo exemplo, pode ser visto que foi feita uma pergunta sobre “O que é uma cadeira?” e o usuário não identificou nenhuma relação com os assuntos conhecidos pelo Duca. Tanto no primeiro quanto no segundo caso, as sugestões são deixadas para análise posterior, a ser feita pelo administrador, para uma possível inclusão nas bases de conhecimento, sendo que no primeiro caso, tem-se um tema que pode ser vinculado a um assunto existente, enquanto no segundo caso o assunto é considerado novo.

Figura 15 – Pergunta sobre um assunto conhecido



Fonte: elaborada pelo autor.

Figura 16 – Pergunta sobre assuntos desconhecidos



Fonte: elaborada pelo autor.

Quando o login é feito por um usuário que tem o papel de administrador, ele é redirecionado para a tela de administração do *chatbot* (Figura 17). Lá ele pode visualizar todas as perguntas feitas que não têm uma resposta direta na base de conhecimentos. O administrador analisa a lista de pendências e decide se, a partir das sugestões dos usuários, irá ou não incrementar a base de conhecimento. O administrador visualiza exatamente o que o usuário questionou ao Duca e o vínculo que o usuário fez, como é o caso do tema “O que são avaliações?” vinculado ao assunto NOTAS e “O que é uma cadeira?”, considerando um assunto novo por não ter vínculo com nenhum outro assunto da base de conhecimento do *chatbot*. Ao clicar no botão incluir, representado pelo sinal +, ou no botão aceitar, representado pelo sinal *check*, as bases de conhecimento serão incrementadas. É importante destacar que o administrador tem total liberdade de alterar o tema sugerido, o assunto sugerido ou a respectiva resposta, para uma melhor adequação do conteúdo, podendo escolher se o assunto será incluído na base de conhecimento do estudante, do responsável ou de ambos. Uma vez incluídos em uma base de conhecimento os assuntos que estavam pendentes, o Duca passa a responder sobre eles, relacionando-os nas saudações iniciais em que é mostrado o novo assunto tratado (CADEIRA) e a respectiva resposta incluída na base de conhecimento (Figura 18a). Já a Figura 18b mostra que foi feita uma pergunta sobre avaliações e agora o Duca sabe responder.

Figura 17 – Lista de pendências do administrador

Lista de pendências

✓ Incluir na base do Estudante ✓ Incluir na base do Responsável

Assunto: notas ✓ ✕

Tema: O QUE SÃO AVALIAÇÕES? _____

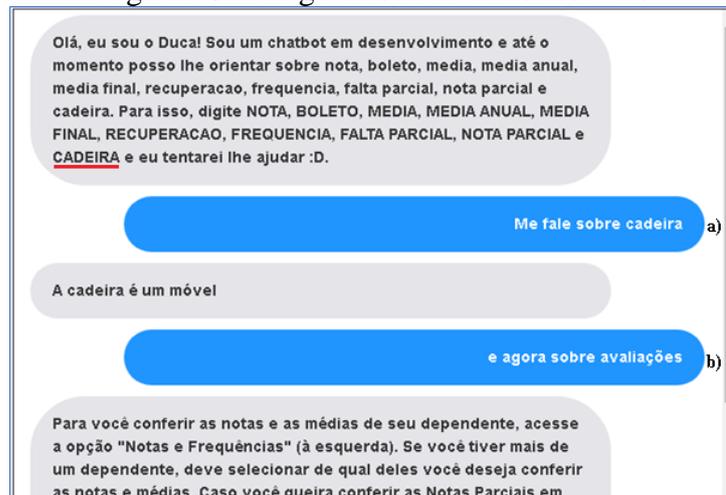
✓ Incluir na base do Estudante ✓ Incluir na base do Responsável

Assunto: O QUE É UMA CADEIRA? + ✕

Resposta: _____

Fonte: elaborada pelo autor.

Figura 18 – Pergunta sobre novos assuntos



Fonte: elaborada pelo autor.

3.4 RESULTADOS E DISCUSSÕES

No Quadro 7 **Erro! Fonte de referência não encontrada.** é apresentado um comparativo entre os trabalhos correlatos descritos e o ChatDuca.

Quadro 7 – Comparativo entre os trabalhos correlatos e o ChatDuca

	ChatterEDU (MARTINS, 2016)	Robô Andrew (OLIVEIRA, 2015)	chatbot (TRENTIN, 2016)	ChatDuca
características				
utiliza AIML	sim	sim	sim	sim
utiliza banco de dados	não	não	sim	sim
utiliza ontologia	não	sim	não	não
conversa sobre diversos assuntos	não	não	não	sim
gera automaticamente a base de conhecimento	sim	não	não	não
possibilita a complementação/manutenção da base de conhecimento	não	sim	não	sim
plataforma	desktop	web	web	web

Fonte: elaborado pelo autor.

No Quadro 7 **Erro! Fonte de referência não encontrada.** pode-se ver que a utilização de AIML é comum em todos os trabalhos, enquanto apenas Trentin (2016) e o ChatDuca utilizam banco de dados. Trentin (2016) usa banco de dados para armazenar conhecimento. No caso do ChatDuca, o banco de dados é utilizado tanto para fazer a autenticação dos usuários que têm permissão de acesso ao *chatterbot*, quanto para armazenar as perguntas que foram feitas e que não puderam ser respondidas por falta de conhecimento do *chatterbot*. Já de forma exclusiva, somente Oliveira (2015) usa ontologia para armazenar conhecimento.

Nenhum dos *chatterbots* descritos, com exceção do ChatDuca, conversa sobre diversos assuntos, embora o ChatterEDU tenha sua base de conhecimento gerada automaticamente a partir de um texto de entrada. No entanto, a cada novo texto analisado, a base de conhecimentos do ChatterEDU é sobrescrita. Para conversar sobre diversos assuntos, o ChatDuca faz uso da interação com usuário. Sempre que o ChatDuca não encontra uma resposta, o usuário pode informar se existe algum assunto relacionado à questão que ele perguntou ou se é um assunto novo e, então, o administrador deve validar a sugestão do usuário, incrementando as bases de conhecimento. Observa-se que o trabalho do Oliveira (2015) também possui funcionalidade para incrementar a base de conhecimento, mas de forma manual pelo administrador do *chatterbot* através de uma área administrativa, desde que seja sobre as maneiras de ingressar em cursos na FURB, único assunto tratado pelo Robô Andrew. Ressalta-se que nenhum dos trabalhos permite a complementação da base de conhecimento de forma automática. Considera-se essa funcionalidade o principal fator de melhoria de *chatterbots*, pois por meio dela o *chatterbot* conseguiria responder sobre mais assuntos de forma mais precisa. Neste formato, o *chatterbot* poderia, por exemplo, aprender a partir da interação com o usuário, armazenando e utilizando a nova informação com outro usuário, sem a interação de um humano neste processo.

Trentin (2016), Oliveira (2015) e o ChatDuca foram desenvolvidos para plataforma web visando a facilidade de utilização, interfaces mais amigáveis e uma melhor usabilidade por parte dos usuários. Importante ressaltar que por conta disso não é necessária a instalação dos softwares no computador do usuário. Já Martins (2016) optou por desenvolver um sistema desktop utilizando Java. Ainda, o ChatDuca foi desenvolvido para dar suporte a usuários de um software de gestão educacional. Visa diminuir o suporte dado aos usuários pela equipe de Tecnologia da Informação e pelas demais pessoas que compõem a equipe de suporte de instituições de ensino que usam esse software. Como a base de conhecimento é complementada a partir da interação com o usuário e posterior validação do administrador, o *chatterbot* pode responder perguntas sobre diversos assuntos relacionados com o software de gestão

educacional, além dos inicialmente relacionados, quais sejam notas, frequência e boletos, sendo, assim, um diferencial sobre os trabalhos expostos anteriormente.

Foram realizados testes de simulação de conversação com dúvidas sobre um software de gestão educacional e foi possível obter a resposta desejada. Também foi realizado um teste com um usuário do sexo feminino, que não possuía conhecimento nem do software nem do *chatbot*, mas com experiência no uso de computadores. O teste foi feito localmente, pois o ChatDuca não está hospedado em nenhum servidor externo. O teste demorou cerca de um minuto e durante o teste o usuário questionou como fazer login na aplicação. Foi orientado como devia proceder para cadastrar um novo usuário e foi solicitado para imprimir um boleto no software de gestão educacional. Dali para frente o teste foi feito de forma totalmente autônoma. O resultado foi positivo, pois o usuário conseguiu sanar suas dúvidas e encontrar a funcionalidade que estava procurando dentro do software de gestão educacional a partir da interação com o *chatbot*. De forma mais detalhada, esse teste é apresentado no Apêndice B.

4 CONCLUSÕES

Este projeto teve como objetivo construir um *chatbot* que conseguisse conversar na língua portuguesa e sanar dúvidas dos usuários sobre um software de gestão educacional, com a possibilidade de complementar a base de conhecimento do *chatbot* a partir da interação com o usuário.

Para tanto, para armazenar as credenciais dos usuários que têm permissão de acessar o chat foi utilizado um banco de dados e para armazenar o conhecimento do *chatbot* foram utilizadas bases de conhecimento estruturadas com o padrão AIML. As bases de conhecimento foram construídas manualmente utilizando a língua portuguesa e, com isso, o *chatbot* só se comunica em português. Se for desejado que o *chatbot* se comunique em outro idioma, é necessário reescrever as bases de conhecimento. Além disso, sempre que o usuário faz uma pergunta que o *chatbot* não sabe responder, ela é incluída no banco de dados para posterior análise pelo administrador do ChatDuca. Somente após a validação do administrador, é que as bases de conhecimento são incrementadas.

Ao utilizar a linguagem AIML, foram encontradas algumas dificuldades. Uma delas é a limitação da própria linguagem, que por ser estruturada com base em XML, não é possível fazer uma análise aprofundada dos dados, somente *links* e vínculos de sentenças. Outra limitação da AIML é que as perguntas devem ser escritas pelo usuário do *chatbot* exatamente como relacionadas nas bases. Dessa forma, erros na escrita das perguntas ou palavras sinônimas não são tratados. Nesse último caso, é necessário que todos sinônimos sejam incluídos na base de conhecimento para que o *chatbot* faça a interpretação corretamente. Em função disso, estruturou-se as bases de conhecimento com assuntos e temas.

Por fim, considera-se que o desenvolvimento de um *chatbot*, como opção para suporte técnico de softwares, no caso desse trabalho de um software de gestão educacional, apresenta um ganho para empresas e usuários. O *chatbot*, ao conseguir responder às perguntas dos usuários, dispensa a necessidade de suporte por parte da equipe de Tecnologia da Informação da instituição de ensino ou da empresa responsável por manter o software. Além disso, o usuário terá sua dúvida respondida de forma mais rápida, assertiva e até mesmo disponibilizando *links*, se assim for necessário. Dessa forma, pode-se dizer que existe uma redução de custos envolvida com o desenvolvimento/uso de *chatbots*.

4.1 EXTENSÕES

O projeto apresentado neste trabalho atingiu os objetivos propostos. Mesmo assim, existem pontos que podem ser melhorados e novas funcionalidades que podem ser implementadas. São eles:

- a) formular uma rotina de busca de assuntos para completar o preenchimento das questões do usuário, sugerindo o que pode ser perguntado;
- b) realizar testes com usuários do software de gestão educacional para validar o *chatbot*;
- c) incluir funcionalidade para permitir que todos os assuntos das bases de conhecimento possam ser alterados;
- d) permitir que o administrador possa incluir um assunto novo, sem precisar que o usuário tenha tido contato com o *chatbot*;
- e) incluir funcionalidade para permitir que o administrador possa incluir outros administradores;
- f) permitir que o administrador possa conversar com o ChatDuca;
- g) reformular a organização das bases de conhecimento para permitir que o usuário possa fazer duas perguntas diferentes com respostas também diferentes sobre um mesmo assunto, como por exemplo “Onde encontro minhas mensalidades?” ou “Qual é o valor das minhas mensalidades?”.

REFERÊNCIAS

- BRANSKI, Regina M. Recuperação de informações na web. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 9, n. 1, p. 70-87, jan./jun. 2004. Disponível em: http://www.brapci.inf.br/_repositorio/2010/11/pdf_7b0e618ad3_0012984.pdf. Acesso em: 19 ago. 2018.
- COMARELLA, Rafaela L.; CAFÉ, Ligia M. A. **Chatterbot: conceito, características, tipologia e construção**. *Informação & Sociedade: estudos*, João Pessoa, n. 2, p. 55-67, 2008. Disponível em: <http://www.periodicos.ufpb.br/index.php/ies/article/view/1758/2110>. Acesso em: 15 set. 2018.
- MACHADO, André. **Estudo da EMC prevê que volume de dados virtuais armazenados será seis vezes maior em 2020**. São Paulo, 2014. Disponível em: <https://oglobo.globo.com/sociedade/tecnologia/estudo-da-emc-preve-que-volume-de-dados-virtuais-armazenados-sera-seis-vezes-maior-em-2020-12147682>. Acesso em: 09 set. 2018.
- MARTINS, Camila V. **Ferramenta de auxílio acadêmico utilizando chatterbot**. 2016. 51 f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade Regional de Blumenau, Blumenau.
- MAULDIN, Michael L. Chatterbots, tinymuds, and the turing test entering the loebner prize competition. In: THE AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 12., 1994, Pittsburgh. **Proceedings...** Seattle: The MIT Press, 1994. p. 16-21. Disponível em: <http://new.aaai.org/Papers/AAAI/1994/AAAI94-003.pdf>. Acesso em: 14 jul. 2019.
- MOURA, Thiago J. M. **Um chatterbot para aquisição automática de perfil do usuário**. 2003. 127 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife. Disponível em: <http://www.liber.ufpe.br/teses/arquivo/20050228150106.pdf>. Acesso em: 15 set. 2018.
- OLIVEIRA, Bruno. **Chatterbot para esclarecimento de dúvidas sobre as formas de ingresso em cursos da FURB**. 2015. 66 f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade Regional de Blumenau, Blumenau.
- SENIOR SISTEMAS. **CRM da Senior ganha assistente virtual**. Blumenau, 2017. Disponível em: <https://www.senior.com.br/noticias/sara-crm>. Acesso em: 19 set. 2018.
- SILVA, Gleidistone. **Quem é a BIA? Conheça a inteligência artificial do Bradesco**. São Paulo, 2018. Disponível em: <https://inteligencia.rockcontent.com/bia-bradesco>. Acesso em: 19 set. 2018.
- TEIXEIRA, Sérgio; MENEZES, Crediné S. Facilitando o uso de ambientes virtuais através de agentes de conversação. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 14., 2003, Rio de Janeiro. **Anais eletrônicos...** Rio de Janeiro: SBC/UFRJ, 2003. Não paginado. Disponível em: <http://www.nce.ufrj.br/sbie2003/publicacoes/paper48.pdf>. Acesso em: 19 ago. 2018.
- TRENTIN, Carlos E. **Chatterbot para auxiliar o usuário no atendimento ao processo de atualização de software**. 2016. 60 f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade Regional de Blumenau, Blumenau.
- WALLACE, Richard. **Official Alicebot AIML Wiki**. [S.l.], 2018. Disponível em: <http://alicebot.wikidot.com/>. Acesso em: 14 jul. 2019.

WEIZENBAUM, Joseph. ELIZA: a computer program for the study of natural communications between man and machine. **Communications of the ACM**, Cambridge, v. 9, n. 1, p. 36-45, Jan. 1966. Disponível em: <http://web.stanford.edu/class/linguist238/p36-weizenbaum.pdf>. Acesso em: 14 jul. 2019.

APÊNDICE A – Descrição dos Casos de Uso

A seguir são detalhados os quatro casos de uso deste projeto: UC01: Cadastrar usuário (Quadro 8), UC02: Fazer login (Quadro 9), UC03: Dialogar com o ChatDuca (Quadro 10) e UC04: Manter as bases de conhecimento (Quadro 11).

Quadro 8 – UC01: Cadastrar usuário

UC01 – Cadastrar usuário	
Descrição	Permitir que o usuário efetue um cadastro para utilização do ChatDuca.
Autor	Usuário
Cenário Principal	1. Usuário preenche nome, login (usuário), senha e seleciona o seu perfil (estudante ou responsável).
Pré-condições	Conexão com o banco de dados.
Pós-condições	Usuário cadastrado.

Fonte: elaborado pelo autor.

Quadro 9 – UC02: Fazer login

UC02 – Fazer login	
Descrição	Permitir que o usuário faça login no ChatDuca.
Autor	Usuário
Cenário Principal	1. Usuário preenche login e senha e clica em acessar. 2. Se o usuário tiver o papel de administrador, é redirecionamento para área de administração. 3. Se o usuário tiver o papel de estudante ou responsável, é redirecionamento para a área de conversação.
Pré-condições	Conexão com o banco de dados. Usuário cadastrado.
Pós-condições	Usuário logado.

Fonte: elaborado pelo autor.

Quadro 10 – UC03: Dialogar com o ChatDuca

UC03 – Dialogar com o ChatDuca	
Descrição	Permitir que o usuário dialogue com o ChatDuca.
Autor	Usuário
Cenário Principal	1. O usuário faz uma pergunta sobre o algum dos assuntos conhecidos do ChatDuca. 2. ChatDuca apresenta a resposta para a pergunta feita.
Fluxo alternativo 01	1. O usuário faz uma pergunta ao ChatDuca sobre um assunto que ele não tem conhecimento. 2. ChatDuca informa que não sabe responder o que foi perguntado e pede para o usuário informar se o que ele perguntou tem ou não relação com notas, frequência ou boletos. 3. O usuário responde ao ChatDuca. 3.1. Se o usuário informar que o que ele perguntou tem relação com os assuntos conhecidos, o ChatDuca insere a pergunta no painel de administração (como um tema) e apresenta a resposta correspondente ao assunto relacionado. 3.2. Se o usuário informar que o que ele perguntou não tem relação com os assuntos conhecidos, o ChatDuca insere a pergunta no painel de administração (como um assunto novo).
Pré-condições	Usuário logado como responsável ou como estudante. Bases de conhecimento, do estudante e do responsável, com assuntos já pré-estabelecidos.
Pós-condições	Histórico da conversação.

Fonte: elaborado pelo autor.

Quadro 11 – UC04: Manter as bases de conhecimento

UC04 – Manter as bases de conhecimento	
Descrição	Permitir que o usuário insira novos conteúdos nas bases de conhecimento.
Autor	Administrador
Cenário Principal	<ol style="list-style-type: none"> 1. Administrador avalia assunto, temas e respostas da lista de pendências. 2. Administrador refina o texto escrito pelo usuário, se for o caso. 3. Administrador insere o conteúdo nas bases de conhecimento desejadas (do responsável, do estudante ou de ambos).
Pré-condições	<p>Usuário logado como administrador.</p> <p>Bases de conhecimento, do estudante e do responsável, com assuntos já pré-estabelecidos.</p> <p>Itens pendentes para análise, a partir do Fluxo Alternativo do UC03.</p>
Pós-condições	Bases de conhecimento atualizadas.

Fonte: elaborado pelo autor.

APÊNDICE B – Teste do *chatbot*

Este apêndice apresenta o resultado do teste realizado com um usuário do sexo feminino sem conhecimento do software de gestão educacional, mas com experiência no uso de computadores. O teste foi realizado e, ao final, foram feitas as perguntas do **Erro! Fonte de referência não encontrada.** a fim de obter a experiência do usuário na conversação com o *chatbot*. O **Erro! Fonte de referência não encontrada.** também apresentada as respostas dadas pelo usuário, enquanto a Figura 19 mostra a conversação do usuário com o ChatDuca.

Quadro 12 – Avaliação do ChatDuca pelo usuário

1ª Você conseguiu dialogar com o Chatterbot? R: Sim
2ª Qual foi sua dúvida? R: Sobre cadeira e recuperação.
3ª Você acredita que o ChatDuca pode sanar as dúvidas do usuário e orientá-lo de forma mais rápida e precisa? R: Com certeza, tendo em vista que otimiza o tempo, pois não depende de realizar ligações ou ler algum manual para tirar a dúvida.

Fonte: elaborado pelo autor.

Figura 19 – Conversação do usuário com o ChatDuca

Olá, eu sou o Duca! Sou um chatbot em desenvolvimento e até o momento posso lhe orientar sobre nota, boleto, media, media anual, media final, recuperacao, frequencia, falta parcial, nota parcial e cadeira. Para isso, digite NOTA, BOLETO, MEDIA, MEDIA ANUAL, MEDIA FINAL, RECUPERACAO, FREQUENCIA, FALTA PARCIAL, NOTA PARCIAL e CADEIRA e eu tentarei lhe ajudar :D.

cadeira

A cadeira é um móvel

recuperação

Infelizmente não encontrei nenhuma resposta sobre: RECUPERAÇÃO. Se o que você me perguntou tem relação com boletos, mensalidade, anuidade ou pagamento de parcelas, digite BOLETOS. Se o que você me perguntou tem relação com avaliações, notas ou médias, digite NOTAS. Se o que você me perguntou tem relação com presença, ausência ou frequência nas aulas, digite FREQUÊNCIA. Caso contrário, digite OK.

recuperacao

Iremos analisar sua sugestão e, talvez, incluiremos em nossa base de conhecimento :D. Obrigada pela colaboração.

A recuperação trimestral (REC) é a nota que você obteve na atividade de recuperação de uma disciplina em um trimestre. Pode ser uma atividade no AVA que acrescenta pontos na média, pode ser uma prova cuja nota substituirá a menor nota entre as provas do trimestre, entre outras atividades possíveis, dependendo do ano em que você está matriculado.

Fonte: elaborada pelo autor.