



**CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS DE LINGUAGENS**

Leonardo Drummond Vilaça Lima Camargo

**CRITÉRIOS PARA AVALIAÇÃO DE *CHATBOTS* DIDÁTICOS**

Belo Horizonte  
2022

Leonardo Drummond Vilaça Lima Camargo

## **CRITÉRIOS PARA AVALIAÇÃO DE *CHATBOTS* DIDÁTICOS**

Tese de doutorado apresentada como requisito parcial para a conclusão do programa de doutorado em Estudos de Linguagens do Centro Federal de Educação Tecnológica de Minas Gerais.

Orientador: Prof. Dr. Vicente Aguiar Parreiras  
Coorientadora: Prof. Dra. Terhi Katriina Skaniakos

Belo Horizonte  
2022

C173c Camargo, Leonardo Drummond Vilaça Lima.  
Critérios para avaliação de chatbots didáticos / Leonardo  
Drummond Vilaça Lima Camargo. – 2022.  
183 f. : il.

Orientador: Vicente Aguiar Parreiras  
Coorientadora: Terhi Katriina Skaniakos

Tese (Doutorado) – Centro Federal de Educação Tecnológica de  
Minas Gerais, Programa de Pós-Graduação em Estudos de  
Linguagens, Belo Horizonte, 2022.  
Bibliografia.

1. Tecnologia educacional. 2. Inteligência artificial - Aplicações  
educacionais. 3. Material didático - Avaliação. 4. Processamento de  
Linguagem Natural. I. Parreiras, Vicente Aguiar. II. Skaniakos, Terhi  
Katriina. III. Título.

CDD: 006.35

Ficha elaborada pela Biblioteca - Campus I – CEFET-MG  
Bibliotecário: Wagner Oliveira Braga CRB6 - 3261



**CEFET-MG**

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS  
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM ESTUDOS DE LINGUAGENS

**Leonardo Drummond Lima Camargo**

## **CRITÉRIOS PARA AVALIAÇÃO DE CHATBOTS DIDÁTICOS**

Tese apresentada ao Programa de Pós-Graduação em Estudos de Linguagens do Centro Federal de Educação Tecnológica de Minas Gerais em 09 de maio de 2022, como requisito parcial para obtenção do título de Doutor em Estudos de Linguagens, aprovada pela Banca Examinadora constituída pelos professores:

Prof. Vicente Aguiar Parreiras, Dr. CEFET-MG - Orientador  
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Edson Martins Júnior, Dr. - USJT

Prof. Bruno Alves Rodrigues, Dr. TRT

Prof<sup>a</sup>. Tehri Skaniakos, Dr<sup>a</sup>. Universidade de Jyvaskyla - Finlândia

Prof. Renato Caixeta da Silva, Dr. CEFET-MG

## AGRADECIMENTOS

À minha mãe, Maria Cecília, maior incentivadora para que eu me inscrevesse neste programa de doutorado. Ela pôde compartilhar a alegria da minha entrada no programa, mas apesar de fisicamente não poder acompanhar a finalização deste processo, sinto sua felicidade e orgulho comigo também nesta etapa de conclusão.

Agradeço também a minha família, amigos e todas as pessoas que fizeram este processo mais leve, alegre e produtivo, especialmente ao meu marido Arthur.

Meu orientador Vicente Parreiras e minha co-orientadora Terhi Skaniakos, obrigado por tanto apoio, carinho e contribuições que tornaram possível esta pesquisa. Trabalhar com vocês é um presente que recebi da vida.

Aos professores e colegas do Posling - CEFET-MG, por diálogos construtivos e aprendizagens tão relevantes na minha construção profissional e pessoal. Aos colegas da Universidade de Jyväskylä por me acolherem tão bem e proporcionarem uma experiência profissional e de vida tão marcante.

À banca de avaliação, agradeço antecipadamente por todas as contribuições. Jerônimo Coura-Sobrinho e Renato Caixeta, as suas contribuições na banca de qualificação foram essenciais para o resultado desta tese.

À Samira Spolidorio, grande amiga e companheira na produção do *chatbot* e no apoio de tradução e revisão desta tese.

Aos colegas que trabalharam na produção de Katri e aos alunos do curso Teaching and Learning in Higher Education. À equipe de gestão da área em que o curso foi realizado, especialmente Inês Barreto, Rafael Ávila, Rogério Loureiro, Edson Martins e Clara Teixeira pelo companheirismo, apoio e pelos instigantes projetos que desenvolvemos juntos.

Às professoras Terhi, Milla, Kaisu e Johanna: vocês são grandes amigas e inspirações para mim, além de excelentes colegas de jornada pedagógica.

Agradeço também à CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela concessão de bolsa de pesquisa no Programa de Doutorado Sanduíche no Exterior<sup>1</sup>, por meio da qual pude fazer meu estágio de doutoramento na Finlândia.

---

<sup>1</sup> O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001"/"This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001"

## RESUMO

Esta pesquisa discute a produção e a avaliação de chatbots usados para fins de ensino e aprendizagem. Esse tópico foi pesquisado com o objetivo de criar e validar o uso de uma matriz com critérios e descritores para avaliação de *chatbots* didáticos. A pesquisa apresenta um mapeamento de critérios para produção e avaliação desse tipo de material, tendo como base critérios para avaliação de materiais didáticos de outras naturezas e de critérios para avaliação de *chatbots* usados para fins não didáticos. Também se apresenta o estudo de caso do chatbot Katri, desenvolvido durante esta pesquisa para uso em um curso finlandês de formação docente. A matriz de avaliação de chatbots didáticos foi utilizada para avaliar esse chatbot, considerando as percepções de produtores do sistema e de alunos brasileiros desse curso. A análise dos dados coletados foi baseada no Sistema de Avaliabilidade proposto por Martin e White (2005). Como resultados da pesquisa, validou-se o uso da matriz, composta por 5 categorias e 30 descritores, por meio da qual foi possível sugerir melhorias e revisões no *chatbot* estudado.

**Palavras-chave:** *Chatbot*, *Chatbot* educacional, Avaliação de Material Didático, Inteligência Artificial na Educação, Processamento de Linguagem Natural.

## ABSTRACT

This research discusses the production and evaluation of chatbots used for teaching and learning purposes. This topic was researched with the objective of creating and validating the use of a matrix with criteria and descriptors for the evaluation of didactic chatbots. The research presents a mapping of criteria for the production and evaluation of this type of material, based on criteria for evaluating didactic materials of other natures and criteria for evaluating chatbots used for non-didactic purposes. This thesis also presents the case study of the chatbot Katri, developed during this research for a Finnish teacher training course. The evaluation matrix of didactic chatbots was used to evaluate this chatbot, considering the perceptions of the system's producers and the Brazilian students of this course. The analysis of the collected data was based on the Appraisal System proposed by Martin and White (2005). As a result of this research, the use of the matrix, composed by 5 categories and 30 descriptors, was validated, through which it was possible to suggest improvements and revisions in the studied chatbot.

**Keywords:** *Chatbot*, Educational *Chatbot*, Didactic Material Evaluation, Artificial Intelligence in Education, Natural Language Processing.

## LISTA DE FIGURAS

FIGURA 1 - Tipos de MD segundo Graells.....	15
FIGURA 2 - Tipos de Materiais Didáticos (proposta atualizada).....	16
FIGURA 3 - Aprendizagem por <i>chatbots</i> educacionais.....	22
FIGURA 4 - Teste de interação com Eliza.....	31
FIGURA 5 - Diálogos de estudantes com o Jill Watson.....	40
FIGURA 6 - Exemplos de troca de mensagens em inglês e árabe no SIAAA-C.....	43
FIGURA 7 - <i>Chatbot</i> envia conteúdos por texto escrito e vídeo aos alunos em escola turca.....	44
FIGURA 8 - Fluxograma de avaliação de materiais didáticos.....	48
FIGURA 9 - Sistema de Avaliatividade.....	64
FIGURA 10 - Exemplos de expressões textuais que demonstram Afeto.....	66
FIGURA 11 - Exemplos de expressões textuais que demonstram afetos positivos e negativos.....	67
FIGURA 12 - Tipos de Apreciação.....	68
FIGURA 13 - Subsistema de Gradação.....	70
FIGURA 14 - Semana de Design Sprint do <i>Chatbot</i> .....	86
FIGURA 15 - Exemplos de temas tratados pelo <i>chatbot</i> .....	87
FIGURA 16 - Identidade de Katri.....	88
FIGURA 17 - Fluxo para representar a hierarquia de informação do <i>chatbot</i> .....	89
FIGURA 18 - Piloto da página inicial do <i>chatbot</i> .....	90
FIGURA 19 - Tela do piloto de Katri respondendo à questão Para Fazer Nessa Semana.....	90
FIGURA 20 - Nova imagem de Katri.....	103
FIGURA 21 - Trecho de diálogo com orientação sobre como voltar ao menu inicial.....	106
FIGURA 22 - Introdução a Katri no Moodle.....	109
FIGURA 23 - Menu inicial e primeira mensagem sobre <i>Pedagogical Project</i> .....	110
FIGURA 24 - Continuidade de diálogo sobre o <i>Pedagogical Project</i> .....	111
FIGURA 25 - Exemplos de conteúdos criados para treinar a Katri.....	117
FIGURA 26 - Katri chama o usuário pelo nome “Léo”, fornecido por ele na primeira interação.....	127
FIGURA 27 - Versão web de Katri.....	138
FIGURA 28 - Versão mobile de Katri.....	138
FIGURA 29 - Katri fornece link e material para download.....	140
FIGURA 30 - Resposta a questões desconhecidas.....	149
FIGURA 31 - Pedido de feedback.....	151

## LISTA DE QUADROS

QUADRO 1 - Estrutura do curso .....	18
QUADRO 2 - Instrumento para avaliação de coleções didáticas de língua inglesa.....	50
QUADRO 3- Critérios para produção e avaliação de <i>chatbots</i> usados como material didático .....	57
QUADRO 4 - Etapa 1: versão piloto de Katri.....	75
QUADRO 5 - Etapa 2: Versão definitiva de Katri.....	77
QUADRO 6 - Atitudes e Gradação na Avaliação de Katri .....	79
QUADRO 7 - Exemplo de preenchimento do quadro de descritores para avaliação.....	81
QUADRO 8 - Exemplo de preenchimento da matriz de avaliação.....	82
QUADRO 9 - Descritores analisados no piloto .....	91
QUADRO 10 - Categoria: Aparência e persona - piloto de Katri.....	92
QUADRO 11 - Afeto e Apreciação: menu inicial .....	93
QUADRO 12 - Afeto e Apreciação: conhecimentos especializados .....	94
QUADRO 13 - Categoria: Discurso e linguagens - piloto de Katri.....	96
QUADRO 14 - Afeto e Apreciação: Clareza .....	97
QUADRO 15 - Afeto e Apreciação: Suficiência de Informações.....	98
QUADRO 16 - Categoria: Aparência e persona – piloto de Katri.....	100
QUADRO 17 - Afeto e Apreciação: Persona do <i>chatbot</i> .....	101
QUADRO 18 - Afeto e Apreciação: Papel representado por Katri.....	102
QUADRO 19 - Categoria: Navegação e Usabilidade - piloto de Katri.....	105
QUADRO 20 - Categoria: Funcionalidades adicionais - piloto de Katri.....	105
QUADRO 21 - Matriz de avaliação da versão piloto de Katri.....	107
QUADRO 22 - Síntese das interações entre usuários e Katri .....	112
QUADRO 23 - Número de interações por aluno .....	115
QUADRO 24 - Avaliação da categoria Conteúdos.....	119
QUADRO 25 - Avaliação do descritor Conteúdos Especializados.....	121
QUADRO 26 - Avaliação do descritor Clareza .....	123
QUADRO 27 - Avaliação da categoria Discurso e Linguagens .....	124
QUADRO 28 - Avaliação do descritor Clareza .....	126
QUADRO 29 - Avaliação do descritor Suficiência de informações .....	129
QUADRO 30 - Avaliação do descritor Adequação do vocabulário.....	131
QUADRO 31 - Avaliação do descritor Formalidade .....	132
QUADRO 32 - Avaliação do Categoria Navegação e Usabilidade .....	133
QUADRO 33 - Avaliação do descritor Layout.....	135
QUADRO 34 - Avaliação do descritor Adaptação/individualização do processo .....	136
QUADRO 35 - Avaliação do descritor Facilidade de Uso do Sistema .....	141
QUADRO 36 - Avaliação do Categoria Aparência e Persona .....	142
QUADRO 37 - Avaliação do descritor Persona do <i>chatbot</i> .....	143
QUADRO 38 - Avaliação do descritor Papel desempenhado pelo <i>chatbot</i> .....	145
QUADRO 39 - Avaliação do Categoria Funcionalidades Adicionais .....	147
QUADRO 40 - Avaliação Completa da Matriz .....	152

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>11</b>
<b>1.1 CONTEXTUALIZAÇÃO DO ESTUDO .....</b>	<b>12</b>
<b>1.2 O objeto do estudo de caso.....</b>	<b>16</b>
<b>1.3 Problema e Objetivos de Pesquisa .....</b>	<b>19</b>
<b>1.4 Justificativa .....</b>	<b>21</b>
<b>2. REFERENCIAL TEÓRICO.....</b>	<b>26</b>
<b>2.1 Inteligência Artificial e Processamento de Linguagem Natural na Educação.....</b>	<b>26</b>
<b>2.2 Breve histórico da evolução dos <i>chatbots</i>.....</b>	<b>29</b>
<b>2.3 Desambiguação no Processamento de Linguagem Natural por <i>Chatbots</i> .....</b>	<b>33</b>
2.3.1 Análise de intenções .....	34
2.3.2 Análise de contexto.....	36
2.3.3 Resolução de correferência .....	37
<b>2.4 <i>Chatbots</i> usados na educação .....</b>	<b>38</b>
<b>2.5 <i>Chatbots</i> didáticos no contexto da pandemia de Covid-19.....</b>	<b>42</b>
<b>3. CHATBOT DIDÁTICO: CRITÉRIOS PARA AVALIAÇÃO .....</b>	<b>46</b>
<b>3.1 Critérios para Avaliação de Materiais Didáticos.....</b>	<b>46</b>
<b>3.2 Critérios para Produção e Avaliação de <i>Chatbots</i> .....</b>	<b>52</b>
Critério 1: Aparência visual do chatbot .....	52
Critério 2: Forma de implementação no website .....	53
Critério 3: Unidade de produção de fala .....	53
Critério 4: Base de conhecimento .....	53
Critério 5: Apresentação de conhecimentos e funcionalidades adicionais .....	54
Critério 6: Habilidades conversacionais, habilidades de linguagens e sensibilidade ao contexto.....	54
Critério 7: Traços de personalidade .....	54
Critério 8: Opções de personalização .....	54
Critério 9: Respostas de emergência em situações inesperadas.....	55
Critério 10: Possibilidade de atribuição de notas pelo usuário ao chatbot .....	55
<b>3.3 Proposta de matriz para avaliação de <i>chatbots</i> didáticos .....</b>	<b>56</b>
<b>4. LINGUAGEM E AVALIAÇÃO .....</b>	<b>61</b>
<b>4.1 Sistema de Avaliatividade.....</b>	<b>63</b>
4.1.1 Atitude: Afeto e Apreciação .....	65

4.1.2 Gradação: foco e força .....	69
<b>5. METODOLOGIA .....</b>	<b>72</b>
<b>5.1 Enquadramento metodológico .....</b>	<b>72</b>
<b>5.2 Etapas da pesquisa, participantes e procedimentos .....</b>	<b>74</b>
<b>5.3 Análise dos dados.....</b>	<b>79</b>
<b>5.4 - Considerações sobre ética na pesquisa .....</b>	<b>82</b>
<b>6. RESULTADOS .....</b>	<b>84</b>
<b>6.1 Piloto: primeira versão de Katri .....</b>	<b>84</b>
6.1.1 Teste e avaliação da versão piloto de Katri .....	91
<b>6.2 Resultados: versão definitiva de Katri .....</b>	<b>112</b>
6.2.1 Dados de interação .....	112
6.2.2 Avaliação da versão definitiva de Katri: análise e discussão .....	116
<b>7. CONSIDERAÇÕES FINAIS .....</b>	<b>159</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>162</b>
<b>APÊNDICES .....</b>	<b>174</b>
<b>Apêndice 1 .....</b>	<b>174</b>
<b>Apêndice 2 .....</b>	<b>178</b>
<b>Apêndice 3 .....</b>	<b>179</b>
<b>Apêndice 4 .....</b>	<b>183</b>

## 1. INTRODUÇÃO

Os *chatbots* têm sido constantemente utilizados no comércio, principalmente no atendimento inicial a clientes, mas já começam a aparecer as primeiras experiências na educação, em instituições como Georgia Institute of Technology, nos Estados Unidos (RAHMAN; SUGUNA, 2017); Prepathon, comunidade de educação da Índia (JOHNSON, 2016); e Universidade Federal de Minas Gerais (LIMA, 2014). É neste contexto de novas tendências em tecnologias educacionais que se insere esta pesquisa sobre *chatbots*, o que pode contribuir para que a comunidade acadêmica entenda e se aproprie desta tecnologia que utiliza inteligência artificial (IA)<sup>2</sup>.

A produção de Materiais Didáticos (MD) utilizando IA é prática recente, mas cada vez mais pode-se encontrar esse tipo de produção. É, segundo Vicari (2018), uma tendência no campo da educação. Segundo Casagrande (2019), instituições de ensino apostam nessa tecnologia e os professores e alunos já começam a utilizá-las para apoio aos processos de ensino e aprendizagem. Ainda segundo o autor, entre os exemplos de tecnologias que utilizam IA que começam a ser usados no apoio a cursos de diferentes instituições estão os *chatbots*, agentes de conversação que tentam reproduzir a cognição humana para que possam aprender e recuperar informações em situações de interação. Mas como se dá o processo de avaliação desses *chatbots* usados especificamente para apoio aos estudos?

Após busca criteriosa nos principais repositórios de teses, dissertações e periódicos de divulgação científica (ver página 24), não foram encontrados critérios consolidados para avaliar *chatbots* usados como materiais didáticos. Assim, a produção de uma matriz com critérios e descritores é uma contribuição para o campo dos estudos de linguagens e da educação mediada por tecnologias digitais. Portanto, este estudo contribui para o preenchimento de uma lacuna na produção acadêmica sobre o ensino e aprendizagem mediados por tecnologias, especialmente por inteligência artificial. Será criada e validada uma matriz para se contemplar as especificidades do *chatbot* usado como MD, ou *chatbot* didático (ver proposta de uso dessa nomenclatura do capítulo 3), o que se configura como parte do ineditismo e contribuição desta tese.

---

<sup>2</sup> Conceito de Inteligência Artificial e discussões sobre o seu uso na educação são apresentados no capítulo 2.

Como a produção de uma matriz de avaliação de *chatbots* didáticos é uma contribuição da pesquisa, faz-se necessário o esclarecimento do termo matriz. Entende-se que matriz é um instrumento que “apresenta o objeto de uma avaliação e é formada por um conjunto de descritores” que guiam a avaliação desse objeto. (PORTAL DA AVALIAÇÃO, 2014).

Nesta tese, além de apresentar discussões e reflexões sobre o processo de avaliação de *chatbots* didáticos, apresenta-se também um estudo de caso sobre a experiência com a utilização desse tipo de *chatbot*<sup>3</sup> em um curso de formação de professores.

Esta tese insere-se no conjunto de pesquisas do Grupo Infortec, do CEFET-MG, liderado pelo orientador Prof. Dr. Vicente Aguiar Parreiras, que têm o intuito de investigar os usos das tecnologias digitais da informação e da comunicação como instrumento de mediação e de apoio ao ensino e à aprendizagem na perspectiva da robótica pedagógica, fundamentado na linguística aplicada.

## 1.1 Contextualização do Estudo

As tecnologias digitais de informação e comunicação – TDIC - dão origem a novos tipos de materiais didáticos, seja pela criação de novos produtos editoriais ou tecnológicos, seja pela apropriação didática de artefatos já existentes. Reflete-se nesta tese sobre o uso do *chatbot* como material didático (MD). Deryugina (2010) e Duncan (2018) utilizam o termo *didactic chatbot*, que será usado em livre tradução nesta tese como *chatbot* didático.

Usar o conceito de material didático para se falar dos *chatbots* usados para fins de aprendizagem é uma escolha deste pesquisador, baseado em autores que já utilizam o termo, principalmente em literatura publicada em língua inglesa, pois acredita-se ser necessário assumir que diversas mídias, tecnologias e suportes podem ser usadas com este fim e que novos tipos de MD serão criados com base no desenvolvimento tecnológico. Usar o termo material didático, sem apego somente aos materiais considerados tradicionais e sem pudor, para falar sobre aquilo que de fato é um MD, é uma valorização da produção e pesquisa sobre MDs, não o contrário.

---

<sup>3</sup> A discussão sobre o uso do termo material didático para *chatbots* com fins pedagógicos é apresentada no capítulo 3.

*Chatbots* são *softwares* que simulam um ser humano na interação/conversa o com seres humanos reais. Ferreira e Uch a (2008) explicam que se trata de uma entidade de intelig ncia artificial capaz de simular um interlocutor humano atrav s do Processamento de Linguagem Natural (PLN). Em suma, o sistema tem o objetivo de responder  s perguntas como se fosse uma pessoa, utilizando uma linguagem pr xima a de falantes ou redatores humanos. Assim, quem interage com o *software* t m a impress o de que est  conversando com uma pessoa, quando na verdade quem responde   uma m quina, um rob  virtual.

Essa fa anha   obtida ao se manipular, atrav s de uma linguagem de programa o, express es, termos e outras inst ncias de uma determinada l ngua humana atrav s de c digos especiais, que criam n o apenas padr es de buscas lingu sticas, mas implementam tamb m recursos de mem ria, de compara es, de restri es de assuntos e at  de indaga es sobre a declara o feita pelo usu rio humano produzindo respostas e, portanto, simulando di logos t picos de uma intera o entre duas pessoas. (LIMA, 2014, p. 37)

Apesar de Ferreira e Uch a (2008) falarem que o *chatbot* simula ser uma pessoa, h  *chatbots* que se apresentam de outras maneiras, assumindo sua condi o de m quina, se apresentando como “assistente virtual” ou “rob ”, por exemplo. Pode-se criar diferentes “personas” para um *chatbot*, mas a inten o   sempre a conversa o com um ser humano. Lima (2014) explica que, para simular a conversa o com um ser humano, o *chatbot* pode usar, al m da “modalidade de bate-papo textual e online”, outros recursos como reconhecimento de voz e at  mesmo express es faciais, no caso de *chatbots* que contam com anima es ou avatares.

Para falar sobre o uso do *chatbot* como MD, faz-se necess rio esclarecer porque se classifica o *chatbot* como did tico e n o como pedag gico e o que se considera material did tico nesta pesquisa.

Nesta pesquisa, adota-se as defini es de Meyer (2007) e L on-Henri (2016) sobre o que   did tico e o que   pedag gico. Enquanto o pedag gico envolve processos e rela es de ensino e aprendizagem, considerando n o apenas professores e estudantes, mas a sociedade e as pessoas envolvidas nesse processo de maneira mais ampla, o did tico cuida mais diretamente dos processos de ensino (L ON-HENRI, 2016).

O pedag gico, ent o,   uma complexa mistura de conhecimentos te ricos e habilidades pr ticas envolvidas nos processos do campo da Educa o (L ON-HENRI, 2016). A did tica, por sua

vez, é uma disciplina essencialmente preocupada com as instruções que se dão em uma área de estudo, com o conteúdo, as estratégias e técnicas envolvidas no fazer docente (MEYER, 2007). O didático, então, se apresenta como parte da complexa gama de estudos e fazeres que compõem o pedagógico. Entende-se que o modelo de *chatbot* proposto para estudo nesta pesquisa está envolvido diretamente com os métodos e tecnologias de apoio à docência, mas não contempla de maneira ampla outros processos que fazem parte do campo pedagógico.

Em relação à definição sobre MD, não existe um consenso sobre o que é ou não um material didático. Para Trojan e Rodríguez (2008, p.57), MD é “qualquer tipo de material que se destina a ser utilizado por alunos e professores quando tem como finalidade específica auxiliar no processo de planejamento, desenvolvimento e avaliação da proposta curricular”. Tomlinson (1998) e Rangel (2005), também defendem que qualquer instrumento utilizado para fins de ensino e aprendizagem é considerado um MD. Nessa perspectiva, “uma garrafa de refrigerante torna-se um MD quando o professor de Química a utiliza em sala de aula para ajudar a explicar como se dá o processo de gaseificação de alimentos” (CAMARGO, 2013).

Outros autores discordam dessa perspectiva e, para eles, essa mesma garrafa de refrigerante seria um recurso didático ou material de apoio, mas não um MD. Graells (2000) e Bandeira (2009) afirmam é necessário que o objeto seja construído com intencionalidade para uso didático para que seja considerado um MD.

Bandeira (2009, p. 14) explica que MD é “o material instrucional que se elabora com finalidade didática”. Opta-se nesta tese por trabalhar nesta perspectiva, assumindo como base o conceito de Graells, que distingue os conceitos de material didático e recurso educativo, considerando que “material didático é todo material elaborado com a intenção de facilitar os processos de ensinar e aprender”, enquanto o recurso educativo seria qualquer material utilizado com finalidade didática, mesmo que não tenha sido criado com esse objetivo. Ou seja, é a intenção pedagógica em seu processo de elaboração o que faz de um objeto um material didático.

Assim, entende-se que o *chatbot* pode ser considerado um material didático se é criado especificamente para facilitar os processos de ensino e aprendizagem. Portanto, um *chatbot* já existente, como um atendente virtual de uma rede varejista, se usado numa aula de marketing para realização de atividade sobre como o mercado tem utilizado a inteligência artificial na relação com os consumidores, não seria considerado um material didático, mas um recurso

educativo, de acordo com a definição de Graells. Já um *chatbot* criado especialmente para a disciplina de Marketing, com o objetivo de auxiliar os alunos esclarecendo dúvidas sobre os textos e atividades disponibilizados pelo professor, seria um material didático, já que a elaboração desse *chatbot* teria como objetivo central contribuir para os processos de ensino e aprendizagem nessa disciplina.

Para Graells (2000), os MD podem ser utilizados com diferentes funções em situações de ensino e aprendizagem: fornecer informações, guiar a aprendizagem, exercitar habilidades, motivar, avaliar, promover simulações e proporcionar oportunidades para a expressão e criação. O pesquisador divide os MD em três categorias, conforme a

FIGURA 1.

**FIGURA 1 - Tipos de MD segundo Graells**

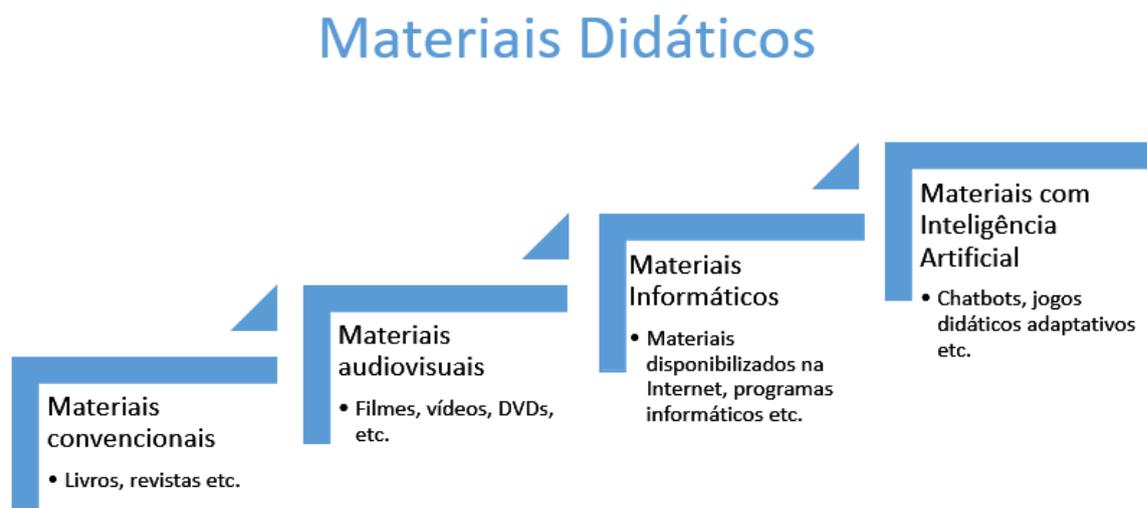


Fonte: Camargo (2013) baseado em Graells (2000)

Graells fez a proposta de categorias há mais de duas décadas e as TDICs tiveram grandes avanços nesse tempo, o que impactou também na criação de novos tipos de MD. Por isso, propõe-se nesta tese a amplificação e atualização das categorias de Graells. Primeiramente, revisou-se a redação da terceira categoria, pois considera-se que Novas Tecnologias já não é mais um termo adequado na terceira década do século XXI, visto que a disponibilização de materiais na Internet já não é algo tão recente ou “nova”, principalmente considerando-se que grande parte dos jovens e adultos que hoje estão nas escolas e universidades já nasceram quando

programas informáticos e a internet eram usados na educação. Assim, alterou-se a redação da categoria para “Materiais Informáticos”. Propõe-se também a criação de uma quarta categoria: Materiais com Inteligência Artificial. São materiais informáticos, mas com tecnologias específicas que utilizam Inteligência Artificial na programação dos sistemas, como os *chatbots*, os jogos didáticos adaptativos, ambientes virtuais de aprendizagem adaptativos e smartbooks, que de acordo com Vicari (2018), são os livros didáticos customizados/adaptáveis de acordo com o perfil do estudante. Este trabalho enquadra-se nesta proposta de quarta categoria de MD. Ao propor essa atualização, apresenta-se um novo desenho para representar os tipos de MD, conforme mostra a figura 2.

**FIGURA 2 - Tipos de Materiais Didáticos (proposta atualizada)**



Fonte: Baseado em Graells (2000)

## 1.2 O objeto do estudo de caso

Esta pesquisa apresenta o estudo de caso da implementação de um *chatbot* no curso *Teaching and Learning in Higher Education* (TLHE), um curso de formação de professores ofertado pelo

Instituto Ânima<sup>4</sup> em parceria com a Finland University<sup>5</sup> e Universidade de Tampere (Finlândia) para docentes brasileiros. O curso tem carga horária de 270 horas em modelo híbrido (*blended learning*), com 70 horas de aulas presenciais e 200 horas de estudos a distância, com o uso do ambiente virtual de aprendizagem Moodle. O público do curso são professores universitários e gestores do ensino superior brasileiros. As aulas são ministradas em inglês por professoras da Faculdade de Educação da Universidade de Tampere.

Segundo o site oficial do curso<sup>6</sup>, os objetivos de aprendizagem são:

- conceber, conduzir e avaliar atividades de aprendizagem flexíveis/adaptativas, em ambientes de aprendizagem versáteis e inovadores;
- dominar questões teóricas centrais relativas ao ensino e aprendizagem na educação superior;
- identificar e superar desafios pessoais e institucionais relacionados ao desenvolvimento de competências pedagógicas teóricas e práticas. (INSTITUTO ANIMA, 2020)

O curso é dividido em quatro unidades de ensino: 1) Aprendizagem e Engajamento dos Alunos; 2) Criando Processos e Ambientes de Aprendizagem; 3) Tópicos Especiais do Ensino na Universidade; e 4) Desenvolvendo Expertise em Ensino e Aprendizagem no Ensino Superior. O Quadro 1 detalha o conteúdo das unidades.

---

<sup>4</sup> O Instituto Ânima é uma instituição sem fins lucrativos que tem por finalidade a formação de professores, o fomento, o estímulo e o desenvolvimento de atividades ligadas à cultura e desenvolvimento social e sustentável, à ciência e à educação do nível básico ao superior. É ligado ao Grupo Ânima Educação, grupo educacional privado que conta com cerca de 110 mil alunos em cinco estados brasileiros, em quatro instituições: Una (MG e GO), Universidade São Judas (SP), UniBH (MG) e UniSociesc (SC e PR). Também são parte do grupo a Escola Brasileira de Direito (EBRADI), a HSM Educação Executiva e a filial brasileira da escola internacional de gastronomia Le Cordon Bleu, em São Paulo. Informação disponível em: [www.nossoinstitutoanima.com.br](http://www.nossoinstitutoanima.com.br).

<sup>5</sup> A Finland University é um consórcio formado por quatro das principais universidades públicas finlandesas: Universidade do Leste da Finlândia, Universidade de Tampere, Åbo Akademi University e Universidade de Turku. Fundada em 2014, a instituição conta com mais de 100 mil alunos. Segundo a própria instituição, seu objetivo é levar a instituições do mundo as melhores práticas do sistema de ensino finlandês, referência mundial em educação. Informação disponível em: [www.finlanduniversity.com](http://www.finlanduniversity.com)

<sup>6</sup> <https://www.nossoinstitutoanima.com.br/capacitamos-formamos/especializacao-em-ensino-e-aprendizagem-na-educacao-superior-2/>

**QUADRO 1 - Estrutura do curso**

<b>Unidade de ensino</b>	<b>DESCRIÇÃO</b>
Aprendizagem e engajamento dos alunos	Abordagem de ensino centrada no aluno; Estudantes de hoje e de amanhã; Teorias de ensino e aprendizagem e suas aplicações em diferentes contextos.
Criando processos e ambientes de aprendizagem	Alinhamento construtivo; Modelos curriculares alternativos de apoio à aprendizagem; Ambientes flexíveis de aprendizagem e aprendizagem <i>online</i> .
Tópicos especiais do ensino na universidade	Processos de supervisão e orientação acadêmica; Métodos para relacionar ensino e pesquisa; Sala de aula invertida.
Desenvolvendo expertise em ensino e aprendizagem no ensino superior	Identificação e superação de desafios pessoais e institucionais no desenvolvimento da prática docente; Entendimento da competência pedagógica como uma competência da comunidade acadêmica; Criação de projeto pedagógico.

Fonte: Instituto Ânima (2019)<sup>7</sup>

As aulas presenciais são realizadas em São Paulo, na Universidade São Judas Tadeu, uma das instituições de ensino superior mantidas pelo grupo Ânima Educação. As professoras vêm da Finlândia exclusivamente para ministrar os três módulos presenciais do curso, que somam 10 dias de aulas (70 horas). As outras 200 horas do curso são realizadas em módulos *online* por meio do ambiente virtual de aprendizagem Moodle. Nesses módulos a distância são realizadas discussões, estudos de caso, fóruns e webinários sobre os temas dos módulos. Nas aulas presenciais acontecem exposições das professoras, discussões, dinâmicas e atividades práticas. O curso é ministrado em inglês, portanto a proficiência no idioma é pré-requisito para matrícula.

Durante todo o curso, são realizados estudos individuais, atividades avaliativas individuais e em grupo. Uma atividade chamada de Projeto Pedagógico percorre todo o curso e consiste no planejamento, execução e compartilhamento dos resultados de um projeto prático realizado pelos alunos (todos os alunos já atuam como docentes no Ensino Superior) com seus próprios alunos em contexto real de sala de aula, utilizando as teorias e metodologias aprendidas no curso. Os alunos que concluem o curso recebem certificação pela Universidade de Tampere, uma das quatro universidades públicas finlandesas que formam a Finland University.

A primeira turma do curso, realizada em 2017, formou 38 professores. O pesquisador foi um dos alunos desta turma. Em seguida, o pesquisador assumiu o papel de coordenador do curso

<sup>7</sup> <https://www.nossoinstitutoanima.com.br/capitamos-formamos/especializacao-em-ensino-e-aprendizagem-na-educacao-superior-2/>

nas turmas subsequentes. Foi nesse momento que surgiu a ideia de criar o *chatbot*. A instituição já planejava testar novas tecnologias educacionais e experimentar o uso de inteligência artificial como suporte aos cursos. Como o pesquisador-coordenador já integrou a equipe de Inovação Acadêmica e tem experiências em pesquisas sobre tecnologias educacionais, sentiu-se motivado a abraçar o desafio. Entre as possibilidades do uso de inteligência artificial, o pesquisador propôs a implementação de *chatbot*, pois havia um incômodo pessoal, compartilhado também por outros membros da equipe do Instituto Ânima, em ter que responder diversas vezes, para diferentes alunos, as mesmas perguntas. O pesquisador e a equipe acreditavam que esse trabalho, que passa a ser repetitivo, poderia ser realizado por máquinas, deixando mais tempo para que o professor e a equipe pedagógica pudessem se ocupar com esforços criativos. Haveria então benefícios pedagógicos, pelo apoio do *chatbot* em tempo integral aos estudantes, e benefícios administrativos.

A partir desse desafio, durante o período de aulas da turma de 2018 do curso TLHE, o *chatbot* Katri foi criado como MD do curso. A turma teve 32 alunos e, ao final do curso, quando a primeira versão ou versão piloto de Katri estava pronta, quatro alunos foram convidados para testá-la, contribuindo para melhoramentos. No capítulo 6 desta tese é apresentado o processo de criação de Katri, inclusive os resultados do teste dessa versão piloto. A versão definitiva de Katri foi disponibilizada para todos os 32 alunos da turma de 2019 do curso TLHE. No capítulo 7, os resultados da interação com esses alunos e avaliação de Katri serão apresentados. Os alunos e também os produtores do *chatbot* avaliaram Katri baseados em uma matriz que foi criada durante esta pesquisa. A matriz é apresentada no capítulo 3.

### **1.3 Problema e Objetivos de Pesquisa**

As professoras finlandesas que ministram o curso TLHE não são proficientes em português e, por isso, o curso é realizado em inglês para professores brasileiros. Apesar de todos os alunos/professores brasileiros conseguirem se comunicar em inglês, há diferentes níveis de proficiência. Alguns alunos, inclusive, apresentam mais dificuldades em falar e escrever com fluidez, apesar de ler e escutar mais facilidade (analisar a proficiência dos alunos não faz parte do escopo desta pesquisa). Esses é um dos motivos pelos quais os alunos buscaram o coordenador do curso (pesquisador/autor desta tese) para esclarecer dúvidas administrativas e de conteúdo.

Este coordenador é o único brasileiro que trabalha diretamente com as professoras finlandesas no curso; então todas as dúvidas, quando enviadas em português, são esclarecidas apenas por uma pessoa.

Diante desta situação, considerou-se importante manter um “monitor” ou “assistente” virtual em prontidão para esclarecimento de dúvidas 24 horas por dia. Trabalhou-se com a expectativa de que, com um *chatbot* atuando neste papel, as respostas poderiam chegar mais rápido aos alunos e o fluxo de demandas que o coordenador precisaria responder poderia diminuir. Ao iniciar a pesquisa sobre como fazer um *chatbot* que conseguisse responder dúvidas de conteúdo, se apresentando então como um material didático do curso, o pesquisador se deparou com a dificuldade de encontrar critérios para produção ou avaliação de *chatbots* usados para este fim. Este problema foi o principal motivador desta pesquisa.

Partiu-se da hipótese de que era possível criar uma matriz com critérios para avaliação de *chatbots* usados como material didático, tendo como base critérios para avaliação de materiais didáticos de outras naturezas e de critérios para avaliação de *chatbots* usados para fins não didáticos. A proposta nesta pesquisa foi criar essa matriz e validar a sua utilização a partir dos resultados de um estudo de caso baseado no uso do *chatbot* do curso TLHE.

O objetivo geral desta pesquisa é criar e validar o uso de uma matriz com critérios e descritores para avaliação de *chatbots* didáticos. Os objetivos específicos são:

1. mapear critérios para produção e avaliação de *chatbots* didáticos;
2. observar a produção e o uso de um *chatbot* didático em um curso de formação de professores;
3. utilizar a matriz proposta, com seus critérios e descritores, para avaliar o *chatbot* do curso TLHE considerando as percepções de produtores deste *chatbot* e alunos do curso;<sup>8</sup>
4. sugerir melhorias e revisões para o *chatbot* com base na avaliação realizada com o uso da matriz proposta.

Com base em quais critérios e descritores deve-se avaliar a qualidade de um *chatbot* didático?

---

<sup>8</sup> A todos os participantes foi solicitada a leitura e assinatura do termo de consentimento livre e esclarecido (Apêndice 1).

A possível ausência desses critérios foi sentida pelo pesquisador e por produtores de *chatbot* envolvidos em um desafio: criar um *chatbot* didático para um curso de formação de professores. Atingir o objetivo geral e relatar o percurso desta pesquisa pode ser útil para facilitar ou inspirar produções de *chatbots* para serem usados como MD em outros cursos.

#### 1.4 Justificativa

Ensinar o estudante e não a turma toda, essa é a premissa de uma educação adaptativa. Ter como atingir o indivíduo, em vez da turma toda, é uma questão cada vez mais presente e possibilitada por tecnologias como a big data, pois possibilita que cada indivíduo aprenda da melhor maneira possível e, mais importante que isso, da sua maneira. (SIMON, 2017, p. 66)

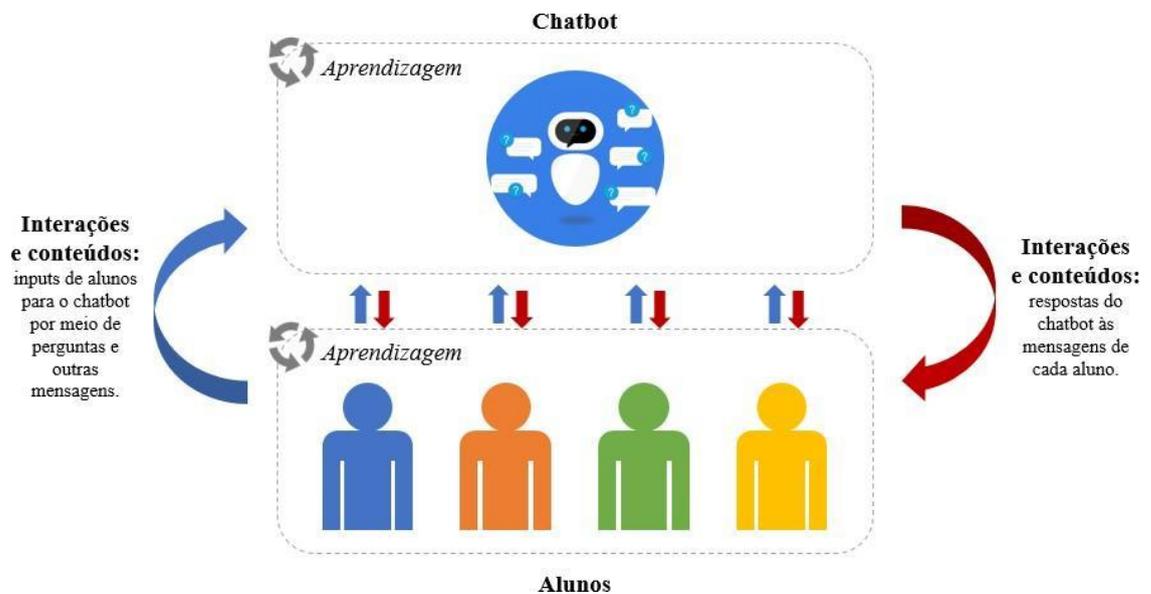
Esta pesquisa justificou-se pela necessidade de se investigar a produção, avaliação e o uso de uma tecnologia digital que tem alcançado destaque recentemente: os *chatbots*, ou agentes de conversação, que em contextos educacionais podem ser entendidos como um tipo de sistema de aprendizagem adaptativo. Acredita-se que essa investigação é pertinente e relevante, pois o caráter adaptativo desse tipo de sistema se faz necessário à medida em que se entende que a aprendizagem dos estudantes não se dá de maneira ordenada, fixa e estável, mas muitas vezes de forma caótica, visto que essa aprendizagem - e o estudante também - pode ser considerado um sistema complexo (PAIVA; NASCIMENTO, 2009).

Se cada estudante tem em si um próprio processo (e sistema) de aprendizagem, é relevante pensar que os sistemas de estudo *online* não devem tratar os processos de aprendizagem de seus alunos como se todos aprendessem da mesma maneira, ofertando a grandes grupos de alunos os mesmos textos, no mesmo tempo, dentre outros elementos que fazem parte do processo de aprendizagem *online*. Assim, o que se propõe é uma pesquisa que amplia o entendimento do funcionamento de um sistema de aprendizagem que entrega conteúdos de acordo com as intenções de estudos manifestadas por cada aluno. Sendo assim, é um sistema em que cada estudante pode ter uma experiência diferente, com diálogos e acesso a conteúdos que outros alunos da mesma turma podem nunca ter visto, pois os textos e outros recursos que são visualizados são respostas deste sistema de aprendizagem adaptativo aos estímulos que a máquina recebeu do próprio estudante/usuário.

Um processo como esse pode ser considerado um sistema complexo formado por outros sistemas complexos, pois além de se considerar a aprendizagem dos alunos, é necessário lembrar que o próprio *chatbot* também tem o seu próprio sistema de aprendizagem, também

complexo, visto que tem uma condição inicial (programada em sua primeira versão pelos produtores), mas ele aprenderá e se desenvolverá de acordo com as interações com os usuários. Então, um sistema alimenta o outro. O aluno pode aprender com o *chatbot* ao interagir com ele para estudar. Essa interação com o aluno aumentará o volume de informações do banco de dados do *chatbot* e aprimorará a maneira como ele consegue entender os diálogos, o que também vai gerar um aprendizado para a máquina. A figura 3 ilustra esse processo de retroalimentação e aprendizagem na interação entre estudantes e *chatbot*.

**FIGURA 3 - Aprendizagem por *chatbots* educacionais**



Fonte: Elaborado pelo autor

Medeiros, Júnior e Moser (2017, p. 2) afirmam que na segunda década do século XX observa-se um “avanço vertiginoso das tecnologias educacionais, potencializado pelo uso da Inteligência Artificial e permitindo novas abordagens com relação à aprendizagem adaptativa”. Entre as tecnologias que os autores apresentam para a promoção dessa aprendizagem que se adapta ao percurso e processo individual de cada estudante estão os *chatbots*. Segundo os autores, “tecnologias que oferecem interfaces com a linguagem natural são uma boa opção para obter melhores interações pelo uso de agentes de interface conversacional incorporados” (MEDEIROS; JÚNIOR; MOSER, 2017, p. 2).

Lima (2014, p. 14) trabalhou na implantação de um *chatbot* numa disciplina do curso de Letras numa universidade pública de Minas Gerais. Ele afirma que “esse tipo de ferramenta de conversação automatizada pode se tornar um recurso versátil dentro do processo de ensino- aprendizagem *online*” e se tornar um “grande aliado e facilitador de interação”. O autor elenca quatro possibilidades que deveriam ser entregues por essa ferramenta. Para ele, é preciso observar se ela:

1) proporciona uma interface de conversação interessante (natural); 2) fornece conteúdo relevante (especializado) e adaptável (inteligência artificial); 3) dá liberdade de diálogo aos alunos (quebra de timidez, medo de perguntar); e 4) pode atendê-los em todos os horários e a partir de todos os locais com acesso *online* (disponibilidade 24/7 – vinte quatro horas por dia, sete dias por semana). (LIMA, 2014, p.15)

Apesar das indicações de que o bom uso dos *chatbots* pode trazer contribuições para o processo de aprendizagem, encontrou-se ainda pouca investigação científica sobre o fenômeno, talvez justificado por ser uma tecnologia recente, com ainda menos tempo de uso para fins educacionais (VICARI, 2018; CASAGRANDE, 2019). Ampliando o foco da pesquisa e ao buscar produção sobre a utilização de materiais didáticos que utilizam inteligência artificial, percebeu-se que essa produção é relativamente escassa. Apesar disso, instituições de ensino têm experimentado o uso desse tipo de tecnologia. Essas experimentações, entretanto, nem sempre se tornam públicas por meio de pesquisas e divulgação científica.

Em pesquisa de acesso livre no Portal de Periódicos de Capes<sup>9</sup>, considerando artigos acadêmicos publicados em periódicos revisados por pares, os termos de pesquisa “*chatbot*” e/ou “*chatbots*” resultam num total de 923 publicações, a maioria escrita em língua inglesa. Dentre o total de artigos, destacam-se as áreas de inteligência artificial e ciências da computação com 267 e 198 artigos, respectivamente. Dentro da área de linguagens, foram encontrados 87 artigos, todos indexados dentro da subcategoria de processamento de linguagem natural. Nenhum dos artigos encontrados tem foco na avaliação de *chatbots* didáticos.

No banco de dados do Catálogo de Teses e Dissertações da Capes<sup>10</sup>, em que são indexadas dissertações de mestrado e teses de doutorado, ao realizar a mesma busca por “*chatbot*” e/ou “*chatbots*”, encontra-se 37 dissertações e apenas 8 teses já publicadas de pesquisas realizadas no Brasil. Nenhuma dissertação ou tese encontrada discute a avaliação de *chatbots* didáticos.

Da mesma forma que os artigos, a maioria das dissertações e teses que estudam *chatbots* são da área de Ciência/Engenharia da Computação e Sistema de Informação/Informática, com um total de 16 trabalhos. Foram apenas cinco trabalhos em Linguística/Estudos Linguísticos.

---

<sup>9</sup> Pesquisa realizada em 10/01/2021 que inclui “*chatbot*” e “*chatbots*” como palavras-chave mencionadas em qualquer parte do texto.

<sup>10</sup> Pesquisa realizada em 12/01/2021.

Os artigos, dissertações e teses considerados pertinentes para esta pesquisa são apresentados ao longo desta tese.

Para uma simples comparação, o termo “*artificial intelligence*” resulta em 275.702 publicações em periódicos revisados por pares no Portal de Periódicos da Capes de acesso livre, sendo que 1082 estão em português. Ou seja, Inteligência Artificial já é um tema explorado em um número maior de pesquisas, mas a aplicação desta tecnologia em *chatbots* ainda carece de referências, assim como sua utilização didática. Dentre essas pesquisas, 6057 se concentram na área de Linguística e/ou de Educação. Entre as teses e dissertações, temos mais de 21.071 resultados para “*inteligência artificial*” (14.347 dissertações e 5.349 teses), das quais 484 são da área de Educação e 316 são da área de Letras/Linguística, mas nenhuma delas apresenta a intersecção com os termos “*chatbot*” ou “*chatbots*”.

Nos anais de cinco edições (de 2015 a 2019) do Congresso Internacional ABED de Educação a Distância - CIAED, evento organizado pela Associação Brasileira de Educação a Distância - ABED, encontrou-se apenas 3 textos que discutem a utilização de inteligência artificial na educação. Esse número representa 0,2% dos 1454 trabalhos publicados nos anais dessas cinco edições do evento. As propostas de discussões desses textos são apresentadas a seguir.

Medeiros, Júnior e Moser (2017) relataram no CIAED o uso do *chatbot* THOTH no apoio às aulas da disciplina Inteligência Artificial Aplicada, no curso de graduação em Análise e Desenvolvimento de Sistemas de um centro universitário do Paraná. Aline Silva et. al. (2017) escreveram sobre um sistema de IA que foi usado em uma universidade federal do Piauí para acompanhar determinados resultados de aprendizagem dos estudantes e descobrir a correlação entre o Índice de Rendimento Acadêmico e alguns aspectos sociais desses alunos, como estado civil, ano de conclusão do ensino médio, raça, gênero etc, o que segundo os autores poderia auxiliar na tomada de decisão da equipe de gestão acadêmica. Dois anos antes, Xavier, Netto e Alencar (2015) apresentaram no evento um trabalho sobre um sistema que usa inteligência artificial para facilitar o processo de gestão de cursos por coordenadores em um Instituto Federal de Educação. Apenas no primeiro trabalho pode-se ver a utilização dessa tecnologia como MD, visto que nos outros dois casos o uso da tecnologia se prestava exclusivamente a fins administrativos.

Esta tese apresenta-se como uma contribuição para o campo dos estudos de linguagens,

principalmente para as pesquisas relacionadas ao ensino mediado pelas tecnologias digitais da informação e da comunicação, por trazer tema e abordagem ainda novos e pouco explorados, mas que se conectam diretamente com a educação e a sociedade contemporânea. Nesta pesquisa, é relatado o processo de criação, uso e avaliação de uma tecnologia emergente, o que pode gerar *insights* para o desenvolvimento de novos projetos educacionais com o uso de *chatbots* e inteligência artificial.

No próximo capítulo desta tese, discute-se o uso de Inteligência Artificial e Processamento de Linguagem Natural na Educação, com foco especial nas questões relacionadas às estratégias de desambiguação por máquinas e na contextualização desse tema que se apresenta com uma tendência em tecnologias educacionais. No capítulo 3, discute-se Critérios de Avaliação de MD e de *chatbots*. Com base nesses critérios, é apresentada a matriz proposta nesta pesquisa. O capítulo 4 apresenta o Sistema de Avaliatividade proposto por Martin e White (2005) como uma extensão da Linguística Sistêmico-Funcional (HALLIDAY; MATTHIESSEN, 2014). Esse sistema será utilizado como referência para a análise dos resultados desta pesquisa. O capítulo 5 descreve os procedimentos metodológicos adotados e apresenta os participantes da pesquisa. No capítulo 6, é apresentado primeiramente o estudo piloto realizado com a primeira versão de Katri, *chatbot* do curso TLHE. Em seguida, são apresentados e analisados os resultados da pesquisa: a validação da matriz criada para avaliação de *chatbots* a partir da avaliação da versão definitiva de Katri por alunos do curso e produtores do *chatbot*. Por fim, são apresentadas as considerações finais desta pesquisa.

## 2. REFERENCIAL TEÓRICO

### 2.1 Inteligência Artificial e Processamento de Linguagem Natural na Educação

Inteligência Artificial é um termo cunhado por John McCarthy em 1956 (apud NAVEGA, 2000), em um famoso *Workshop* do Dartmouth College em que ele e mais nove pesquisadores discutiram o tema por dois meses. Foi nesse evento que McCarthy designou a IA como a capacidade de uma máquina realizar funções consideradas inteligentes quando são realizadas por seres humanos. Em definição mais recente, Vicari (2018, p. 11) afirma que a IA “do ponto de vista simbólico, pode ser definida como a arte de se construir programas que se adaptem e aprendam, com a finalidade de prolongar o seu ciclo de vida”.

Antes mesmo da definição de McCarthy, entretanto, já havia estudos iniciais sobre IA. Alan Turing é reconhecido como o primeiro pesquisador a articular uma visão geral sobre IA no seu artigo “*Computing Machinery and Intelligency*”, de 1950. Nele, foi apresentado o Teste de Turing, em que uma máquina seria considerada inteligente se for aprovada num jogo: um jogador humano participa de uma conversa em linguagem natural com uma máquina e outro jogador humano. Os três atores estão em espaços separados. Se o primeiro jogador não conseguir distinguir com segurança as respostas elaboradas pela máquina das elaboradas por outro ser humano, considera-se que a máquina foi aprovada no teste. Percebe-se então como o processamento de linguagem natural faz parte dos estudos de inteligência artificial desde as pesquisas iniciais sobre o tema.

Processamento de Língua Natural é um dos temas que surgiu com a IA e, embora tenha ficado latente durante vários anos, ressurgiu com força nos últimos anos graças a dois fatores: o poder computacional das máquinas atuais e dos produtos e serviços de empresas. O Processamento de Língua Natural trata da compreensão da língua escrita e falada, da tradução e da geração da língua (falada e escrita). Para isso, são utilizadas várias tecnologias da IA e da Estatística. Essa tecnologia já está começando a ser usada na Educação para correção de textos escritos pelos alunos e tradução simultânea de texto e voz. (VICARI, 2018, p.32).

Ao longo das últimas décadas, já foram apresentadas diferentes linhas de pensamento para definir o que é a IA. Gomes (2010, p. 235) categorizou essas linhas em quatro definições:

I. Sistemas que pensam como seres humanos: ‘o novo e interessante esforço para fazer os computadores pensarem... máquinas com mentes, no sentido total e literal’. (HAUGELAND, 1985). II. Sistemas que atuam como seres humanos: ‘a arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.’ (KURZWEIL, 1990). III. Sistemas que pensam racionalmente: ‘o estudo das faculdades mentais pelo seu uso de modelos computacionais.’ (CHARNIAK;

MCDERMOTT, 1985). IV. Sistemas que atuam racionalmente: ‘a Inteligência Computacional é o estudo do projeto de agentes inteligentes’ (POOLE et al., 1998).

Para Gomes, enquanto as linhas de pensamento I e III estão ligadas ao pensamento e raciocínio, as linhas II e IV estão ligadas ao comportamento. As linhas I e II mediriam o sucesso de acordo com o grau de fidelidade ao desempenho humano nas funções que realizam, enquanto as linhas III e IV teriam sucesso de acordo com uma comparação a um modelo ideal de inteligência ou racionalidade. Seria considerado um modelo ideal/racional o sistema que “faz tudo certo, com os dados que tem” (GOMES, 2010, p.235). Entende-se então que o campo da inteligência artificial, independente da linha adotada, além de buscar entender a cognição humana, tenta construir sistemas inteligentes.

As quatro linhas têm sido seguidas em diferentes projetos que usam IA, realizadas atualmente em todas as áreas do conhecimento. Para Gomes (2010, p. 234), “a inteligência artificial sistematiza e automatiza tarefas intelectuais e, portanto, é potencialmente relevante para qualquer esfera da atividade intelectual humana”.

No campo da educação, diversas utilizações são possíveis, muitas já em utilização, e várias outras em fases de desenvolvimento ou em testes. Sistemas de aprendizagem são um rico campo para a IA. Vicari (2018) afirma que “a IA aplicada à Educação é uma área de pesquisa multi e interdisciplinar, pois contempla o uso de tecnologias da IA em sistemas cujo objetivo é o ensino e a aprendizagem”.

Os sistemas de inteligência artificial aplicados à educação podem ser divididos em duas categorias amplas: sistemas de IA para gestão escolar e sistemas de IA para apoio à processos de aprendizagem. Na primeira categoria, alguns exemplos são o uso de *big data* para análise de desempenho dos alunos, o uso de reconhecimento facial para controle de presença e evasão, e também o atendimento por *chatbots* para esclarecer dúvidas sobre demandas administrativas e de secretaria. Já na segunda categoria, estão o uso de *big data* para promoção da aprendizagem adaptativa em ambientes virtuais de aprendizagem, onde conteúdos e avaliações são customizadas automaticamente para cada aluno, a legendagem automática de vídeos ao vivo em diferentes línguas, o que pode promover a inclusão na sala de aula, e também o uso de *chatbots* para esclarecimentos de dúvidas individuais de alunos sobre conteúdos, agendas de avaliações, fontes de estudos, entre outros apoios ao processo de aprendizagem. Vicari (2018) apresenta o que ela considera como principais sistemas educacionais que utilizam a IA:

Sistemas Tutores Inteligentes Afetivos, os Learning Management Systems, a Robótica Educacional Inteligente e os Massive Open *Online* Course (MOOCs), no que se refere a Learning Analytics (LA). Entretanto cada uma dessas aplicações faz uso de tecnologias da IA de formas distintas. (VICARI, 2018, p. 12)

Sobre o uso de IA na educação, Cabral (2018, n.p.) afirma que “as decisões pedagógicas estão e estarão cada vez mais amparadas em dados” e que “o professor já conta com algoritmos que permitem, cada vez melhor, a personalização das experiências e da aprendizagem de cada aluno”.

Parte significativa dos sistemas educacionais que utilizam IA dependem do Processamento de Linguagem Natural (PLN). A maior parte trabalha com o texto escrito, mas há também os sistemas que utilizam voz e vídeo. Vieira e Lopes (2010, p.184) conceituam o PLN como uma área “que estuda o desenvolvimento de programas de computador que analisam, reconhecem e/ou geram textos em linguagens humanas”. As autoras explicam que alguns dos objetivos usuais do PLN são a recuperação de informação a partir de textos, a tradução automática, a interpretação de textos, a realização de inferências a partir de textos e a construção de antologias a partir de textos. Esses objetivos de uso já são (mas não se limitam a eles) algumas das aplicações do PLN na educação em sistemas que utilizam IA.

Ao usar PLN em sistemas educacionais com IA, pode-se, inclusive, utilizar tecnologias que usam estratégias de afetividade em textos criados pelas máquinas e o reconhecimento por elas da afetividade e emoções dos usuários humanos. Há sistemas educacionais que já trabalham nessa perspectiva. “Em grande parte dos artigos e produtos, essas duas tecnologias caminham juntas. Esses fatos levam à prospecção de que eles podem ser incorporados aos diferentes sistemas de ensino-aprendizagem em curto espaço de tempo” (VICARI, 2018, p. 21).

Vicari (2018) apresenta um estudo em que foram prospectadas as tendências do uso de IA na educação até 2030. As conclusões da pesquisadora foram que até lá, o campo da educação vai ser impactado pelo forte desenvolvimento:

- Do uso generalizado dos produtos do Processamento de Língua Natural (tradução, reconhecimento, geração da fala e da escrita), tanto em sistemas educacionais (*Massive Open Online Courses*, *Learning Management Systems* e Sistemas Tutores Inteligentes), quanto na Robótica.

- Da integração das tecnologias da Afetividade/Emoções, tanto nos sistemas educacionais quanto no Processamento da Língua Natural e na Robótica.
- De sistemas educacionais com um formato híbrido, contemplando a integração de aplicativos nas plataformas, similar aos ecossistemas.
- De maior utilização, nos sistemas educacionais, da visão computacional vinda da Robótica, proporcionando a identificação dos alunos em sistemas *online*.
- Para a adoção de material educacional digital personalizado, como os Smartsbooks, Objetos de Aprendizagem e Recursos Educacionais Abertos.
- Para a reconfiguração da sala de aula em espaços de aprendizagem no formato *Fab Lab*, sala de aula invertida e ambientes virtuais de ensino- aprendizagem que acompanham os alunos em qualquer local e que incorporam, além da IA, a mobilidade e as interfaces 3D com o uso de vários hardwares. (VICARI, 2018, p. 43)

Chama atenção o fato de que nos 6 itens listados, 2 citam o PLN. Ou seja, segundo esse estudo, PLN na IA é umas das principais tendências nos sistemas de aprendizagem para os próximos anos. Esses resultados foram suportados por pesquisas em bases de patentes registradas na última década. Entre as empresas que registraram patentes de sistemas e produtos tecnológicos relacionados à área de PLN e Afetividade Computacional estão a Kairos, nViso, Affective, Emotion API, EmoVoice, Vokaturi, Google, IBM, Nuance, Microsoft e Skype<sup>11</sup>. “À medida que essas tecnologias continuam a se desenvolver, a IA aplicada à Educação tem o potencial de melhorar o ensino *online* por meio dos *softwares* de aprendizagem adaptativa (personalização do ensino).” (VICARI, 2018, p. 41).

Produtos educacionais com IA que usam PLN, como os *chatbots*, ainda apresentam muitos desafios para pesquisadores e desenvolvedores de *softwares* devido à complexidade das interações em linguagem natural. Um dos principais desafios é tratar a ambiguidade presente nas interações entre seres humanos e máquinas. “Essa ambiguidade torna o PLN diferente do processamento das linguagens de programação de computador, as quais são formalmente definidas evitando, justamente, a ambiguidade”. (VIEIRA; LOPES, 2010, p. 184). Na próxima seção, discutiremos questões relacionadas às estratégias de desambiguação por *chatbots*.

## 2.2 Breve histórico da evolução dos *chatbots*

Dentre um grande número de *chatbots* já criados em todo o mundo, foram escolhidos para ilustrar esta seção *chatbots* que trouxeram algum incremento linguístico ou nova funcionalidade

---

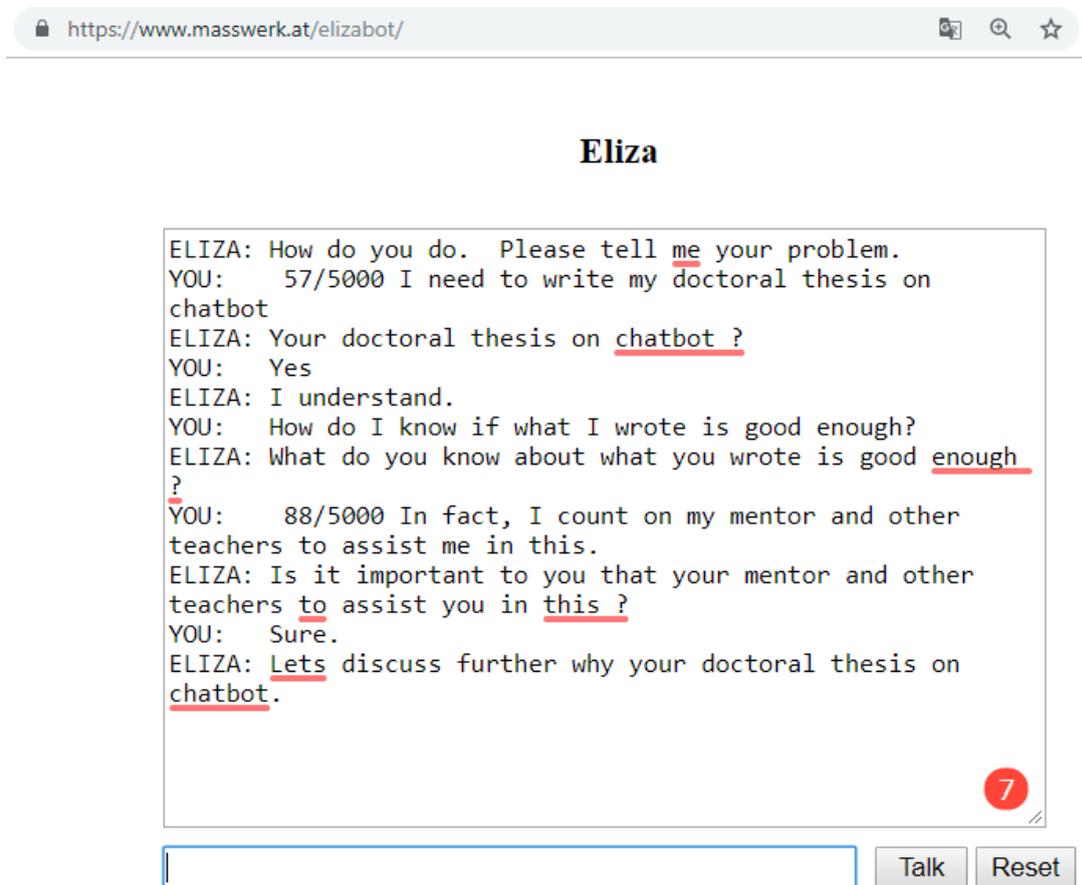
<sup>11</sup> A autora não explica quais são os produtos e não há informação se essas patentes são de produtos que já estão acessíveis ao público ou não.

e, assim, contribuíram para o desenvolvimento dessa tecnologia como se conhece atualmente.

O primeiro *chatbot* encontrado na literatura sobre o tema é o Eliza, desenvolvido no Instituto de Tecnologia de Massachussetts pelo professor Josep Weizenbaum (1966). O objetivo era que o *chatbot* simulasse a conversa de uma psicanalista (Eliza) com um paciente (o usuário). Eliza é considerado um sistema pequeno, com aproximadamente apenas 200 linhas de código, e basicamente estimula por meio de perguntas ou afirmações que o usuário fale mais sobre os seus problemas. A arquitetura de *software* de Eliza é formada por uma base de regras e um programa que funcionam da seguinte maneira: 1) o usuário envia um input textual; 2) O programa lê o texto do usuário e identifica a regra a ser aplicada para respondê-lo; 3) Eliza responde o usuário e o ciclo pode se repetir. A base de regras é o que faz que Eliza consiga interagir simulando ser uma psicanalista.

Segundo Leonhardt (2005, p. 24), as regras de Eliza se baseiam em decomposição e recomposição dos textos (*inputs*) fornecidos pelo próprio usuário. Eliza tenta encontrar as palavras-chave das sentenças e também entender um contexto mínimo sobre elas: “a palavra ‘você’ seguida da palavra ‘é’ são consideradas uma afirmação”. Ao encontrar uma palavra-chave e entender um contexto mínimo, Eliza modifica a frase usando uma de suas regras, como, por exemplo, perguntar o porquê da pessoa pensar aquilo que ela acabou de escrever. Quando Eliza não consegue reconhecer palavras-chave ou o contexto, a regra é o envio de uma mensagem genérica como “Me fale mais sobre isso”. Assim, ela estimula a continuação do diálogo (LEONHARDT, 2005).

Segundo Lima (2014, p. 40) “por ter sido o primeiro projeto de simulação de bate-papo disponibilizado publicamente, esse *chatbot* é até hoje considerado a “mãe” de todos os *chatbots*”. Eliza continua disponível em diferentes sites que integraram seus códigos, então ainda é possível testá-la. A figura 3 mostra um teste de interação feito pelo autor desta tese com Eliza.

FIGURA 4 - Teste de interação com Eliza<sup>12</sup>

Fonte: Diálogo construído pelo autor em interação com Eliza no site <https://www.masswerk.at/elizabot/#EL>

Eliza tem algumas limitações, como não ter memória. Ou seja, ela não aumenta o seu banco de dados à medida que interage com novos usuários e também não se lembra do que o usuário já conversou com ela. Portanto, não pode desenvolver melhor os diálogos se baseando em

<sup>12</sup> Tradução da língua inglesa para língua portuguesa:

Eliza: Como você está? Por favor, me fale o seu problema.

Você: Eu preciso escrever a minha tese de doutorado sobre *chatbot*. Eliza: Sua tese de doutorado é sobre *chatbot*?

Você: Sim. Eliza: Entendi.

Você: Como eu posso saber se o que eu escrevi é bom o suficiente? Eliza: O que você sabe sobre o que você escreveu ser bom o suficiente?

Você: Na verdade, eu conto com o meu orientador e outros professores para me auxiliarem nisso. Eliza: É importante para você que o seu orientador e outros professores te auxiliem nisso?

Você: Claro.

Eliza: Vamos discutir mais porque a sua tese de doutorado é sobre *chatbot*.

conversas anteriores. Apesar das limitações, as regras de decomposição e recomposição de diálogos serviram como base para o desenvolvimento de outros *chatbots* e continuam sendo usadas atualmente.

Na década seguinte, em 1971, foi criado por Colby o *chatbot* Parry, que simula uma pessoa com comportamento paranoico. A ideia é que os diálogos escritos por Colby sempre simulem um ar de mistério ou de paranoia (LEONHARDT, 2005). O banco de dados de Colby é muito maior que o de Eliza, com cerca de 6 mil padrões textuais que podem ser combinados aos *inputs* dos usuários, mantendo o padrão de decomposição e recomposição de diálogos já visto em Eliza. A simulação de emoções foi uma grande contribuição de Colby para os experimentos seguintes. De acordo com as palavras-chave identificadas, o *chatbot* tentava simular emoções como raiva, medo ou desconfiança na escrita de suas mensagens.

Em 1988, o programador inglês Rollo Carpenter criou o Jabberwacky, inteligência artificial que também tinha o objetivo de passar no Teste de Turing, mas diferenciando-se dos anteriores por usar humor nas construções textuais.

Em 1994, Michael Maldin publicou o *chatbot* Júlia como recurso do jogo digital Tynymud. Júlia tinha o objetivo de responder aos jogadores perguntas sobre outros jogadores, objetos do jogo e caminhos para se deslocar no jogo. Assim como Parry, ela também tenta demonstrar emoções. “Ela assume um papel feminino e, num determinado período do mês, ela afirma estar com tensão pré-menstrual e, inclusive, seu humor podia ser alterado”, explica Almeida Júnior (2017, p. 37).

A maioria dos padrões existentes na base de conhecimento de JULIA tinha múltiplas respostas associadas, que eram mostradas para o usuário de forma randômica e não repetida. (...) Julia tem algumas limitações: repetição de algumas perguntas após um tempo decorrido; o domínio sobre o que Julia podia falar era pequeno, assim como seu vocabulário; como Julia tinha um conhecimento específico sobre o domínio onde ela atuava, os usuários acabavam reconhecendo que ela era um robô, por não conseguir responder a perguntas fora desse domínio específico. (ALMEIDA JÚNIOR, 2017, p. 37)

Um dos avanços que Julia representou para a evolução dos *chatbots* foi o fato de ter memória. Ela guardava em seu banco de dados as conversas com os usuários e podia acessar essas informações para desenvolver melhor diálogos seguintes.

Alice, ou *Artificial Linguistic Internet Computer Entity*, é um *chatbot* criado por Wallace em

1995 e que foi disponibilizado com código aberto. Por isso, recebeu contribuições de desenvolvedores de diferentes países e tem o seu código usado para a criação de novos *chatbots* desde então. Uma grande contribuição de Alice tem sido à utilização da linguagem AIML (*Artificial Intelligence Markup Language*), desenvolvida para proporcionar a criação de diálogos semelhantes à linguagem natural (ALMEIDA JÚNIOR, 2017). Assim, Alice proporciona uma sofisticação de PLN, usando processamento morfológico, sintático e semântico, por exemplo. Alice também tem uma base de regras muito maior que os *chatbots* antecessores, com mais de 50 mil regras para definir as interações com usuários.

Em 2006, a IBM lançou a plataforma de serviços de inteligência artificial Watson, que teve repercussão mundial ao vencer seres humanos na edição de 2011 de um programa de perguntas e respostas da TV norte-americana chamado Jeopardy. Essa plataforma passou a ser usada por muitas instituições para criação de *chatbots* para atendimento ao público. O Watson é a plataforma usada para a produção de Katri.

### **2.3 Desambiguação no Processamento de Linguagem Natural por *Chatbots***

A ambiguidade ou duplicidade de sentidos presentes no texto pode confundir o leitor ou gerar interpretações distintas das que foram pretendidas pelo autor. Na interação com a máquina, como no caso dos *chatbots*, não há uma situação em que o leitor tem a chance de conversar com outro ser humano para tentar esclarecer o sentido do texto. Por isso, a desambiguação torna-se um tema relevante ao se construir um *chatbot*.

Entre os tipos de ambiguidade estão a lexical, quando a duplicidade de sentidos resulta dos diferentes significados de uma mesma palavra ou unidade lexical, como nos exemplos das palavras “banco” (que pode se referir ao banco de uma praça ou a uma instituição bancária) ou “manga” (que pode ser uma fruta ou uma parte de uma camisa); e a sintática, que ocorre devido às funções que determinados termos exercem nas orações, como no exemplo da frase “A mulher viu o acidente do carro”. Neste caso, a mulher pode ter visto como aconteceu um acidente automobilístico ou ela, de dentro de um carro, pode ter visto algum acidente que acontecia em outro lugar. Há diversos outros tipos de ambiguidade textual, como a pragmática, a semântica,

a morfológica ou a predicativa<sup>13</sup> e fatores causadores da ambiguidade, como a ambivalência, o duplo sentido, a indeterminação e vagueza<sup>14</sup>. Segundo Demarques (2011), muitas vezes a ambiguidade não é percebida pelo autor dos textos e o problema é passado para o leitor-interpretador que terá três possíveis soluções para resolvê-lo:

a) a ambiguidade é resolvida pelo leitor por algum mecanismo desambiguador, usando dados que estão além da frase; b) a ambiguidade não é resolvida pelo leitor, com prejuízo da compreensão do texto; e c) a ambiguidade é resolvida erroneamente pelo leitor, com a adoção do significado não pretendido pelo autor. (DEMARQUES, 2011, p. 40)

Em uma situação em que o usuário/aluno interage com o que o autor (professor/produtor de *chatbot*) programou, não há a possibilidade de solução imediata de problemas de ambiguidade pelo autor. Corre-se o risco, então, de haver ruído na comunicação. Como solucionar ou diminuir problemas causados pela ambiguidade na interação entre um ser humano e um *chatbot*, criando mecanismos de desambiguação programáveis, é então uma questão relevante para se refletir. Entre as técnicas de desambiguação utilizadas em *chatbots* estão as análises de intenções, a análise de contexto e a resolução de correferência (INBOT, 2019).

### 2.3.1 Análise de intenções

Em relação ao método de análise de intenções, Cho et al (2019) explicam que estudar e prever as possíveis intenções de fala dos usuários é uma importante ação para que os *chatbots*, possam fornecer respostas adequadas em diálogos. A capacidade de reconhecimento de intenções, por meio das mensagens enviadas por um ser humano a uma máquina, é um ponto chave para o sucesso desse tipo de sistema. Com base em Allen et al (2001), Pardo (2005) e Neves (2005), entende-se no contexto desta pesquisa que intenção é o objetivo da interação do usuário, considerando-se o efeito pretendido na ação de resposta da máquina.

Para Allen et al (2001), para que um sistema de conversação seja capaz de promover uma interação mais realista (ou menos distante da interação entre seres humanos) é necessário que a tecnologia utilizada consiga prover mecanismos para reconhecimento de intenções. Os autores

---

<sup>13</sup> Ver Demarques (2011).

<sup>14</sup> Ver Silva (2006).

indicam que esses mecanismos precisam identificar qual obrigação (ou problema para a máquina resolver) emerge a partir da intenção do usuário. Assim, o sistema tenta entregar uma resposta relacionada à intenção, resolvendo a obrigação ou problema identificado.

Vieira e Lima (2001, p. 2) explicam que, como uma sentença pode expressar uma intenção do usuário, é necessário “especificar algoritmos capazes de estabelecer relações entre os diversos componentes e segmentos de um texto”. Para isso, o banco de dados do sistema precisa conter um campo específico no qual estejam classificadas as intenções que são possíveis de se prever que os usuários usarão nas interações. Assim, com essa programação, o sistema é capaz de reconhecer a intenção, no caso de o usuário ter utilizado uma intenção já classificada. Nesse caso, o *chatbot* busca em seu banco de dados a resposta mais adequada. Sem algoritmos programáveis capazes de identificar corretamente as intenções, a interação se torna menos eficiente.

Uma maneira encontrada por desenvolvedores de *chatbots* para amplificar o sucesso na identificação das intenções é criar uma correlação entre a classificação de intenções e entidades. Segundo Oliveira e Marques (2018, p. 8) “as entidades representam uma classe de objeto ou um tipo de dado que é considerado relevante para o propósito do usuário”.

A classificação de entidades auxilia na desambiguação e no entendimento com mais acurácia de qual é de fato a intenção do usuário. Por exemplo, se o usuário pergunta ao *chatbot* “Qual é a data da prova de matemática?”, o sistema poderia identificar “prova de matemática” como intenção, se ela está classificada como uma das intenções possíveis no banco de dados do sistema. O *chatbot* poderia responder com um dos vários conteúdos relacionados à prova de matemática que fazem parte do seu banco de dados, como os conteúdos da prova, local em que ela será realizada, referências de materiais para estudar, dentre outras opções. Entretanto, se além das intenções há entidades cadastradas e uma delas é “data”, o sistema poderá combinar a intenção “prova de matemática” com a entidade “data” e, assim, chegar a uma resposta possivelmente correta, respondendo ao usuário a data dessa prova. Se o usuário em seguida pergunta “Qual é o conteúdo da prova de matemática?”, o sistema pode buscar a mesma intenção, mas combinando-a com a entidade “conteúdo” e achar outra resposta em seu banco de dados para fornecer ao usuário.

É comum também o cadastro de sinônimos para as entidades e intenções. Assim, a intenção

“prova de matemática” poderia ter como sinônimos “teste de matemática” ou “avaliação de matemática”, enquanto a entidade “conteúdo” poderia ter equivalência a outras entidades como “matéria” e “o que vai cair”. Assim, diferentes maneiras de escrever uma pergunta podem ser reconhecidas pelo *chatbot*. Questões que combinam as entidades e intenções, como “O que vai cair no teste de matemática?”, “Qual o conteúdo da avaliação de matemática?” ou “Quero saber a matéria da prova de matemática” seriam reconhecidas pelo sistema e levariam à mesma resposta. Portanto, inferir e alimentar o banco de dados de um *chatbot* com o maior número de intenções e entidades é uma estratégia de desambiguação para aumentar as chances de sucesso das interações.

### 2.3.2 Análise de contexto

A análise de contexto é um método para desambiguação em *chatbots*, mas considera-se importante relativizar o uso da palavra “contexto”, no atual estágio de desenvolvimento das tecnologias de inteligência artificial e de processamento de linguagem natural usadas em *chatbots*. Questiona-se o quanto os *chatbots* são capazes, de fato, de entender o contexto, considerando a situação de comunicação, ou se têm capacidade de entender apenas o cotexto, que segundo Mmaingueneau (2013) é o contexto linguístico. Ou seja, os *chatbots* podem entender o discurso de maneira mais ampla ou apenas a língua na perspectiva do processamento semântico?

Ao pesquisar o que desenvolvedores de *chatbot* dizem sobre a análise de contexto, percebe-se que o conceito é usado para falar sobre o contexto linguístico. Sindhu Joseph (2018) explica esse recorte ao tratar da importância de um *chatbot* entender o contexto do discurso dos usuários. Ao exemplificar o que significa contexto, a autora diz que no jargão bancário a palavra da língua inglesa “*interest*” pode se referir à remuneração cobrada em intervalos regulares devido a um empréstimo de dinheiro ou ao interesse por saber mais sobre um produto ou serviço. O *chatbot*, então, precisa ter estratégias, como relacionar “*interest*” com as outras palavras presentes no diálogo, para entender se o usuário se refere a juros ou curiosidade. Não se consideram outros elementos contextuais, como o tom de voz e a expressão facial do usuário, mas apenas o cotexto/contexto linguístico.

Joseph (2018), explica que há dificuldades para que as experiências conversacionais com

*chatbots* sejam significativas e personalizadas devido ao entendimento limitado do contexto de uma conversa por parte do robô de conversação. A autora acredita que a evolução dos *chatbots* levará a entendimentos mais amplos dos contextos. Uma solução possível apresentada por ela é o *chatbot* “se lembrar” de conversas anteriores com um mesmo usuário e, assim, aprimorar o entendimento do que é redigido por ele.

Para Makadia (2019), o PLN de um *chatbot* é baseado nas aprendizagens que a máquina tem com produtores e usuários, pois o *chatbot* aprende mais contextos linguísticos a medida em que interage com mais pessoas, o que faz seu banco de dados aumentar (pois guarda e processa os *inputs* que recebeu dos usuários), ter mais informações e assim aprimorar o entendimento completo de uma sentença. Segundo o autor, num diálogo, o *chatbot* é capaz de avaliar a intenção do texto de um usuário e criar respostas específicas com base em análise do contexto, caso ele já tenha recebido *inputs* suficientes para conseguir processar essa análise. Makadia (2019) afirma que embora o PLN não tenha o nível de compreensão da linguagem comparável ao de um ser humano, um *chatbot* pode imitar bem esse nível de entendimento. O contexto linguístico do *chatbot* auxilia na desambiguação, mas é restrito apenas ao repertório linguístico que já foi adicionado ao banco de dados do sistema.

Entende-se nesta pesquisa que o método de análise de contexto é uma análise cotextual no atual estágio tecnológico em que se apresentam os *chatbots* como Katri, que será objeto do estudo de caso apresentado adiante.

### 2.3.3 Resolução de correferência

Outra técnica de desambiguação utilizada em *chatbots* é a resolução de correferência. Zhu et al (2018) e Agrawal et al (2017) apresentam exemplos de *chatbots* que utilizam a resolução de correferência como estratégia para que a máquina consiga buscar em seu banco de dados anáforas e antecedentes como estratégia para melhorar o PLN.

A correferência ocorre quando duas ou mais expressões se referem ao mesmo termo em um texto. Anáfora e antecedente são conceitos essenciais para se entender a correferência. Anáfora é a retomada de um termo já referenciado anteriormente num texto. O antecedente é exatamente esse termo retomado por meio da anáfora. Ou seja, a anáfora e o antecedente se

correferenciam. Por exemplo:

O aluno estudou muito para a prova, porém não recebeu uma boa nota.

Ele justificou que o teste estava muito difícil.

Nas sentenças acima, “O aluno” e “a prova” são respectivamente os antecedentes de “Ele” e “o teste”, termos retomados por meio de anáfora. Para Moraes (2010, p.17), a identificação ou resolução automatizada de correferência em textos por sistemas computacionais é uma tarefa importante em vários sistemas de PLN, “pois dela depende uma compreensão adequada dos textos”, o que vai contribuir para solucionar problemas como os de ambiguidade.

Agrawall et al (2017) definem então a resolução de correferência como a tarefa automatizada de se encontrar entidades que se referem a outras em um texto. Considera-se que a resolução de correferência é uma complexa tarefa nos *chatbots*, pois pode ser necessário encontrar correferências em diferentes partes do diálogo com o usuário ou em diferentes textos que fazem parte do banco de dados do sistema.

É importante salientar que todos os métodos de desambiguação apresentados são aplicados por meio de programação computacional, mas também pelo treinamento do *chatbot*, que se dá por meio de informações escritas por seres humanos (produtores e usuários do sistema) e inseridas em seu banco de dados. Esses métodos de desambiguação precisam ser previstos pelos produtores para que o cadastro de entidades, intenções e a produção dos textos de respostas que o *chatbot* dará aos usuários já contemplem esses métodos.

#### **2.4 Chatbots usados na educação**

Nos últimos anos, o uso de *chatbots* em instituições de ensino tem se tornado uma realidade. Pérez, Daradoumis e Puig (2020) propõem que os *chatbots* usados na educação sejam divididos em duas subcategorias, de acordo com suas finalidades: *service-oriented chatbot* ou *teaching-oriented chatbots*.

Na primeira categoria, *service-oriented chatbot*, estão os *chatbots* que dão suporte aos serviços,

como uma espécie de FAQ (perguntas frequentes). Eles proveem informações, por exemplo, sobre processos de matrícula, emissão de certificados, cursos ofertados pela instituição etc. Esta abordagem orientada a serviços, segundo os autores, já encontrou sucesso em outras áreas, como no uso de *chatbots* por instituições bancárias e na medicina.

No Brasil, *Chatbot Admin* (SILVA, 2021) é um *chatbot* orientado a serviços do curso de graduação em Sistemas de Informação da Universidade Federal de Santa Catarina. Ele responde aos alunos questões sobre realização e trancamento de matrículas, grade de horários de aulas, nomes do coordenador de curso e professores das disciplinas ofertadas e endereço da universidade.

Na Espanha, *Lola* (REN et al, 2019) é um *chatbot* usado na Universidade de Murcia como uma assistente virtual para auxiliar os alunos nos processos de pré-inscrição e matrículas nos cursos da instituição.

Na Itália, *Lisa* (Debitonto et al, 2018) é um *chatbot* da Link Campus University para fornecer informações de serviços para os estudantes. Em pesquisa realizada com alunos no processo de criação do *chatbot*, descobriu-se que entre as informações que os alunos mais gostariam de receber de *Lisa* estavam notificações sobre as aulas, oportunidades emprego e estudos no exterior, eventos, atividades extracurriculares e agendamento de avaliações.

Na segunda categoria apresentada por Pérez, Daradoumis e Puig (2020), *teaching-oriented chatbots*, estão os *chatbots* que dão suporte a processos de aprendizagem, que atuam na construção de conhecimento junto aos estudantes. Segundo os autores,

Houve um aumento considerável do número de *chatbots* voltados para o ensino, isto é, cujo principal propósito é gerar conhecimento para alunos específicos, geralmente a respeito de um tópico específico. Tais *chatbots* são usados tanto em contextos educacionais formais quanto informais. O objetivo desses *chatbots* é gerar conhecimento da maneira como um tutor humano faria <sup>15</sup> (PERÉZ, DARADOUMIS; PUIG, 2020, p. 5, tradução minha<sup>16</sup>)

---

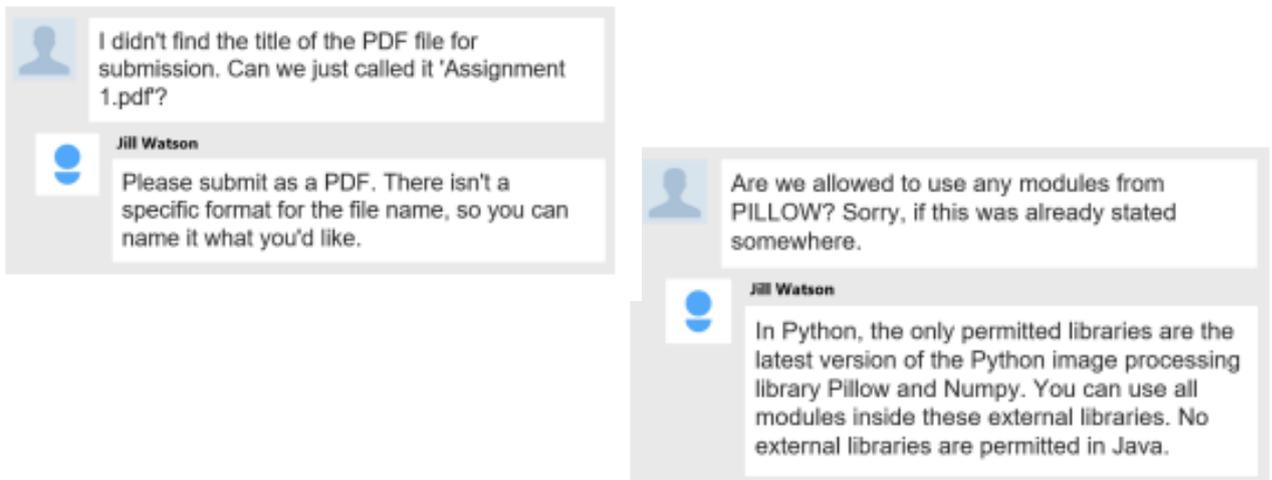
<sup>15</sup> Original: “There is a considerable increase in the number of *chatbots* whose purpose is teaching, that is, *chatbots* whose main purpose is to generate knowledge for specific students, usually on a specific topic. Such *chatbots* are used in formal education and nonformal education. The purpose of these *chatbots* is to generate knowledge like a human tutor would.”

<sup>16</sup> Exceto quando indicado o nome do tradutor ou tradutora na bibliografia, todas as outras traduções nesta tese são de minha autoria.

É nesta categoria que se encontra Katri, caso estudado nesta tese, pois mesmo considerando-a um material didático e assumindo o posicionamento de que um *chatbot* não deve ser encarado como um substituto de um tutor ou professor humano, reconhece-se que Katri contribui para a construção do conhecimento, assim como outros materiais didáticos o fazem. Desta maneira, Katri não é orientada a serviços, mas ao processo de ensino e aprendizagem.

Como exemplo de *chatbot* classificado nesta categoria, nos Estados Unidos, Jill Watson (EICHER; POLEPEDDI; GOEL, 2018) é um *chatbot* usado no curso de Inteligência Artificial do Georgia Institute of Technology. Na interação com os alunos, ele é capaz de tirar dúvidas sobre conteúdos do curso e também explicar questões da rotina de entrega de atividades, como pode ser visto Figura 5.

**FIGURA 5 - Diálogos de estudantes com o Jill Watson<sup>17</sup>**



Fonte: Eicher, Polepeddi e Goel (2018)

Na Tailândia, ScratchThAI (KATCHAPAKIRIN; ANUTARIYA, 2018) é um *chatbot* usado para ensinar linguagem de programação a crianças como parte das aulas de pensamento computacional em diferentes escolas. O *chatbot* tira dúvidas dos alunos, fornece conteúdos e

<sup>17</sup> Tradução do inglês para o português:

Aluno: Eu não encontrei o título do arquivo PDF para enviar. Posso nomeá-lo só de 'Tarefa 1.pdf'?

Jill Watson: Envie no formato PDF. Não há um formato específico para o nome do arquivo, então você pode nomeá-lo como quiser.

Aluno: Podemos usar algum módulo do PILLOW? Desculpe se essa informação já está disponível em algum outro lugar.

Jill Watson: Em Python, as únicas bibliotecas permitidas são a última versão da biblioteca de processamento de imagens do Python, Pillow e Numpy. Você pode usar todos os módulos dentro dessas bibliotecas. Nenhuma biblioteca externa é permitida em Java.

indica fontes de estudos.

Por meio dos exemplos citados, implementados nos últimos anos em diferentes países e continentes, é possível perceber como o uso dos *chatbots* na educação é uma tendência global. Na revisão de literatura realizada por Pérez, Daradoumis e Puig (2020), foram encontradas 80 publicações sobre *chatbots* usados na educação, sendo que nelas são citados exemplos de 18 diferentes *chatbots* (há diferentes publicações que citam os mesmos *chatbots*), sendo 8 *service-oriented chatbots* e 10 *teaching-oriented chatbots*.

Nesta revisão de literatura feita por Pérez, Daradoumis e Puig (2020), há publicações que não citam exemplos de *chatbots*, mas discorrem sobre as possibilidades de uso e até mesmo sobre avaliação de *chatbots*, tema desta pesquisa. Gonda et al (2018) sugerem que *chatbots* sejam avaliados de acordo com princípios de para uma boa docência (*Principles for Good Teaching*), como encorajar a aprendizagem ativa e dar feedbacks aos alunos. Bahja, Hammad e Butt (2020) apresentam um framework de design de *chatbots* educativos centrado nos usuários, porém não apresentam uma proposta de matriz para ser usada na avaliação desses *chatbots* educativos. Esses estudos são, entretanto, referências para a reflexão sobre avaliação de *chatbots* usados na educação.

Peréz, Daradoumis e Puig (2020) explicam que a maneira mais comum de avaliar os *chatbots* usados na educação é por meio do uso de questionários de percepção dos estudantes sobre a qualidade dos *chatbots*, mas dentre todas as publicações que encontraram não há uma proposta de matriz com critérios de avaliação de *chatbots* didáticos. Os autores, inclusive, recomendam que futuras pesquisas busquem preencher esta lacuna, podendo “usar este estudo como referência para entender o potencial que tutores conversacionais têm ao atuar como assistentes educacionais em diferentes níveis e para obter diferentes critérios de avaliação desses *chatbots*<sup>18</sup>” (p. 13). Portanto, a recomendação dos autores se conecta diretamente com os objetivos desta tese.

No próximo capítulo, apresenta-se a matriz proposta nesta pesquisa para avaliação de *chatbots* didáticos, como Katri, assim como as referências que foram base para a criação dessa matriz.

---

<sup>18</sup> Original: “refer to this study to understand the potential that conversational tutors can have as teaching assistants at different educational levels and to obtain different evaluation criteria of these *chatbots*.”

## 2.5 *Chatbots* didáticos no contexto da pandemia de Covid-19

A pandemia de Covid-19<sup>19</sup>, além de todas as sérias questões relacionadas à saúde pública, trouxe relevantes consequências para a educação. De acordo com a ONU (2021), depois de mais de um ano de pandemia, em 2021, mais de 800 milhões de estudantes ainda estavam em situação atípica relacionada aos espaços e oportunidades de aprendizagem, com instituições de ensino parte fechadas, parte funcionando remotamente e parte funcionando em tempo parcial, em modalidade presencial ou online.

No início da pandemia, em março e abril de 2020, Marinoni, Land e Jensen (2020)<sup>20</sup>, conduziram um estudo para verificar como 424 universidades de 109 países estavam lidando com as restrições impostas pela Covid-19. Apenas 2% dos respondentes afirmaram que não houve mudanças no processo de aprendizagem, enquanto 7% disseram que todas as aulas haviam sido canceladas, 24% informaram que as aulas estavam suspensas enquanto a universidade planejava soluções digitais ou de estudos autônomos para continuação dos estudos e 67% disseram que as aulas já haviam sido migradas para formatos com o uso de tecnologias digitais. Este estudo mostra o quanto as instituições de ensino precisaram avançar rápida e significativamente na adoção de tecnologias educacionais.

Para Pantelimon et al. (2021), essa pandemia mudou e reescreveu as regras sobre como os sistemas de educação funcionam, assim como em que espaços a aprendizagem acontece e como estudantes e professores interagem entre si. E diante de todas as dificuldades para se manter as aulas presenciais, a pandemia potencializou o uso dos sistemas digitais voltados à educação, assim como acelerou o processo de implementação de *chatbots* didáticos.

Sweidan et al. (2021) explicam que já planejavam o desenvolvimento de um *chatbot* para uso dos mais de 40 mil alunos da Univeristy of Jordan (Jordânia), mas que a pandemia de Covid-19 fez com que o desenvolvimento do sistema precisasse ser acelerado. Eles afirmam que essa aceleração para uso de *chatbots* aconteceu em todo o mundo durante a pandemia. Os autores desenvolveram o SIAAA-C (Student Interactive Assistant Android Application with *Chatbot*),

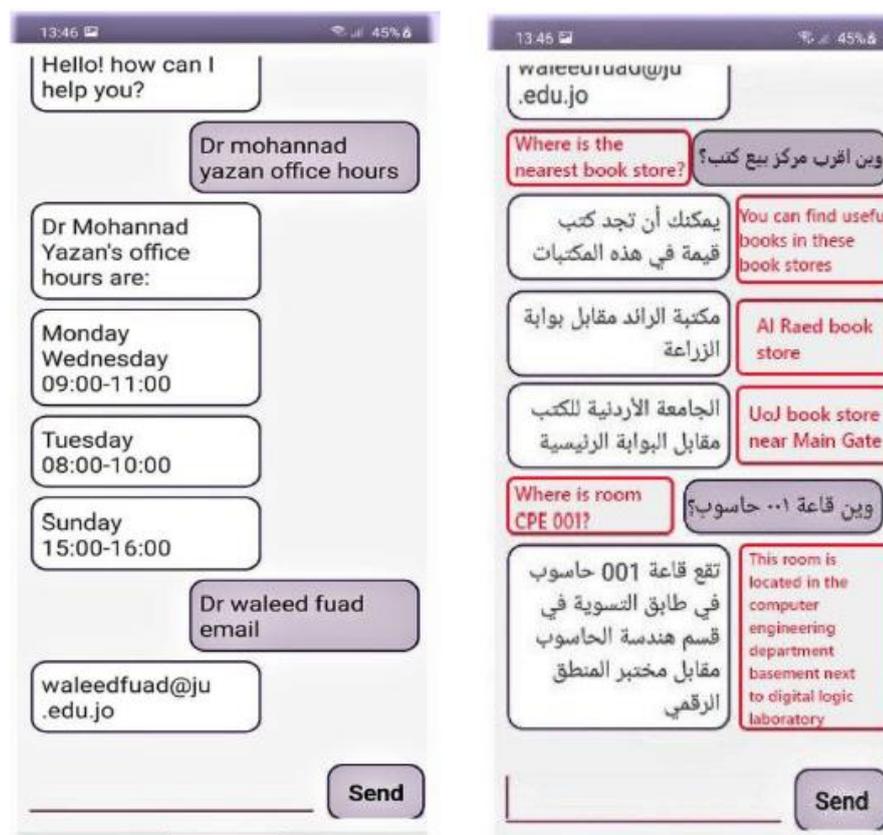
---

<sup>19</sup> Para ler sobre o contexto da pandemia de COVID-19: <https://www.paho.org/pt/covid19/historico-da-pandemia-covid-19> e [https://pt.wikipedia.org/wiki/Pandemia\\_de\\_COVID-19](https://pt.wikipedia.org/wiki/Pandemia_de_COVID-19)

<sup>20</sup> A pesquisa de Marinoni, Land e Jensen (2020) mostra também que há uma correlação entre as chances de professor adotarem o uso de *chatbots* e outras tecnologias com inteligência artificial e a vontade de serem vacinados contra COVID-19. Segundo os resultados da pesquisa, professor e alunos que querem ser vacinados estão mais propensos a usarem esse tipo de tecnologia que os que não querem se vacinar.

um *chatbot* bilíngue que funciona em árabe e em inglês para atender também aos alunos estrangeiros da universidade. Ele auxilia os alunos com a agenda de aulas e testes avaliativos, informações sobre cursos disponíveis, informações de contato de professores e outros colaboradores da universidade, assim como o mapa do campus e dicas sobre onde comer e encontrar livros. O *chatbot* também dá sugestões de cursos para os alunos se matricularem, a partir da análise automatizada dos pré-requisitos necessários e a correlação com os componentes curriculares já cursados pelo aluno.

FIGURA 6 - Exemplos de troca de mensagens em inglês e árabe no SIAAA-C



Fonte: Sweidan et al. (2021, p. 1736)

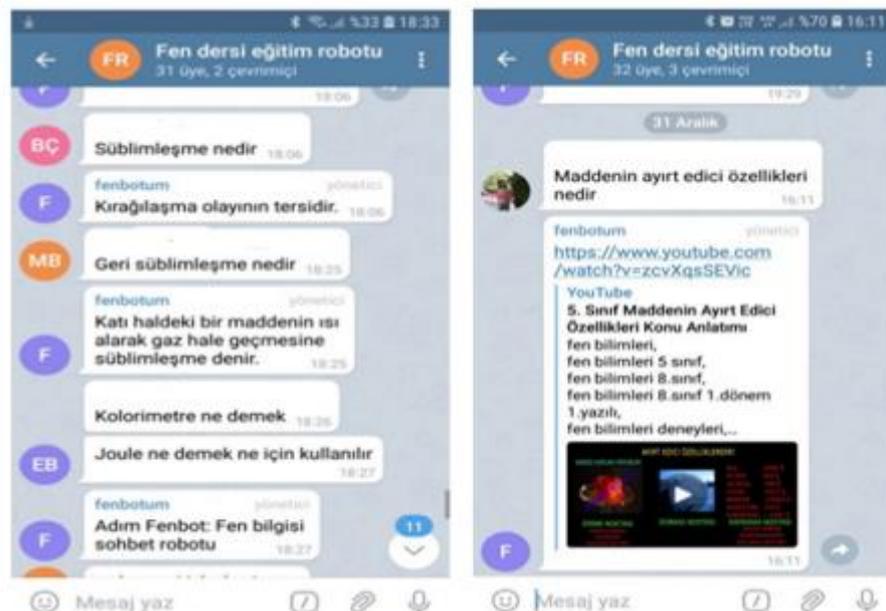
Pesquisa feita com 102 estudantes da University of Jordan mostrou que 80% deles preferiu tirar dúvidas com o *chatbot* a com atendentes humanos durante a pandemia. Fatores como a maior disponibilidade de horários e a acurácia das respostas foram considerados os mais significativos para essa preferência. Segundo Sweidan et al. (2021), a situação da pandemia confirmou a importância desse tipo de sistema nas instituições de ensino.

Ranjan et. al. (2021) conduziram pesquisa com alunos de universidades indianas para entender suas percepções sobre o uso de tecnologias aplicadas à educação durante a pandemia de Covid-

19. Os resultados mostram que 61% dos estudantes acreditam que o uso de *chatbots* como assistentes do processo de ensino e aprendizagem é uma boa alternativa, o que reforça o argumento de Sweidan et al. (2021).

Topal, Eren e Geçer (2021) apresentam uma experiência de uso de *chatbot* didático na educação básica, especificamente com alunos do 5º ano de uma escola da cidade de Darica, na Turquia, para aprendizagem de Ciências. O *chatbot* enviava mensagens em texto escrito e em vídeo para os alunos. Os pesquisadores trabalharam com um grupo de 41 estudantes que usaram o *chatbot* e um grupo controle, que não usou o *chatbot*, para aprender os mesmos conteúdos. Os resultados mostraram uma pequena diferença nos resultados de aprendizagem, mais positiva para o grupo que usou o *chatbot*, e os pesquisadores ressaltam que mais significativo foi o quanto o *chatbot* afetou positivamente a experiência de aprendizagem dos alunos no estudo remoto durante a pandemia. Esses alunos disseram que o *chatbot* foi útil, divertido e que gostariam de usar em mais disciplinas escolares. Ressaltaram, também, que foi útil ter um assistente de estudos virtual e poder assistir repetidamente os materiais de aulas que o *chatbot* enviava.

**FIGURA 7 - Chatbot envia conteúdos por texto escrito e vídeo aos alunos em escola turca**



Fonte: Topal, Eren e Geçer (2021, p. 6250)

No Brasil, durante a pandemia, escolas de todo o país também passaram a usar sistemas inteligentes, como os *chatbots*. Redes de ensino, como a estadual de São Paulo, usaram sistemas como o AprendiZap, *chatbot* que funciona no aplicativo de trocas de mensagens WhatsApp e

que em agosto de 2021 já tinha conteúdos de sete disciplinas da educação básica: língua portuguesa, língua inglesa, matemática, história, geografia, ciências e artes. Os alunos recebiam textos, vídeos e outros materiais didáticos, além de responderem a testes sobre o conteúdo por meio da interação com o *chatbot* (RIGA, 2021).

Pantelimon et al. (2021) afirmam que a Covid-19 tornou possível introduzir nas escolas mais rapidamente sistemas avançados de inteligência artificial, como os *chatbots*, e que o uso de sistemas de inteligência artificial na educação se tornou um tema de maior interesse na comunidade científica. Assim, com a pandemia, é possível verificar exemplos como os apresentados previamente que demonstram o rápido avanço nas ações práticas e nos estudos sobre *chatbots* usados em contextos de educação.

### **3. CHATBOT DIDÁTICO: CRITÉRIOS PARA AVALIAÇÃO**

Como avaliar *chatbots* didáticos? Por não terem sido encontradas referências de matrizes de avaliação de *chatbots* usados para fins educacionais, buscou-se referências de critérios de avaliação de materiais didáticos, especialmente os interativos *online*. Referências sobre livros didáticos também foram usados nesta pesquisa, pois devido ao protagonismo que esse tipo de MD têm nas práticas pedagógicas (e possivelmente continuará tendo, convivendo com os MD essencialmente digitais), encontram-se mais estudos sobre critérios de avaliação de livros didáticos que de outros MD. Também foram pesquisados critérios para avaliação de *chatbots* usados para fins não pedagógicos. Com base em todas essas referências, na última seção do capítulo é apresentada uma proposta de matriz de avaliação de *chatbots* didáticos, que no capítulo de resultados será validada a partir da experiência com Katri no curso TLHE.

#### **3.1 Critérios para Avaliação de Materiais Didáticos**

Alguns autores, como Cabero (2001) falam sobre critérios de avaliação de MD. Outros, como Belisário (2003), apresentam critérios para a produção de MD. Considera-se que critérios de avaliação e de produção de MD têm convergências, visto que os critérios que guiam a produção influenciam a avaliação da qualidade desses materiais, assim como os critérios para avaliação servem como guia para a sua produção. Opta-se nesta pesquisa pelo foco nos critérios para avaliação de MD, mas sem desconsiderar que eles podem influenciar também a produção.

Tomlinson e Masuhara (2004) sugerem que a avaliação de MD seja feita em até três momentos: antes de utilizar os materiais, durante o uso e/ou após o uso. Geralmente, o primeiro momento, antes da utilização dos materiais “envolve a elaboração de previsões sobre o valor potencial dos materiais para seus usuários (...) baseando-se em meras impressões pessoais, o que é subjetivo e pouco confiável” (TOMLINSON; MASUHARA, 2004, p. 5). Por isso, segundo os autores, é importante que seja utilizada uma lista de critérios para tornar a avaliação mais objetiva e confiável.

Já a avaliação no momento de uso dos MD envolveria a observação em tempo real, o que contribui para valorar o real impacto e qualidade dos materiais no momento em que professores

e alunos interagem com ele. O momento final, após o uso do material, é considerado o mais importante por Tomlinson e Masuhara (2004, p. 9), “pois pode medir os efeitos reais dos materiais nos usuários”. Essas duas fases também devem ser feitas seguindo listas de critérios.

Para Cabero (2001, p. 260), os resultados da avaliação de um MD indicam as melhores formas de usar o material e as necessidades de revisão, ou seja, indicam questões relacionadas a sua produção. Para ele, a avaliação de MD “se refere à emissão de um julgamento de valor sobre a qualidade científico-técnica e estética do meio”.

As principais estratégias de avaliação de MD são divididas por Cabero (2001) em três categorias: 1 – autoavaliação pelos produtores; 2 - consulta a especialistas externos; 3 - avaliação pelos usuários. Elas não se opõem umas às outras e podem ser combinadas, usando-se duas ou até mesmo as três categorias para se avaliar um MD. O autor concorda com Tomlinson e Masuhara (2004) ao afirmar que há avaliações que podem ser feitas pré-uso, como a pelos produtores, e outras em uso ou pós-uso, como a avaliação feita pelos usuários.

Richter (2006, p. 5) explica que a avaliação pré-uso “embora provavelmente a mais comum, é também a modalidade mais difícil”, pois ainda não há experiências de uso real do MD que sirvam de base para as decisões do avaliador. O avaliador “olha para o futuro”, considerando seus conhecimentos, mas há também “uma espécie de aposta”. A avaliação em uso diz respeito à avaliação do MD enquanto ele é empregado nas aulas e ajuda o educador a “decidir sobre alguma adoção ou mudança no início ou no transcorrer do curso”. A avaliação pós-uso “fornece avaliação retrospectiva da ‘pilotagem’ já realizada - e considerada completa ou suficiente - ” do MD, o que permite identificar vantagens ou desvantagens após um período de uso constante do MD e tomar decisões sobre a continuidade ou não do uso do material. A abordagem adotada nesta tese é a de avaliação pós-uso para que a instituição que oferta o MD aos seus alunos e os produtores possam tomar decisões sobre melhorias a serem implementadas e também decisões de continuidade ou descontinuidade de seu uso.

Abre-se um parêntese para destacar que, mesmo quando se assume a abordagem de avaliação pós-uso, os produtores, devido à natureza de seu trabalho, podem formal ou informalmente já adotar também a avaliação pré-uso, o que é recomendado por Cabero (2001, p. 267).

Não acreditamos que nenhum meio seja introduzido no contexto de utilização sem ter recebido alguma revisão de seus elementos técnicos e/ou estéticos e das suas potencialidades didáticas pelas próprias pessoas que os realizam. E podemos dizer também que a auto-avaliação é processual e começa desde a elaboração do roteiro e das decisões adotadas para a incorporação de alguns elementos e não de outros. (Cabero, 2001, p. 267)

Ou seja, se os produtores têm ciência dos critérios de avaliação do MD, eles normalmente farão análises pré-uso do material durante o processo de produção, realizando revisões e implementando melhorias que visem ao desenvolvimento de um MD de alta qualidade, mesmo que oficialmente não se esteja trabalhando com uma etapa de avaliação do MD. A Figura 6 é um fluxograma que demonstra o ciclo de ações a serem utilizadas para avaliação de MD, com base nas recomendações de Cabero (2001).

**FIGURA 8 - Fluxograma de avaliação de materiais didáticos**



Fonte: Vilaça (2013) baseado em Cabero (2001)

Trata-se de um ciclo em que primeiramente os MD são avaliados pelos produtores, por especialistas técnicos e/ou docentes externos à instituição e também pelos usuários dos MD (alunos e/ou professores). Em seguida, é feita a análise e sistematização dos resultados dessas avaliações. São identificadas as necessidades de adaptação, revisão ou mesmo de inutilização dos MD e também de eventual necessidade de criação de novos MD. Após essa etapa, é observado o processo de produção (revisão, adaptação e/ou criação) dos MD. O ciclo se repete, pois ao se adaptar ou criar novos MD, é necessário novamente avaliá-los, seguindo os mesmos procedimentos. A avaliação de Katri, relatada nesta tese, cumpriu o ciclo duas vezes. No primeiro ciclo, a avaliação da versão piloto de Katri foi feita seguindo o procedimento-padrão utilizado pela equipe de produção do *chatbot*, com o uso da metodologia *Think Aloud Test*,

conforme relato de experiência apresentado no item 6.1 desta pesquisa. No segundo ciclo, a versão definitiva de Katri foi analisada a partir do Sistema de Avaliatividade proposto por Martin e White (2005) como uma extensão da Linguística Sistêmico-Funcional (HALLIDAY; MATTHIESSEN, 2014), conforme será apresentado adiante no capítulo de Metodologia e nos Resultados desta pesquisa (item 6.2).

Cabero (2001, p. 264-266) propõe oito critérios para a avaliação de MD:

1. Conteúdos
2. Aspectos técnicos-estéticos
3. Organização interna da informação
4. Material de acompanhamento
5. Custo econômico
6. Ergonomia do meio
7. Aspectos físicos
8. Público ao qual se destina

Para cada critério, Cabero (2001) apresenta uma série de descritores a ser avaliados. A título de exemplificação, o critério Conteúdos tem como descritores: a) Qualidade científica dos conteúdos; b) Atualização; c) Conhecimentos técnicos exigidos do aluno para acompanhar o material; d) Sequenciação e estruturação corretas; e) Originalidade da forma como os conteúdos são apresentados. Os critérios apresentados acima são indicados de maneira genérica, podendo ser usados na avaliação de livros didáticos, videoaulas, atividades interativas, entre outros tipos de MD.

Em relação especificamente a critérios de avaliação de livros didáticos, há autores de obras amplamente conhecidas, consideradas clássicas, como Cunningsworth (1995). Ao mesmo tempo, pesquisas recentes como a de Bandeira, Stange e Santos (2012), apresentam novas propostas. Cunningsworth (1995) propõe uma seleção de 45 questões para avaliação de livros didáticos, cobrindo critérios como os objetivos, o design, a linguagem, os conteúdos, as habilidades trabalhadas, a metodologia e os custos. Bandeira, Stange e Santos (2012) apresentam uma lista de critérios mais curta, com 25 itens, classificadas em 4 categorias: 1) Conteúdos: aspectos teórico-metodológicos; 2) Recursos Visuais; 3) Atividades: aspectos pedagógico-metodológicos; 4) Avaliação.

Silva, Parreiras e Fernandes (2015) discutem propostas de avaliação e escolha de coleções de livros didáticos de inglês. Os autores explicam que “o livro didático deve conter o que o professor considera importante à sua prática pedagógica e o que o aluno considera indispensável ao desenvolvimento da sua competência comunicativa no idioma” (SILVA; PARREIRAS; FERNANDES, 2015, p.371). Os autores apresentam uma proposta de instrumento de avaliação qualitativa e se preocupam com a extensão dos instrumentos apresentados por outros autores, visto que no dia a dia da profissão, trabalhando normalmente dois ou três turnos por dia, os docentes não têm tempo de avaliar cada livro ou coleção que usarão baseados em listas muito extensas de critérios. Ou seja, a preocupação dos autores é criar um instrumento que permita avaliar as coleções de livros a serem ou não adotados, mas que se conecte com a realidade do profissional, considerando a pouca disponibilidade de tempo para proceder a avaliação. Assim, eles sugerem um instrumento com 21 itens, conforme mostra o Quadro 2.

**QUADRO 2 - Instrumento para avaliação de coleções didáticas de língua inglesa**

1. A coleção é linear no que se refere ao conteúdo de cada unidade? A organização é clara, coerente e funcional, do ponto de vista da sua proposta didático-pedagógica?	2. A disposição das imagens, a qualidade e legibilidade gráfica das mesmas é adequada aos alunos do Ensino Médio?	3. A coleção oferece algum material de apoio, como indicação de leitura complementar, livro digital, acesso aos OEDs (Objetos Educacionais Digitais), etc.?
4. A coleção aborda temas interessantes aos olhos dos alunos? Os assuntos presentes nos textos são apropriados para a idade?	5. A coleção contempla variedade de gêneros verbais, não verbais e verbo visuais?	6. Os textos apresentam aspectos relacionados ao mundo social, oriundos de diferentes esferas e suportes (cultural, social, étnica, etária e de gênero)?
7. A coleção reúne um conjunto representativo das diferentes comunidades falantes da língua estrangeira?	8. A coleção traz diferentes tipos de estratégias de ensino/aprendizagem, tais como localização de informações explícitas e implícitas no texto, levantamento de hipóteses, produção de inferência, compreensão detalhada e global do texto, dentre outras?	9. A coleção propõe atividades que convidam o aluno a usar a língua de forma crítica?
10. A coleção aborda atividades que promovem o processo de reescrita do próprio texto?	11. Há atividades relativas a diferentes situações de comunicação, que estejam em inter-relação com necessidades de fala compatíveis com as do estudante do Ensino Médio?	12. Os conceitos e informações são apresentados com instruções claras e contextualizados?

13. Há relevância entre as atividades de leitura, escrita e oralidade, integrando propósitos e finalidades da aprendizagem da língua estrangeira?	14. Há uma boa disposição no que se refere ao conteúdo (foco na língua e na habilidade/criatividade do aluno)?	15. A coleção propõe atividades de avaliação e de autoavaliação?
16. O livro do professor explicita de forma clara e objetiva a proposta didático- pedagógica?	17. A coleção integra as 4 habilidades no processo de aprendizagem da língua estrangeira no ensino médio?	18. Há uma coerência entre sua fundamentação e a proposta de ensino presente no livro do aluno, levando em consideração as necessidades dos alunos?
19. Qual abordagem embasa a coleção em questão?	20. É possível observar uma flexibilidade no material didático (possibilitando ao professor diferentes estilos de ensino, assim como suprimir e/ou incluir conteúdos)?	21. A coleção é isenta de estereótipos e preconceitos (socioeconômico, regional, étnico-racial, de gênero, etc.)?

Fonte: Silva, Parreiras e Fernandes (2015, p. 367-369)

Ao propor esses itens, Silva, Parreiras e Fernandes (2015, p.362) afirmam que o foco norteador da análise é que o professor consiga responder “Para que este livro será usado? Por que ele é necessário?”. Uma preocupação dos autores é a possibilidade de o próprio professor conseguir avaliar os livros didáticos:

Este pode ser mais um instrumento em meio a outros já propostos, mas tem como características diferenciadoras o fato de não ser extenso (...) e principalmente por requerer por parte do docente usuário reflexões e visões críticas a respeito de seu contexto e sua prática. Enfatizamos: do próprio docente. (SILVA; PARREIRAS; FERNANDES, 2015, p.372)

As propostas de avaliação de livros didáticos dão base para muitas reflexões sobre a proposta de avaliação de *chatbots* que se apresenta nesta pesquisa e alguns critérios podem ser utilizados. Entretanto, outros são próprios da materialidade impressa do livro didático e de sua lógica de produção, portanto, não são transferíveis para o objeto desta pesquisa, como por exemplo o projeto gráfico editorial e as questões que diferem o material do aluno e o material do professor. Por isso, foi necessário também buscar fontes que apresentam critérios para avaliações de outros tipos de MD.

Belisário (2003) apresenta critérios específicos para produção de materiais didáticos interativos *online*, o que pode se aplicar ao caso de *chatbots*, já que também são interativos e têm funcionamento *online*. O autor organiza os critérios em três grupos: navegabilidade, estrutura e discurso.

No critério Navegabilidade, são avaliadas as formas como os usuários podem intervir no processo de construção do conhecimento, a interação entre mídias e as ferramentas de interação e reprodução de sons e imagens. Já no critério Estrutura, Belisário (2003) aponta a importância do uso de hipertextos para promoção da não linearidade, a divisão do conteúdo em blocos de informações, a sequencialidade do conteúdo e a possibilidade de antecipação e respostas a possíveis dúvidas dos estudantes. Por fim, no critério Discurso, o autor aponta a relevância de se dar atenção à dialogicidade dos textos, às ferramentas de interação textual e o uso de linguagem escrita e oral.

Diante da escassez ou possível ausência de critérios para produção e avaliação de *chatbots* educacionais, utilizou-se como base as categorias cunhadas por autores como Cabero (2001) e Belisário (2003), mas também foram considerados critérios utilizados para produção e avaliação de *chatbots* usados para fins não didáticos, como o atendimento a clientes e marketing. Assim, foram verificados quais critérios podem ser utilizados ou adaptados para o uso de *chatbots* didáticos.

### **3.2 Critérios para Produção e Avaliação de *Chatbots***

Apesar de pesquisas recentes sobre o uso (PEREZ; DARADOMIS; PUIG, 2020; SMUTNY; SCHREIBEROVA, 2020) e produção (BAHJA; HAMMAD; BUTT, 2020) de *chatbots* com propósitos educacionais, não foram encontradas nas bases de dados voltadas para publicação científica (ver Capítulo 1) referências de critérios de avaliação de *chatbots* usados como MD ou de outros tipos de *chatbots* educacionais. Entretanto, foi possível encontrar na literatura critérios de avaliação de *chatbots* usados para fins comerciais. Kuligowska (2015) avaliou *chatbots* usados na Polônia para interação entre empresas e consumidores. A avaliação foi feita com usuários. Com base nas respostas de entrevistas e de um questionário de 5 pontos (1 para muito ruim e 5 para muito bom), a pesquisadora chegou a conclusões sobre os 10 critérios apresentados a seguir:

#### *Critério 1: Aparência visual do chatbot*

Neste item, de acordo com a pesquisa de Kuligowska (2015), quanto mais semelhante a um ser

humano fosse a imagem que representa o atendente virtual, melhor o *chatbot* era avaliado. Se o atendente virtual é uma pessoa que apresenta movimento (com o uso de vídeo), ele é mais bem avaliado que fotos de seres humanos. Animações semelhantes a desenhos animados seriam consideradas adequadas e *chatbots* com aparência não humana ou sem uma imagem que os representem eram avaliados negativamente.

#### *Critério 2: Forma de implementação no website*

Este item diz respeito a como o *chatbot* era apresentado no site ou aplicação em que a interação acontece. *Chatbots* que ocupam parte da tela, janelas inteiras disponibilizadas via *pop up* ou ainda uma aba lateral dedicadas ao *chatbot* foram bem avaliadas por Kuligowska (2015). Pequenas janelas flutuantes (que podem ser movidas de lugar na tela) foram consideradas inadequadas. Percebe-se, neste item, que a dimensão de espaço dedicado ao *chatbot* foi o principal ponto avaliado.

#### *Critério 3: Unidade de produção de fala*

Neste item, *chatbots* com interação por voz, ou texto escrito e voz, foram considerados mais adequados por Kuligowska (2015) que *chatbots* que interagem apenas por meio de texto escrito.

#### *Critério 4: Base de conhecimento*

Kuligowska (2015) avaliou dois principais aspectos de conhecimento dos *chatbots*.

- a) Conhecimentos básicos relacionados ao que se espera de qualquer pessoa, como “Qual é o seu nome?; O que você sabe fazer?; Para quem você trabalha? e Que horas são?”.
- b) Conhecimentos especializados relacionados à empresa representada, seus produtos e serviços, no qual foram avaliadas as respostas a perguntas sobre o catálogo de produtos, preços, descontos, contato para mais informações e a história da empresa.

Os *chatbots* eram bem avaliados por Kuligowska (2015) quando apresentavam informações corretas e completas. Se as respostas eram parcialmente corretas ou incompletas recebiam uma

avaliação intermediária. Quando não conseguiam responder ou apresentavam respostas incorretas, eram avaliados negativamente.

*Critério 5: Apresentação de conhecimentos e funcionalidades adicionais*

Nesse critério, Kuligowska (2015) avaliou a presença de 5 itens de apoio a navegação do usuário e envolvimento com o *chatbot*:

- Carregamento autônomo e dinâmico de novas subpáginas ou abas.
- Conexão interativa com base de dados externa.
- Botão “voltar” ou rolagem da tela do *chat* para visualizar o histórico de mensagens.
- Apresentação do funcionamento do *chatbot* mediante escrita do termo “ajuda” ou de clique nos botões correspondentes a ajuda.
- Presença de um botão “home” para retorno ao menu principal.

*Critério 6: Habilidades conversacionais, habilidades de linguagens e sensibilidade ao contexto*

A avaliação desse critério, para Kuligowska (2015), consiste em verificar se o *chatbot* é capaz de participar de um diálogo coerente, entender expressões complexas e de redirecionar o diálogo ao perceber que entendeu algo errado e em seguida mudar o rumo da conversa.

*Critério 7: Traços de personalidade*

Kuligowska (2015) avaliou a riqueza da personalidade apresentada pelo *chatbot* em termos de informações sobre seu autoconhecimento, como idade, gênero, interesses, bem como seu contorno psicológico, demonstrando ter personalidade coerente com suas “reações emocionais”.

*Critério 8: Opções de personalização*

Kuligowska (2015) verificou a presença ou ausência de cada dos itens abaixo:

- Possibilidade de identificar o gênero (masculino ou feminino) do usuário e conduzir o diálogo considerando o gênero na escolha de palavras femininas ou masculinas.

- Lembrança do nome do usuário pelo *chatbot*.
- Possibilidade de o usuário visualizar o histórico de mensagens durante a conversa.
- Acesso ao histórico da conversa via impressão ou envio por e-mail.
- Reconhecimento pelo *chatbot* da subpágina em que o usuário navega.

*Critério 9: Respostas de emergência em situações inesperadas*

Kuligowska (2015) avaliou a capacidade ou incapacidade de os *chatbots* superarem erros de digitação ou ortografia (por parte do usuário) no entendimento da mensagem, terem uma resposta caso não tenha conhecimento do assunto abordado pelo usuário, lidarem com insultos e humilhação, reconhecerem línguas estrangeiras e traduzir palavras em inglês para o polonês (língua materna dos usuários).

*Critério 10: Possibilidade de atribuição de notas pelo usuário ao chatbot*

Neste item, Kuligowska (2015) considerou positiva a possibilidade de avaliar a satisfação geral com o diálogo e/ou de avaliar cada resposta fornecida pelo *chatbot* ao usuário. Foi considerada negativa a ausência de mecanismo para coletar o *feedback* do usuário.

Nicole Radziwill e Morgan Benton (2017) organizaram critérios de usabilidade de *chatbots*, considerando usabilidade como a efetividade, eficiência e satisfação com a qual usuários alcançam seus objetivos em certos ambientes. Os itens ou atributos de qualidade para avaliação de *chatbots* foram divididos pelos autores em três categorias, apresentadas a seguir:

- Na categoria Eficiência, os autores verificaram que entre os atributos de qualidade de *chatbots* estavam a robustez do sistema, a capacidade de lidar com mensagens inesperadas e a ausência de uso de expressões consideradas inadequadas ao contexto.
- Na categoria Efetividade, foram elencados itens como a capacidade de apresentar discursos sintéticos, de interpretar os comandos do usuário com precisão, de executar as tarefas solicitadas, a facilidade de uso do sistema e a capacidade de usar graus adequados de formalidade. Nesta categoria, também foram apresentados itens relacionados à “humanidade” do *chatbot*, como a revelação de sua identidade de *chatbot*, a inclusão de

erros para aumentar o realismo da conversa, a naturalidade da interação e a capacidade de manter uma discussão temática.

- A terceira e última categoria, Satisfação, apresenta como atributos de qualidade itens como o *chatbot* enviar mensagens para cumprimentar o usuário, conseguir demonstrar traços de sua personalidade, dar deixas conversacionais para manter o diálogo, transmitir emoção pelo tom do discurso, expressividade e cordialidade, a proteção da privacidade dos usuários e a capacidade de detectar a intenção do usuário.

Radziwill e Benton (2017) sugerem que cada *chatbot* deva ser avaliado com base nos itens mais adequados às suas especificidades, descartando a necessidade de usar-se todos os itens. Os autores demonstram a mesma preocupação de Silva, Parreiras e Fernandes (2015) em relação à extensão dos instrumentos de avaliação. Nesta pesquisa, por adesão à ideia de que a extensão dos instrumentos de avaliação não deve ser demasiada extensa, para que seu uso seja viável no tempo que educadores podem dispor para a atividade no dia a dia da profissão, buscou-se criar uma matriz concisa para avaliação de *chatbots* didáticos, conforme apresenta-se a seguir.

### **3.3 Proposta de matriz para avaliação de *chatbots* didáticos**

Diante da escassez de categorias para guiar a avaliação de *chatbots* didáticos, foram usadas como base para a criação de uma matriz de avaliação os critérios de produção de materiais didáticos, com destaque para os interativos *online*, juntamente com os critérios para produção e avaliação de *chatbots* comerciais apresentados anteriormente neste capítulo.

Opta-se por não utilizar todas as categorias estabelecidas pelos autores anteriormente citados, porque, assim como sugerem Silva, Parreiras e Fernandes (2015), a dimensão das matrizes de avaliação de MD precisam considerar o tempo disponível para que profissionais da educação possam de fato utilizá-las. Além disso, diante da decisão de se apresentar uma matriz enxuta, privilegiou-se as categorias que têm relação direta com as especificidades dos *chatbots* e com os objetivos desta pesquisa. Nas cinco categorias propostas, foram considerados alguns descritores a partir de itens sugeridos pelos autores citados neste capítulo. Outros itens foram descartados e novas categorias e descritores foram criados para se contemplar as especificidades do *chatbot* didático, o que não é feito por nenhum dos autores citados anteriormente e configura-se como parte do ineditismo e contribuição desta tese.

No Quadro 3, são apresentadas 5 categorias que possuem um total de 30 descritores. Os descritores em negrito foram criados durante o processo de realização dessa pesquisa, considerando as especificidades do *chatbot* didático. Os outros foram cunhados por autores citados neste capítulo.

**QUADRO 3- Critérios para produção e avaliação de *chatbots* usados como material didático**

<b>Categoria 1: Conteúdos</b>		
Descritor	Questões norteadoras	Avaliação
Conteúdo de apresentação	O <i>chatbot</i> é capaz de se apresentar, dizendo quem é e para que serve?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
Conhecimentos especializados	O quão completo é o conhecimento do <i>chatbot</i> sobre os conteúdos trabalhados no curso ou disciplina que ele atende, incluindo conteúdo (conceitos, teorias etc), atividades e agenda do curso. Os conteúdos são válidos e embasados cientificamente?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
Atualização dos conteúdos	Os conteúdos são atualizados em relação à versão do curso ofertado, caso o <i>chatbot</i> esteja sendo ofertado em períodos letivos consecutivos? Por exemplo, o calendário do curso foi atualizado para a nova turma?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
Antecipação às dúvidas mais comuns dos estudantes	As informações ofertadas pelo <i>chatbot</i> via opções de menu são capazes de antecipar e responder às possíveis dúvidas mais comuns dos estudantes?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
<b>Conexão com outros conteúdos do curso</b>	Os conteúdos apresentados no <i>chatbot</i> fazem menção e estão diretamente relacionados com outros conteúdos do curso?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
<b>Categoria 2: Discurso e linguagens</b>		
Modalidades oral e escrita de uso da língua	O <i>chatbot</i> aceita <i>inputs</i> por texto oral (áudio) e escrito e dá respostas também com texto oral e escrito?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
Clareza	Os textos são apresentados de maneira clara e coerente para entendimento dos alunos?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
Lembrança do nome do usuário	O <i>chatbot</i> consegue identificar o nome do usuário e se referir a ele pelo nome durante a conversa?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
<b>Suficiência de informações</b>	A quantidade de conteúdo é suficiente para dar apoio aos estudos autônomos dos alunos? A extensão/tamanho dos textos é adequada ao que os leitores precisam, querem ou estão dispostos a ler?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado

<b>Adequação do vocabulário ao curso</b>	O vocabulário usado pelo <i>chatbot</i> está coerente e em sintonia com o vocabulário usado nos outros materiais didáticos do curso? Os diálogos são realizados em linguagem natural com registros próximos aos usados pelos usuários/alunos no curso?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
Formalidade	O grau de formalidade está adequado ao contexto, ao público-alvo e à persona do <i>chatbot</i> ?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
<b>Categoria 3: Navegação e Usabilidade</b>		
<i>Layout</i>	Formato dos textos, imagens, cores, disposição de elementos, entre outras questões estéticas se apresentam de forma agradável e funcional?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
<b>Adaptação/individualização do processo de interação e aprendizagem</b>	Cada experiência de interação é única, diferente das dos colegas de curso? Há a possibilidade de os estudantes intervirem no processo de construção do conhecimento ou de estabelecerem um processo único de aprendizagem?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
Reprodução de sons e imagens	Sons e imagens (por exemplo, mensagens de voz, músicas, sons diferenciados, desenhos, fotos e vídeos) são reproduzidos pelo <i>chatbot</i> ?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
<b>Responsividade</b>	A disposição dos elementos visuais se adequa ao tamanho da tela em que usuário acessa o <i>chatbot</i> ?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
<b>Acessibilidade</b>	Há recursos tecnológicos para que pessoas com deficiência consigam utilizar o <i>chatbot</i> ?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
Uso de hipertextos	São usados hipertextos nas mensagens enviadas pelo <i>chatbot</i> aos alunos para promoção da não linearidade?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
Facilidade de uso do sistema	Os alunos têm facilidade em usar o <i>chatbot</i> , identificando rapidamente onde clicar, escrever e outras maneiras de usar o sistema para atingir os objetivos de interação?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado
<b>Opções de menu e de digitação simultâneas</b>	Ao interagir com o <i>chatbot</i> , os usuários podem escolher entre digitar ou clicar em botões com opções de menu, seja na tela inicial ou em telas subsequentes?	( ) Implementado ( ) Parcialmente Implementado ( ) Ausente
<b>Categoria 4: Aparência e persona</b>		
Persona do <i>chatbot</i>	Se o <i>chatbot</i> se apresenta como um personagem, ele tem nome, profissão, gênero, humor ou outras características demográficas ou comportamentais expressas por meio de sua imagem e discurso? Os alunos gostam dessa persona e da aparência do <i>chatbot</i> ?	( ) Adequado ( ) Parcialmente Adequado ( ) Inadequado

Papel desempenhado pelo <i>chatbot</i>	O <i>chatbot</i> deixa claro qual papel desempenha na situação de interação com o usuário? Por exemplo, é uma colega que vai conversar com o aluno, um atendente administrativo que vai ajudar a resolver questões de secretaria, uma professora que vai dirimir dúvidas pedagógicas etc? Os usuários aprovam esse papel?	<input type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Revelação de sua identidade de <i>chatbot</i>	O <i>chatbot</i> , em algum momento inicial de interação, revela ao usuário que não é um ser humano respondendo às perguntas, mas uma máquina?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Cumprimento	Há uma mensagem para cumprimentar o usuário? Essa mensagem transmite algum aspecto da persona do <i>chatbot</i> ?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Categoria 5: Funcionalidades adicionais</b>		
Resposta a questões desconhecidas	Existe uma ou mais respostas-padrão para os casos em que usuário faz uma pergunta para a qual o <i>chatbot</i> não tem resposta?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Conexão com base de dados externas	O <i>chatbot</i> utiliza bases de dados externas, como o Google, Wikipedia ou outras fontes disponíveis <i>online</i> , e não apenas sua própria base de dados para coletar informações e gerar textos de respostas?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Visualização do histórico de mensagens	O usuário pode usar rolagem de tela ou outro recurso para reler o histórico de mensagens trocadas ao longo do diálogo?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Opção de ajuda	Há algum botão ou opção de ajuda para que o usuário leia informações sobre como utilizar o sistema?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Opção para voltar ao menu inicial	Há algum comando ou opção de onde clicar para voltar ao menu inicial do sistema?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Coleta de <i>feedbacks</i>	Existe algum mecanismo de coleta de <i>feedbacks</i> para que o usuário avalie a qualidade das respostas/ interação?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Direcionamento de dúvidas	Dúvidas não respondidas ou demandas não solucionadas pelo <i>chatbot</i> são direcionadas ao docente, coordenador do curso ou alguém que possa dirimir as dúvidas do usuário?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Fonte: elaborado pelo autor.

No quadro, a coluna Avaliação é apresentada com opções pré-definidas para marcação do resultado da avaliação. Ela será preenchida nesta tese na seção de Resultados, com base na

avaliação do *chatbot* Katri por usuários-alunos, por produtores e pelo pesquisador. Alguns dos itens necessitam objetivamente da verificação se o recurso ou conteúdo está implementado/presente ou ausente no *chatbot*. Outros itens necessitam de uma avaliação qualitativa, na qual são consideradas a avaliação de alunos do curso e produtores do *chatbot*. O avaliador ou produtor de cada *chatbot*, diante do contexto de uso do MD e suas especificidades, é quem pode decidir por uma ou outra abordagem, se haverá ou não análise qualitativa em cada descritor.

Sugere-se o preenchimento da coluna de Avaliação com os indicadores Adequado ou Implementado, no caso de avaliação satisfatória, Parcialmente Adequado ou Parcialmente Implementado, no caso de avaliação considerada intermediária; e Inadequado ou Ausente, no caso de avaliação considerada insatisfatória. O avaliador pode, entretanto, criar seus próprios indicadores para expressar o resultado de sua avaliação. Sugere-se também que o avaliador escreva comentários sobre a sua avaliação de cada item, se julgar que a redação de comentários pode auxiliar no entendimento de como se chegou ao indicador marcado na matriz.

Conforme a sugestão de Radziwill e Benton (2017), nem todos os itens da matriz são obrigatórios. Cada avaliador pode analisar quais são os itens mais relevantes de acordo com as especificidades do *chatbot* a ser avaliado ou de acordo com os objetivos da avaliação.

No próximo capítulo será apresentado o Sistema de Avaliatividade, proposto por Martin e White (2005). É com base neste sistema que serão analisadas as avaliações de Katri por seus produtores e por alunos do curso TLHE. Com base nessas análises, a coluna de Avaliação da matriz proposta será preenchida pelo pesquisador no capítulo de Resultados.

#### 4. LINGUAGEM E AVALIAÇÃO

Para validar a matriz de avaliação de *chatbots* didáticos, propõe-se utilizar como arcabouço teórico o Sistema de Avaliatividade proposto por Martin e White (2005) como uma extensão da Linguística Sistêmico-Funcional (HALLIDAY; MATTHIESSEN, 2014).

A Linguística Sistêmico-Funcional (LSF) é uma vertente da corrente funcionalista da linguística, que concebe a “linguagem como instrumento de comunicação e interação social, cuja forma se adapta às funções que exerce” (MUSSALIM, 2008, p.84-85), rejeitando a separação estruturalista entre sistema e uso da linguagem.

O linguista britânico Michael Halliday lançou a primeira edição de seu livro *Introdução à Gramática Funcional* em 1985, no qual propõe uma visão de linguagem ligada direta e inseparável à sua funcionalidade em contextos de uso. Essa visão é expressa por Halliday em suas definições de texto e linguagem. Para ele, “podemos caracterizar o texto como a linguagem funcionando no contexto” e “a linguagem é, em primeira instância, um recurso para produzir sentido; então o texto é um processo de fazer sentido no contexto” <sup>21</sup> (HALLIDAY; MATTHIESSEN, 2014, p. 3).

Halliday e Matthiessen (2014) explicam que a funcionalidade é um aspecto intrínseco à linguagem. Assim, a organização da arquitetura da linguagem está relacionada às suas funções em contexto. Para eles, a linguagem apenas se apresenta da maneira como é devido à forma como as suas funções evoluíram junto com a espécie humana (HALLIDAY; MATTHIESSEN, 2014, p. 31).

Para Repolês (2019), a LSF é

sistêmica por construir significados interligando os diferentes estratos (fonologia, grafologia, léxico-gramática e semântica) envolvidos pelo contexto e recursos de significação; e funcional porque fazemos escolhas dentro do sistema linguístico para alcançarmos os nossos propósitos: construção e representação do mundo, ação sobre outras pessoas e construção de mensagens (REPOLÊS, 2019, p. 65).

---

<sup>21</sup> Original: “We can characterize text as language functioning in context. Language is, in the first instance, a resource for making meaning; so text is a process of making meaning in context”

Assim, as regras do sistema só são consideradas dentro de seus contextos de uso real para produção de sentido. A LSF “concebe a língua como uma rede de sistemas interligados que o falante faz uso (base funcional) para produzir significados (base semântica) em situações de comunicação” (SANTOS, 2014, p. 166). Ressalta-se também que a língua como sistema linguístico dentro da LSF é interativa e dinâmica, uma vez que os significados são construídos e reconstruídos durante o uso efetivo da língua em situações de comunicação (SANTOS, 2014).

Os significados não são construídos aleatoriamente, afinal, a língua permite escolhas por seus usuários. A LSF se preocupa em entender essas escolhas e não se propõe a estabelecer um conjunto estanque de regras linguísticas, mas a entender as opções linguísticas disponíveis e as escolhas de cada falante realizadas em cada contexto, uma vez que “cada escolha adquire um significado em detrimento a outras escolhas que poderiam ter sido feitas” (SANTOS, 2014, p. 166).

A importância da noção de contexto é fundamental para a LSF, uma vez que é o contexto que determina não só a gama de opções possíveis, mas também a eficiência da escolha feita para produzir o sentido almejado. Essas escolhas são feitas em todos os contextos, inclusive no uso da língua e da linguagem para produção de textos de avaliação, tema que interessa a esta pesquisa.

Osiek (2014), explica que a noção de contexto dentro da LSF é tão importante que, para ser corretamente compreendida, é preciso entendê-la em dois níveis. O primeiro nível, considerado mais abstrato, é o contexto de cultura (ou social), que tem relação com as convenções socialmente adotadas para a comunicação em sociedade, levando em conta a evolução da língua e das relações sociais dos indivíduos. O segundo nível, considerado mais perceptível, é o contexto de situação, que pode ainda ser analisado a partir de três aspectos: “campo, que identifica as pessoas envolvidas e sobre o que elas falam; relação, que identifica o relacionamento entre a pessoa que fala e seu interlocutor; e modo, que consiste na organização e desenvolvimento de uma conversa” (OSIEK, 2014, p. 35).

Segundo Halliday e Matthiessen (2014), há três metafunções da linguagem e todas elas se relacionam com os contextos: a ideacional, que se constitui a partir das enunciações das experiências; a interpessoal, usada nas negociações das relações, constituída por meio da interação, tomando a comunicação como um espaço de troca; e a textual, usada para organizar

os textos simbolicamente, ligada ao sistema de realização léxico-gramatical.

Dessa forma, pode-se dizer que todo uso da língua está inserido em um contexto (HALLIDAY; MATTHIESSEN, 2014) e todo texto está inserido em contextos de cultura e de situação (REPOLÊS, 2019). Os contextos de cultura são mais estáveis, enquanto os de situação são constituídos no momento em que o texto é produzido, de acordo com as características imediatas da interação.

As interações dos alunos com Katri no curso TLHE, por exemplo, estão em contexto de cultura acadêmica, que tem características relativamente estáveis compartilhadas pela comunidade acadêmica em diferentes situações de interação, e há contextos específicos da situação, no que diz respeito ao campo (alunos e “robô-monitor virtual” conversando sobre questões relacionadas especificamente ao curso TLHE), nas relações (há posicionamentos que moldam a interação e estão relacionados aos papéis de alunos e robô-monitor do curso) e no modo (o texto é construído por escrito na relação entre os interlocutores usando uma plataforma digital como recurso do curso, podendo-se escrever palavras ou enviar-receber links e arquivos anexos). Além dessas exemplificações, o contexto de situação é virtualmente imprevisível, pois cada interação tem seus próprios contextos que emergem no momento em que a interação é realizada.

Para Eggins (2004), a principal contribuição de Halliday para a área da linguística foi o desenvolvimento de sua Gramática Funcional, que possibilita fazer análises sensatas e úteis sobre qualquer texto, falado ou escrito. Essa possibilidade da LSF foi usada por Martin e White (2005) para a criação do Sistema de Avaliatividade, que orienta a análise de textos que expressam avaliações feitas pelos usuários da língua (SILVA, 2015). As avaliações dos alunos do curso TLHE e dos produtores do *chatbot* Katri são analisadas nesta pesquisa de acordo com o sistema proposto por Martin e White (2005), que será apresentado a seguir.

#### **4.1 Sistema de Avaliatividade**

Com base na LSF, Martin e White (2005) propuseram o Sistema de Avaliatividade, que sistematiza escolhas que os usuários da língua podem fazer para expressar suas avaliações. A Avaliatividade (do inglês *Appraisal*) está ligada diretamente à metafunção interpessoal da

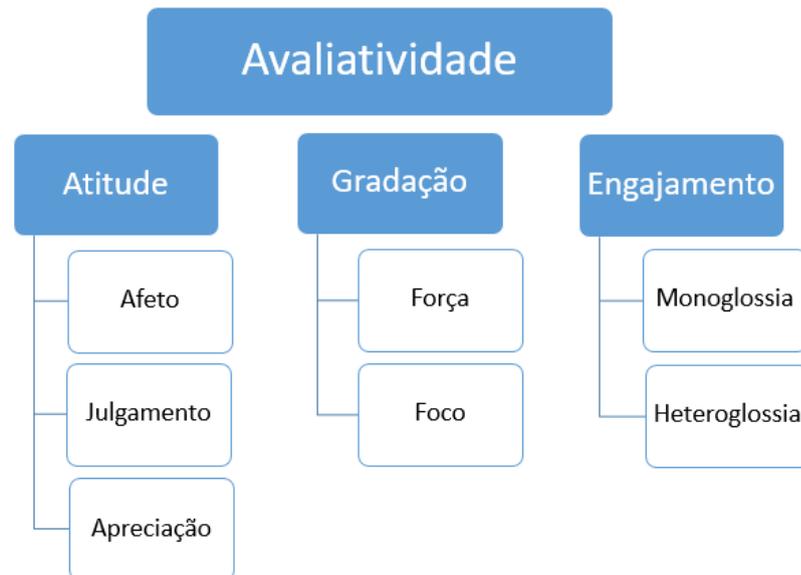
linguagem. É por meio dela que se demonstra aos interlocutores o que se pensa e o que se sente em relação às coisas e às pessoas.

### O Sistema de Avaliatividade

está preocupado com a forma como escritores/oradores aprovam e desaprovam, entusiasmam e abominam, aplaudem e criticam, e com como eles posicionam seus leitores/ouvintes para fazer o mesmo. Está preocupado com a construção por textos de comunidades de sentimentos e valores compartilhados, e com os mecanismos linguísticos para o compartilhamento de emoções, gostos e avaliações normativas<sup>22</sup> (MARTIN; WHITE, 2005, p. 1).

A Avaliatividade é dividida por Martin e White (2005) em três domínios: Atitude, Gradação e Engajamento. A Atitude contempla os sentimentos, incluindo reações emocionais, julgamentos de comportamento e avaliação das coisas e pessoas. A Gradação está relacionada a maior ou menor intensidade demonstradas na avaliação por meio de recursos linguísticos. Já o Engajamento é determinado por quais/quantas vozes estão envolvidas no discurso avaliativo.

**FIGURA 9 - Sistema de Avaliatividade**



Fonte: traduzido e adaptado de Martin e White (2005, p. 38)

O subsistema de Engajamento fornece os meios para a voz autoral se posicionar em relação ao

<sup>22</sup> Original: “It is concerned with how writers/speakers approve and disapprove, enthuse and abhor, applaud and criticise, and with how they position their readers/listeners to do likewise. It is concerned with the construction by texts of communities of shared feelings and values, and with the linguistic mechanisms for the sharing of emotions, tastes and normative assessments.”

objeto avaliado e também a engajar-se com as outras vozes no discurso avaliativo (MARTIN; WHITE, 2005). “Com isso, marca (o avaliador) sua posição em atitude de filiação ao discurso, distanciamento da voz autoral, ou em busca de endossamento de seu discurso” (REPOLÊS, 2019, p. 93). O envolvimento de outras vozes, por exemplo, pode ser feito por meio da citação direta ou indireta de textos de outras pessoas.

O Engajamento pode ser monoglóssico, quando há apenas uma voz no discurso, ou heteroglóssico, quando há mais de uma voz (MARTIN; WHITE, 2005), o que reforça a concepção bakhtiniana da natureza dialógica do discurso. O subsistema de Engajamento, entretanto, não é relevante para a proposta de avaliação de *chatbots* apresentada nesta tese, pois não haverá distinção da Atitude e da Gradação do discurso avaliativo dos usuários de *chatbots*, independentemente do número de vozes envolvidas na expressão da avaliação. Os subsistemas de Atitude e Gradação são detalhados a seguir, pois serão relevantes para análise dos dados da pesquisa.

#### 4.1.1 Atitude: Afeto e Apreciação

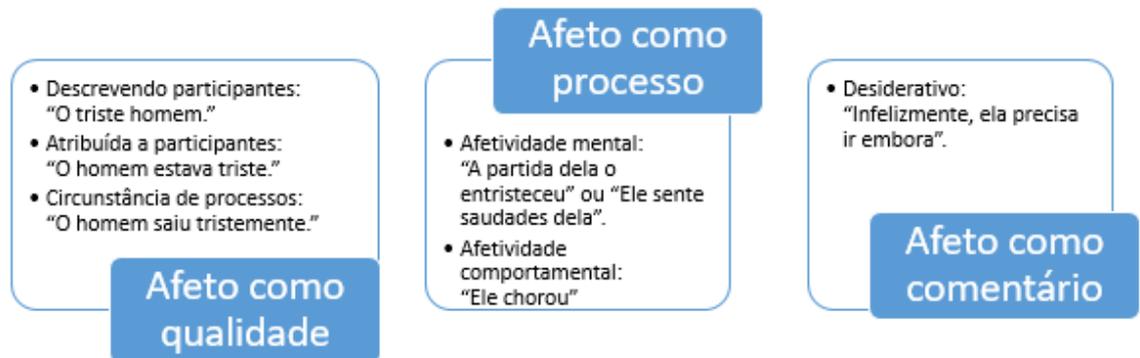
No Sistema de Avaliatividade, o subsistema de Atitude é expresso no discurso quando há elementos que indicam Afeto, Julgamento ou Apreciação. O Afeto está relacionado diretamente à expressão de sentimentos em relação a coisas ou pessoas, o Julgamento é direcionado à avaliação de caráter e conduta social das pessoas e a Apreciação diz respeito ao valor que se dá as coisas (objetos, fenômenos naturais e situações, por exemplo). Como Katri não é uma pessoa, nesta tese não será avaliada a dimensão Julgamento. Apenas Afeto e Apreciação serão dimensões utilizadas na avaliação de Katri, portanto serão detalhadas a seguir.

Em um discurso avaliativo, Afeto compreende a expressão das emoções e sentimentos, que podem ser positivos ou negativos (polaridade), diretos ou implícitos. O Afeto se apresenta de forma direta quando há no discurso o uso de palavras que denotam a emoção, como em “ela ficou **feliz**” ou “ela ficou **triste**”. É implícito quando a emoção é indicada por meio de atitudes expressas no discurso (MARTIN; WHITE, 2005), como em “à medida que o tempo foi passando, **um sorriso foi se abrindo** no rosto dela”. Neste exemplo, dependendo do contexto do enunciado, pode haver a expressão de felicidade (Afeto) de maneira implícita no discurso.

Martin e White (2005) baseados em Halliday (1994), apresentam o Afeto como sistema

semântico do discurso. Os autores apontam que podemos esperar que as realizações de Afeto se diversifiquem através de uma variedade de estruturas gramaticais para descrever qualidades, relatar processos e realizar comentários.

**FIGURA 10 - Exemplos de expressões textuais que demonstram Afeto**

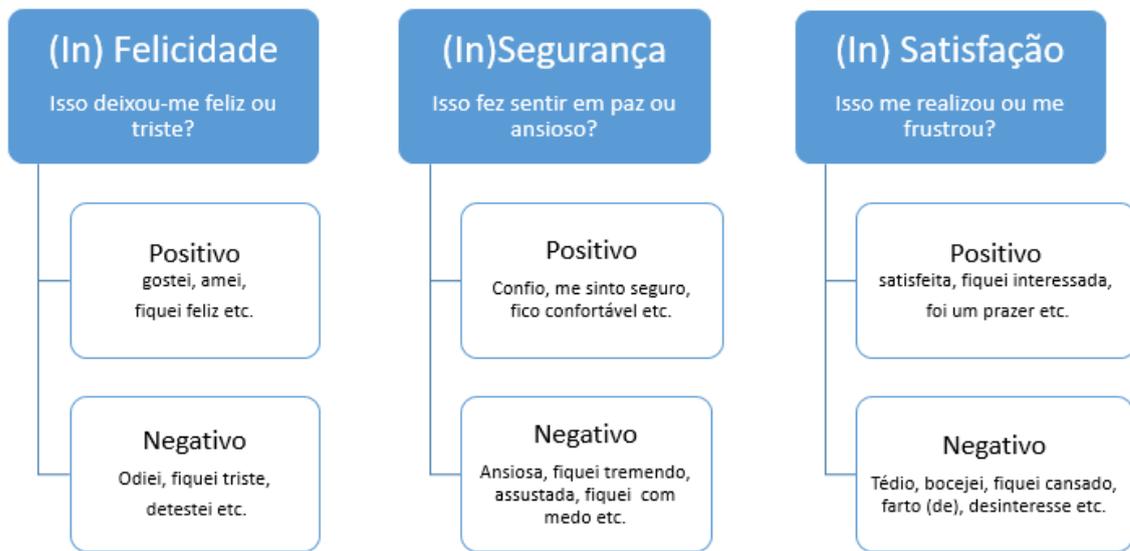


Fonte: Baseado em Martin e White (2005, p.46)

As construções textuais que expressam emoções, relacionadas ao Afeto, revelam o envolvimento sentimental/emocional do produtor do texto com as coisas, pessoas e situações (MARTIN; WHITE, 2005; REPOLÊS; 2019). São inúmeras as possibilidades de sentimentos e emoções expressas pelo ser humano, mas Martin e White (2005) as classificam em três principais grupos: (in)felicidade, (in)segurança e (in)satisfação.

Segundo Martin e White (2005), a (in)felicidade é expressa por reações emocionais de alegria ou tristeza, enquanto a (in)segurança compreende sentimentos de paz e ansiedade com o ambiente, e a (in)satisfação está ligada aos sentimentos de realização e frustração. São todos campos da subjetividade humana e podem ser expressos por um amplo repertório de escolhas lexicais.

**FIGURA 11 - Exemplos de expressões textuais que demonstram afetos positivos e negativos**



Fonte: Baseado em Martin e White (2005, p. 49-51)

Há expressões diretas e implícitas dos três grupos. Eles podem ser expressos pelo uso de adjetivos e verbos, mas também por outras classes gramaticais, pois como preconizam Halliday e Matthiessen (2014), é no contexto que se pode analisar o discurso. Assim um substantivo poderia ser um elemento indicativo de afeto, como na frase “Estou uma pilha”, que poderia ser usada para uma pessoa expressar que está muito ansiosa, o que seria classificado como Insegurança no subsistema de afeto.

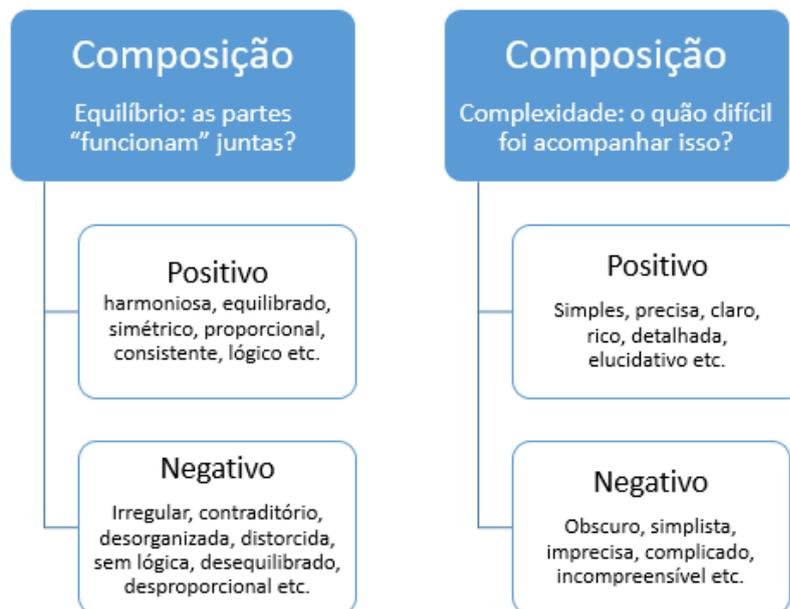
Como exemplos de expressão de afeto, uma das participantes do curso TLHE, ao interagir com Katri pela primeira vez, disse: “Busquei Katri para compreender melhor como elaborar meu projeto pedagógico e fiquei bem satisfeita com as respostas. Minhas dúvidas foram completamente resolvidas” (Participante 18). Este discurso utiliza o afeto para descrever qualidade. A palavra satisfeita aparece como expressão direta e positiva de satisfação.

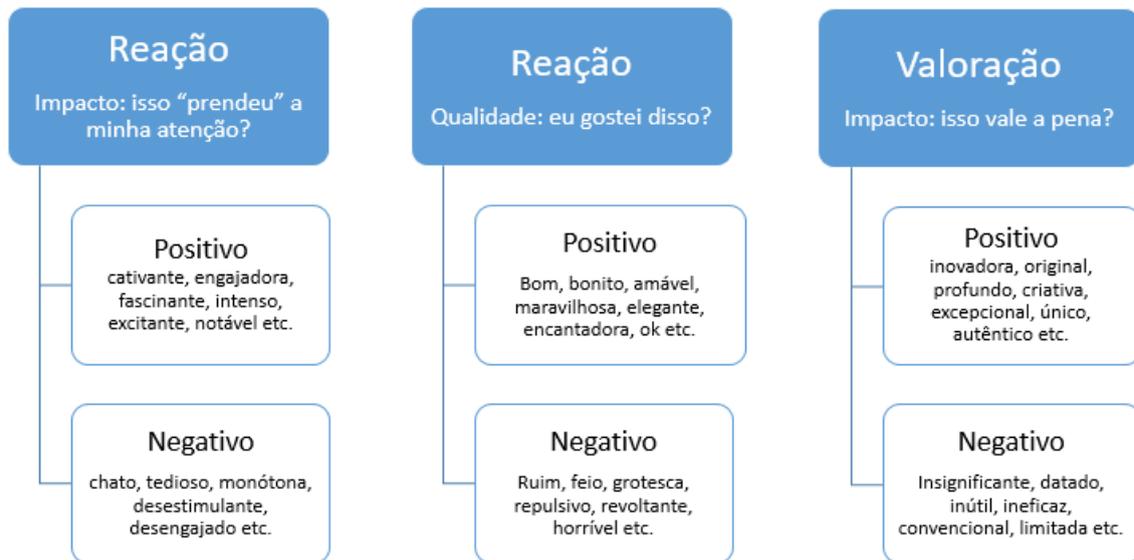
Outra participante afirmou: “Eu sou um pouco incomodada às vezes com essa questão de definição de persona, porque, por exemplo, eu sou mulher e eu não me identifico com a Katri” (Participante 3). Este discurso também utiliza o afeto para descrever qualidade. A palavra incomodada, entendida no contexto da sentença, demonstra indiretamente uma avaliação negativa relacionada a (in)segurança. No contexto, infere-se que a persona de Katri causa desconforto à participante.

Além das expressões de afeto apresentadas acima, os excertos com as opiniões das participantes do curso TLHE também apresentam elementos de Apreciação, que é a Atitude do produtor do texto em relação às “coisas”, termo usado para se referir, por exemplo, a objetos, fenômenos naturais, lugares e instituições (MARTIN; WHITE, 2005). Katri, por ser um sistema computacional, é considerada uma coisa no Sistema de Avaliatividade.

As Apreciações são divididas em três grupos: reações às coisas, a composição delas e o valor que se dá a elas. Podem também ser positivas ou negativas, conforme a figura a seguir.

**FIGURA 12 - Tipos de Apreciação**





Fonte: Baseado em Martin e White (2005, p.56)

Para exemplificar a expressão da Apreciação, retoma-se as opiniões das Participantes do curso TLHE. O excerto “minhas dúvidas foram completamente resolvidas (Participante 18)” traz implicitamente a valoração positiva, visto que a participante expressa que houve bom resultado na interação com Katri, ou seja, indiretamente ela demonstra que valeu a pena usar o *chatbot*. Já no excerto “Eu sou mulher e eu não me identifico com a Katri” (Participante 3)” há uma implícita apreciação negativa da composição do *chatbot*, pois a participante expressa que há algo em desequilíbrio na maneira como Katri se apresenta.

As Atitudes são expressas com diferentes intensidades, o que está ligado ao subsistema de Gradação apresentado abaixo.

#### 4.1.2 Gradação: foco e força

A Gradação “é uma propriedade geral de valor de afeto, julgamento e apreciação que constrói maiores ou menores graus de positividade ou negatividade”<sup>23</sup> no discurso avaliativo (MARTIN; WHITE, 2005, p. 135). O subsistema de Gradação é composto por duas dimensões: Força e Foco.

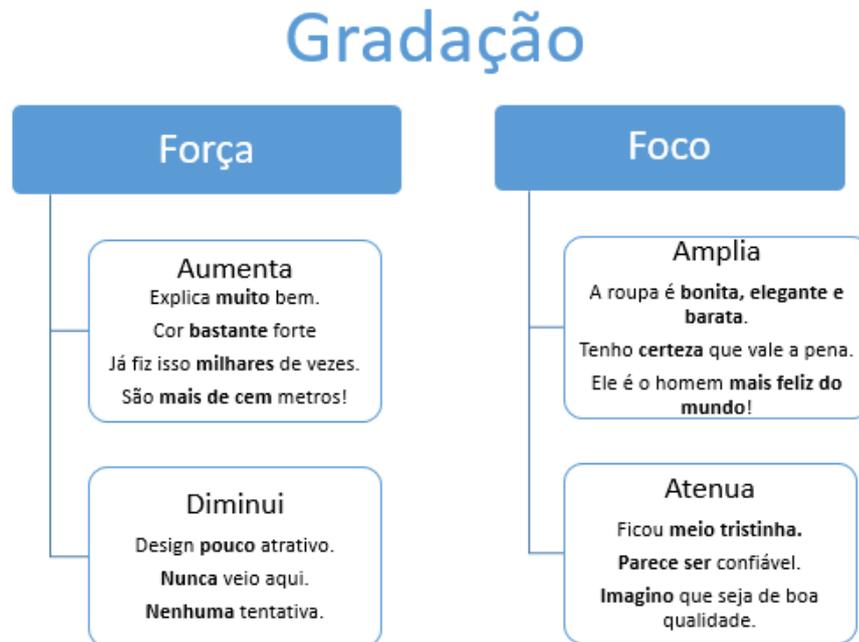
<sup>23</sup> Original: “It is a general property of values of affect, judgement and appreciation that they construe greater or lesser degrees of positivity or negativity”.

A Força se faz presente nos discursos avaliativos por meio de recursos linguísticos que expressam intensificação e quantificação. São exemplos de intensificadores advérbios como “muito” e “pouco”, superlativos, repetições de palavras e o uso de palavrões.

Já a quantificação é usada para amplificar ou diminuir a força das Atitudes por meio da expressão de quantidade, volume ou extensão. Escolhas lexicais que demonstram quantidades determinadas ou indeterminadas (“a professora repetiu 10 vezes essa explicação” ou “a professora repetiu várias vezes essa explicação”), tamanho, altura, peso, tempo e espaço são exemplos de quantificação.

Ao contrário da Força, o Foco é aplicável a experiências não escaláveis ou mensuráveis. É representado por meio de padrões lexicais, como uma sequência de adjetivos em uma frase e palavras que expressam certeza ou incerteza, suavização ou exagero, agravamento etc.

FIGURA 13 - Subsistema de Gradação



Fonte: Baseado em Martin e White (2005) e Osiek (2014)

Ao avaliar Katri, a Participante 3 disse que estava “**um pouco** incomodada **às vezes** com essa questão de definição de persona”. Há no discurso a expressão de diminuição da força por meio das palavras destacadas. Já a Participante 18 afirmou: “fiquei **bem** satisfeita com as respostas”. Neste caso, com o uso do advérbio “bem” há o aumento da força em relação à satisfação (Afeto).

Diante do sistema apresentado neste capítulo, propõe-se analisar os discursos de avaliação de *chatbots* didáticos a partir das Atitudes envolvidas nos discursos avaliativos e a possível Gradação das Atitudes, sendo elas positivas ou negativas. Será investigado, em cada discurso de avaliação do *chatbot* Katri, a presença de elementos textuais que expressem Afeto ou Apreciação, assim como marcadores de Força ou Foco. O instrumento de análise, considerando o Sistema de Avaliatividade, é descrito na sessão de Metodologia, apresentada a seguir.

## 5. METODOLOGIA

Durante o período de realização desta pesquisa, foi criado o *chatbot* Katri, que atuou como uma “monitora” ou “assistente” virtual para esclarecimento de dúvidas administrativas e de conteúdo para alunos do curso TLHE. Por meio da experiência de produção e da observação do uso desse *chatbot*, e também com base em critérios para produção e avaliação de materiais didáticos e de *chatbots*, foi proposta no capítulo 3 desta tese uma matriz de critérios para avaliação de *chatbots* didáticos. Para que a matriz pudesse ser utilizada para avaliar Katri, foram adotados os procedimentos metodológicos descritos a seguir. Neste capítulo, primeiramente, apresenta-se o enquadramento metodológico. Em seguida, são apresentados os participantes da pesquisa, os instrumentos e procedimentos de coleta e análise de dados.

### 5.1 Enquadramento metodológico

Quanto ao enquadramento metodológico, esta é uma pesquisa qualitativa aplicada. Sendo assim, os dados não passaram por tratamentos estatísticos, apesar de a pesquisa apresentar quantificações. A pesquisa visa à geração de “conhecimentos para a aplicação prática, dirigidos à solução de problemas concretos e específicos do mundo real” (CRUZ, 2001, p.20). A proposta foi coletar informações sobre uma realidade educacional e verificar seus pontos fortes e fracos para conseguir propor soluções para os problemas existentes nesse contexto. Há vários tipos de pesquisas qualitativas. Entende-se que esta é um estudo de caso com observação participante, na perspectiva da autonetnografia (KOZINETZ, 2014).

O estudo de caso com observação participante é uma nomenclatura apresentada por Yin (2010) como uma pesquisa na qual o pesquisador é um sujeito ativo no processo investigado e intervém para solucionar problemas que verifica nesse processo. Essa modalidade de estudo de caso, segundo o autor, se caracteriza pela postura ativa do pesquisador que, longe de ser um observador passivo, assume diferentes papéis no caso estudado. Assim, ele atua de maneira efetiva nos eventos observados.

Este estudo de caso foi realizado na perspectiva da autonetnografia. A netnografia é, segundo Corrêa e Rozados (2017, p. 2), “uma adaptação da pesquisa etnográfica que leva em conta as

características dos ambientes digitais e da comunicação mediada por computador”. Podem ser objetos da netnografia os fóruns de discussão online, redes sociais, mundos virtuais, entre outras ferramentas de comunicação mediada por computador. Os autores explicam que a netnografia

é um método de pesquisa, baseado na observação participante e no trabalho de campo *online*, que utiliza as diferentes formas de comunicação mediada por computador como fonte de dados para a compreensão e a representação etnográfica dos fenômenos culturais e comunais. (CORRÊA; ROZADOS, 2017, p. 3)

Enquadra-se esta pesquisa na autonetnografia porque observou-se o comportamento de um grupamento social que existe e interage utilizando a internet e mídias digitais: os alunos que interagem com o *chatbot* do curso TLHE. Para Kozinetz (2014), os grupamentos sociais estudados pela netnografia estão dispostos em rede e podem funcionar interagindo apenas por meios *online* ou podem coexistir fora da rede, em espaços *offline*. É o caso do grupamento observado neste estudo, visto que as interações dos alunos do curso TLHE com o *chatbot* ocorrem apenas *online* - e somente essas interações são objeto dessa pesquisa. No entanto, sabe-se que os alunos interagem entre si em espaços *online* e *offline*, já que o curso é ofertado em modelo híbrido, com módulos em que as interações acontecem a distância por meio de ambiente virtual de aprendizagem e também com módulos presenciais, em que há interações *offline*.

Ao assumir a perspectiva netnográfica neste estudo de caso, apenas as interações mediadas por computador serão consideradas. Dentre as possíveis interações *online*, as dos alunos com o *chatbot* do curso serão consideradas como *corpus* desta pesquisa. Adota-se aqui o termo autonetnografia porque este é um estudo netnográfico sobre um grupamento do qual o pesquisador faz parte, já que é o coordenador do curso. É essa pertença ao grupamento que distingue a netnografia (caso fosse uma pesquisa feita por um observador externo) da autonetnografia (quando o observador faz parte do grupamento observado), seguindo a mesma lógica que separa a etnografia da autoetnografia (AMARAL, 2009).

Retomando o conceito de observação participante que configura este estudo de caso, Yin (2010) observa que, em ambientes organizacionais, a atuação do pesquisador- observador inclui os papéis de membro da equipe que trabalha no caso estudado e de tomador de decisões relacionadas a ele. Esses múltiplos papéis que o pesquisador exerce seriam, então, oportunidades para que a realidade possa ser documentada por alguém “interno” à organização e ao caso/objeto de estudo, o que proporciona coleta e análises de dados distintas das que seriam

possíveis com um observador externo.

Existem estudos de caso que tem o objetivo de dar *feedback* ou informação para promover melhorias na ação/caso pesquisado (ANDRÉ, 2005). Esse é um dos objetivos desta pesquisa, já que o pesquisador realmente trabalha com o *chatbot* estudado e tem interesse no desenvolvimento e processos de melhoria do sistema para que sejam alcançados resultados mais positivos no processo de aprendizagem dos alunos por meio do *chatbot* e no curso em que ele é usado como MD.

Ao integrar a equipe que criou o *chatbot* Katri para o curso TLHE em seu dia a dia de trabalho como coordenador do curso, o investigador não pretende se apresentar como um pesquisador afastado do objeto de estudo, por achar necessário assumir essa condição. De acordo com Raupp e Beuren (2003, p.90-91), essa situação em que o investigador/observador é um ator ativo no processo estudado tem o benefício de apresentar “a possibilidade de aplicação na prática da temática que está sendo investigada”.

## 5.2 Etapas da pesquisa, participantes e procedimentos

Organizou-se o processo de pesquisa em duas etapas:

### **Etapa 1: Versão piloto de Katri**

- Criação e documentação do processo de produção da versão 1 (piloto) do *chatbot*.
- Acompanhamento das interações com a versão 1 do *chatbot* didático por um grupo de quatro alunos, com coleta de dados com usuários (alunos) por meio do método *Think Aloud Test* (roteiro do teste no Apêndice 2).
- Documentação das decisões dos produtores do *chatbot* a partir dos resultados do *Think Aloud Test*.

### **Etapa 2: Versão definitiva de Katri**

- Criação de matriz com critérios e descritores para avaliação de *chatbots* didáticos.
- Uso do *chatbot* com a turma (32 alunos) e análise de dados das interações de estudantes com a versão definitiva do *chatbot*, por meio de banco de dados do sistema.
- Avaliação da versão definitiva do *chatbot*, com preenchimento da matriz proposta.

A coleta de dados foi feita com produtores do *chatbot* e usuários (alunos) por meio de entrevistas semiestruturadas e questionários (roteiro no Apêndice 3).

Nos quadros 4 e 5, são apresentadas as ações a serem realizadas nas duas etapas, assim como as pessoas envolvidas em cada ação e os procedimentos de coleta e análise de dados. Os procedimentos de coleta de dados são detalhados após a apresentação de cada quadro, enquanto os de análise de dados serão descritos em seção posterior neste capítulo.

**QUADRO 4 - Etapa 1: versão piloto de Katri**

Ações	Participantes	Procedimentos de coleta de dados
Criação do <i>chatbot</i>	Equipe de produção: desenvolvedor de sistema, designer, antropólogo, linguista, coordenador-pesquisador	Observação com registro por notas de campo. Documentação de e-mails trocados entre a equipe e arquivos gerados na produção do <i>chatbot</i> .
Uso do <i>chatbot</i> (piloto)	4 alunos da turma de 2018 do curso TLHE. Perfil: Três mulheres e um homem, entre 34 e 45 anos. Todos são professores universitários e lecionam em instituições privadas de ensino superior de Belo Horizonte.	<i>Think Aloud Test</i> . Os alunos interagiram com o <i>chatbot</i> em um computador da equipe de desenvolvimento. Essa interação foi gravada em vídeo com a captura de tela do computador e o áudio das conversas entre participantes e o pesquisador.
Adaptações no <i>chatbot</i> para criação da versão definitiva	Equipe de produção: desenvolvedor de sistema, designer, antropólogo, linguista, coordenador-pesquisador	Com base nos resultados do <i>Think Aloud Test</i> , a equipe de produção tomou decisões quanto às adaptações necessárias no <i>chatbot</i> . A observação do processo foi registrada por notas de campo. Houve também documentação de e-mails trocados entre a equipe e arquivos gerados na produção do <i>chatbot</i> . Foi feita uma análise dos dados coletados a partir de critérios preliminares da matriz de avaliação de <i>chatbots</i> didáticos, que estava em construção no momento de uso da versão piloto.

Fonte: Elaborado pelo autor.

A criação<sup>24</sup> da versão piloto de Katri envolveu como participantes toda a equipe de produção do *chatbot*. Cinco produtores participaram da construção do sistema: um desenvolvedor de *software*, um designer (apenas na fase inicial do projeto), um antropólogo, uma linguista e o coordenador do curso/pesquisador. Os produtores envolvidos na criação do *chatbot* tiveram o seu trabalho observado pelo pesquisador, que por ser um observador participante, esteve presente nas reuniões de trabalho e teve acesso às interações *online*: trocas de e-mails e dados

<sup>24</sup> A produção do *chatbot* foi financiada pelo Instituto Ânima, instituição que oferta o curso em que o MD foi usado.

registrados no sistema de construção do *chatbot*. O processo foi registrado por notas de campo pelo pesquisador, que também cuidou da documentação dos e-mails trocados entre a equipe e arquivos gerados na produção do *chatbot*, como planilhas com o conteúdo a ser disponibilizado via *chatbot*, apresentações de slides usadas nas reuniões e arquivos de texto trocados entre os membros da equipe.

Todos os alunos da turma de 2018 do curso TLHE foram convidados a testar e avaliar a versão piloto do *chatbot* pessoalmente no escritório em que o sistema é desenvolvido, em Belo Horizonte. Os quatro alunos que tinham disponibilidade participaram do teste. Foram três mulheres e um homem, com idades entre 34 e 49 anos. Todos são professores universitários e lecionam em instituições privadas de ensino superior de Belo Horizonte.

Esses quatro participantes foram convidados a participar do *Think Aloud Test* (SCHMIDT; DELAZARI, 2010), método que consiste na observação e análise da verbalização de usuários ao interagir com um sistema. Neste teste de usabilidade, os alunos interagiram com o *chatbot* e falaram em voz alta o que faziam ou pretendiam fazer no sistema, quais os resultados dos seus cliques e outras interações. Eles também contaram quais são as suas opiniões sobre os resultados da interação.

Neste método, durante o teste, os produtores dão algumas tarefas aos usuários. As tarefas foram:

- Escreva uma pergunta à Katri, sobre um assunto do curso que você tem interesse em obter informação.
- Volte ao menu inicial.
- Clique na opção do menu que mais lhe interessa e busque o que você quer descobrir sobre o tema.

Os produtores também podem fazer perguntas. As perguntas feitas a todos os usuários foram:

- As opções que aparecem no menu inicial são as mais relevantes ou você acha que o menu deveria ser diferente?
- Há alguma informação que você não encontrou e deveria ser acrescentada ao *chatbot*?
- O que você acha da imagem ou personalidade da Katri?
- Você prefere interagir com a Katri clicando nas opções de menu ou digitando as suas perguntas?

Além dessas tarefas e perguntas, a interação ao longo teste era livre e o diálogo entre produtores e usuários também foi livre, com o conteúdo da interação surgindo espontaneamente de acordo com as escolhas feitas por cada usuário ao longo do teste. A tela do computador foi gravada em vídeo para registrar a interação realizada com o *chatbot* e o áudio com o diálogo entre produtores e participantes também foi registrado em gravador de áudio.

Segundo Shimidt e Delazari (2010, p. 582), este método “proporciona ao avaliador a clareza na identificação de problemas de interpretação e entendimento de cada etapa do processo pelo usuário”. Assim, identificam-se os problemas de interação e de conteúdo do sistema e entende-se o que o usuário aprova e o que gostaria que fosse apresentado de maneira diferente.

Após a avaliação da versão piloto do *chatbot*, foram observadas quais adaptações foram realizadas no *chatbot* pela equipe de produção para que Katri se apresentasse em sua versão definitiva. Os cinco integrantes da equipe de produção participaram novamente desta etapa. O processo de adaptação foi registrado por notas de campo pelo pesquisador. E-mails trocados entre a equipe e arquivos gerados na produção do *chatbot* também foram documentados pelo pesquisador. Após a realização dessas ações, o pesquisador iniciou a etapa 2 da pesquisa.

**QUADRO 5 - Etapa 2: Versão definitiva de Katri**

Ações	Participantes	Procedimento
Criação de matriz completa para avaliação e produção de <i>chatbots</i> didáticos.	Pesquisador	Criação de critérios e descritores a partir de pesquisa bibliográfica.
Uso do <i>chatbot</i> com a turma e análise das interações com o sistema (pós uso).	32 alunos do curso TLHE. Perfil: 16 mulheres e 17 homens, entre 31 e 60 anos. Todos são professores universitários e lecionam em instituições públicas e privadas de ensino superior dos estados de Minas Gerais, São Paulo, Rio Grande do Sul e Santa Catarina	Apresentação dos dados de interação dos alunos com o sistema (o pesquisador obteve as informações via banco de dados do <i>chatbot</i> ); Aplicação de questionário (Apêndice 3) de avaliação do <i>chatbot</i> pelos alunos;
Avaliação pelos produtores (pós uso)	Equipe de produção: desenvolvedor de sistema, designer, antropólogo e linguista.	Entrevista semiestruturada para avaliação do <i>chatbot</i> (Apêndice 4).

Fonte: elaborado pelo autor

Nesta pesquisa, criou-se uma matriz de categorias e descritores específicos para avaliação de *chatbots didáticos* utilizando-se como base as categorias e descritores já cunhados por autores como Cabero (2001), Belisário (2003) e Kuligowska (2015). A matriz, assim como as

referências que embasaram a sua criação, foi apresentada no Capítulo 3.

A versão definitiva do *chatbot* foi disponibilizada para uso de todos os alunos da turma de 2019 do curso TLHE. A interação dos participantes com o *chatbot* foi observada pelo pesquisador via banco de dados do sistema. Todas as interações ficavam registradas e o pesquisador podia verificar quais foram os cliques de cada usuário, assim como o que cada um digitou e enviou para Katri, e quais foram as respostas que o sistema forneceu a eles. Os dados de interações registrados no sistema do *chatbot* serviram para identificação do nível de engajamento dos alunos com o *chatbot*, como a frequência de uso, a quantidade de alunos que utilizou mais de uma vez o sistema e quais abandonaram o uso do *chatbot* após a primeira experiência. Também foi verificado o nível de assertividade das respostas que o *chatbot* dá aos alunos. Esses dados serão apresentados, para que se possa entender o contexto de uso real do *chatbot*, mas não são analisados, pois não contém em si elementos que expressam atitudes de avaliação do MD, o que é o foco desta pesquisa.

Além das informações fornecidas pelo sistema, o pesquisador também coletou dados diretamente com os alunos, que responderam a questionários de avaliação pós uso do *chatbot*, de acordo com o perfil de interação de cada um. Esses dados serão analisados. Os alunos foram divididos em 4 perfis: 1 - alunos que não interagiram com o *chatbot*; 2 - alunos que interagiram com o *chatbot*, mas que não voltaram a ter interações ao longo do curso; 3 - alunos que interagiram com o *chatbot* em diferentes momentos ao longo do curso; 4 - alunos que só interagiram com o *chatbot* na etapa final do curso. Os questionários foram disponibilizados *online*, com link enviado por e-mail aos alunos.

O pesquisador também convidou todos os produtores do *chatbot* para participação em uma entrevista semiestruturada (roteiro disponível no Apêndice 3), após a finalização do curso, portanto, após o uso do MD pelos alunos. O objetivo da entrevista foi a avaliação do *chatbot*, pós uso, na perspectiva dos produtores. Os dados obtidos nas entrevistas também serão objeto de análise. Concordaram com a participação na entrevista a linguista, o antropólogo e o desenvolvedor de sistemas. A entrevista com a linguista e o antropólogo aconteceu de maneira síncrona, por videoconferência, em outubro de 2019. O desenvolvedor de sistemas não pôde participar da reunião, mas respondeu às perguntas da pesquisa por e-mail, no mesmo mês e ano.

Cumpriu-se o ciclo de avaliação de *chatbots* duas vezes durante a pesquisa, analisando duas

versões do mesmo *chatbot*. A primeira versão ou versão piloto, foi avaliada pelos produtores e por 4 alunos da turma de 2018; e a versão definitiva desse *chatbot*, avaliada pelos mesmos produtores e pelos alunos da turma do ano de 2019. Analisou-se os resultados desta segunda avaliação também para deixar recomendações de melhorias para uma possível terceira versão do *chatbot*, que por limitações de tempo não fará parte do escopo desta pesquisa, já que durante o ano de 2020 não houve turma do curso.

### 5.3 Análise dos dados

Os dados das avaliações de alunos e produtores sobre Katri foram analisados a partir do Sistema de Avaliabilidade (MARTIN; WHITE, 2005), apresentado no capítulo 4. Os discursos verbalizados no *Think Aloud Test* (por alunos) e nas respostas aos questionários (por alunos) e às entrevistas (por produtores) são objeto desta análise, que visa à validação da matriz de avaliação de *chatbots* didáticos.

Nos discursos de avaliação dos participantes sobre Katri, buscou-se elementos linguísticos que expressam Atitudes e a possível Gradação dessas Atitudes, sendo elas positivas ou negativas. Foi investigada a presença de textos que expressem Afeto ou Apreciação e marcadores de Força ou Foco. Para que essa análise fosse feita, construiu-se um quadro para anotação das atitudes e gradação expressas pelos alunos e produtores na avaliação de Katri.

**QUADRO 6 - Atitudes e Gradação na Avaliação de Katri**

Categoria 1: Conteúdos						
Descritor: Conteúdos especializados						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
Avaliação por produtores	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco

Fonte: Elaborado pelo autor

Os atributos identificados no quadro são:

**Avaliação:** dividida em avaliação por alunos e avaliação por produtores. Este atributo é usado somente para que se possa fazer distinção entre os dois grupos de avaliadores do MD.

**Avaliadores:** Quantidade de avaliadores que expressaram atitudes avaliativas sobre o descritor nos questionários ou entrevistas. O número total de avaliadores que será preenchido no quadro não é igual ao número de participantes, pois não há necessariamente atitudes avaliativas sobre todos os descritores no discurso de cada participante. Há participantes não sabiam como responder a algumas perguntas. Outros, mesmo respondendo a todas as perguntas, expressam atitudes avaliativas relacionadas apenas a parte dos descritores. Sendo assim, um descritor pode ter sido alvo de avaliação de 18 alunos, e não dos 32 da turma, por exemplo.

**Polaridade:** Categorização das atitudes como positivas ou negativas de acordo com o discurso dos textos avaliativos.

**Afeto:** Identificação de Felicidade, Segurança e/ou Satisfação em relação à Katri nos discursos dos participantes. Pode haver ou não expressões de Afeto nas avaliações.

**Apreciação:** Identificação de apreciação do Valor, Reação ao Impacto, Reação à Qualidade, Composição de Equilíbrio e/ou Composição da Complexidade de Katri nos discursos dos participantes. Pode haver ou não expressões de Apreciação nas avaliações.

**Força:** Identificação de elementos textuais usados para aumentar ou diminuir a força das Atitudes envolvidas nos discursos avaliativos. No quadro, marca-se “+” para aumento da força e “-” para diminuição da força.

**Foco:** Identificação de elementos textuais usados para ampliar ou atenuar o foco nas Atitudes envolvidas no discurso avaliativo. No quadro, marca-se “+” para ampliação do foco e “-” para atenuação do foco.

Atributos não identificados nos discursos avaliativos de alunos e produtores constarão no quadro, mas serão mantidos sem preenchimento.

Os descritores da matriz de avaliação de *chatbots* didáticos foram divididos em 2 categorias: 1) descritores para avaliação, que demandam a análise qualitativa do discurso dos participantes e 2), descritores para verificação, que demandam exclusivamente a verificação se a função relacionada ao descritor está ou não implementada no sistema, o que dispensa a necessidade de

avaliação por produtores e alunos. Cada descritor para avaliação, entretanto, foi analisado a partir da identificação dos atributos, nos discursos de produtores e alunos, conforme exemplo apresentado no Quadro 7.

**QUADRO 7 - Exemplo de preenchimento do quadro de descritores para avaliação**

<b>Catogoria 1: Conteúdos</b>						
<b>Descritor: Conteúdos especializados</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	1	Negativo		Valoração		
	2,3,5,7,9, 11	Positivo				
	8	Positivo			+	
	10	Positivo				-
	14, 15, 17	Negativo				-
	16	Negativo				
	18	Positivo	Segurança			
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Antropólogo	Positivo		Valoração	+	
	Linguista Desenvolvedor	Positivo				

Fonte: Elaborado pelo autor

No capítulo de análise, além do preenchimento do quadro com os resultados das avaliações, serão apresentados, a fim de exemplificação, alguns excertos das avaliações de cada descritor da matriz para que o leitor desta tese possa ter melhor entendimento de como as avaliações foram expressas nos discursos. Por exemplo: “o produtor-antropólogo expressou Apreciação positiva da Composição de Katri, ao avaliar que o conteúdo é “rico”, com aumento da Força por meio dos advérbios de intensidade “bem” e “muito”.

Essas avaliações dos participantes darão base para que o pesquisador avalie cada descritor da matriz com os indicadores Adequado ou Implementado, Parcialmente Adequado ou Parcialmente Implementado, ou Inadequado ou Ausente.

A título de exemplificação, o descritor Conhecimentos Especializados, conforme o quadro 7, possui majoritariamente avaliações positivas de alunos e produtores, o que sugere que não há evidências nos dados coletados que suportam a avaliação do descritor como Inadequado. O pesquisador-produtor, com base na análise qualitativa de textos avaliativos de alunos e produtores, então, tomará a decisão de avaliar o descritor como Adequado ou Parcialmente

Adequado e fará a marcação do resultado de sua avaliação na matriz, conforme exemplifica o quadro 8.

**QUADRO 8 - Exemplo de preenchimento da matriz de avaliação**

Categoria: Conteúdos		
Descritor	Questões norteadoras	Avaliação
Conhecimentos especializados	O quão completo é o conhecimento do <i>chatbot</i> sobre os conteúdos trabalhados no curso ou disciplina que ele atende, incluindo conteúdo (conceitos, teorias etc), atividades e agenda do curso. Os conteúdos são válidos e embasados cientificamente?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado

Fonte: Elaborado pelo autor

A análise dos dados com a utilização do Sistema de Avaliatividade é o que dá base para uma avaliação não arbitrária, mesmo que subjetiva. O preenchimento da avaliação dos descritores com base na análise linguística das avaliações de alunos e produtores irá validar a matriz proposta e a metodologia de análise dos *chatbots* didáticos.

#### 5.4 - Considerações sobre ética na pesquisa

É essencial considerar a ética em todas as etapas de uma pesquisa científica, principalmente quando ela envolve seres humanos, como é o caso desta tese que apresenta entrevistas, questionários e observação do processo de criação e uso de um MD por produtores e alunos. Por isso, deu-se atenção aos cuidados necessários para “garantir que os interesses e bem-estar das pessoas não venham a ser prejudicados, em razão da realização da pesquisa” (LANKSHEAR; KNOBEL, 2008, p. 93).

Paiva (2005, p. 48) elenca alguns pontos para assegurar o tratamento ético à pesquisa com seres humanos, como a “proteção aos indivíduos ou pessoas cujos pontos de vista ou vozes possam ser identificados” e a “consciência de que a pesquisa pode ser intrusiva”. Para dar a devida atenção a esses pontos, eles foram considerados nos Termos de Consentimento Livre e Esclarecido (TCLE) apresentados aos produtores e alunos/usuários do *chatbot* Katri (ver Apêndice I). Os procedimentos éticos adotados nesta pesquisa, inclusive a redação e forma de apresentação dos TCLE, foram apresentados e aprovados pelo Comitê de Ética em Pesquisa do

CEFET-MG.<sup>25</sup>

A todos os participantes foi solicitada a leitura e assinatura dos TCLEs. O pesquisador foi pessoalmente em sala de aula no primeiro dia de aula do curso para apresentar os objetivos da pesquisa, os riscos e benefícios, e solicitar que os alunos que desejassem participar lessem e assinassem o TCLE. Todos os alunos concordaram com a participação. O pesquisador também conversou por videoconferência com cada produtor do *chatbot* em uma de suas periódicas reuniões de trabalho para apresentar os objetivos da pesquisa, os riscos e benefícios, e solicitou que, se desejarem participar, lessem e assinassem o TCLE. Os produtores também concordaram com a participação na pesquisa.

Paiva (2005, p. 54) também aponta outra importante questão ética: “A Instituição e os informantes/participantes estão devidamente informados sobre os objetivos da pesquisa?”. Para garantir a devida informação à Instituição pesquisada, além de reunião entre o coordenador e diretor da Instituição para apresentação da pesquisa, foi assinada pelo Instituto Ânima uma declaração em que os objetivos de pesquisa estão claramente expostos.

Lankshear e Knobel (2008, p. 94-105) apresentam uma lista de princípios éticos que devem ser seguidos em pesquisa: 1) a construção de um projeto de pesquisa válido; 2) a obtenção do consentimento dos participantes da pesquisa para a coleta de dados; 3) a transparência de informações; 4) a minimização da invasão à privacidade dos participantes e invasões que gerem constrangimentos aos participantes; 5) a garantia de confiabilidade; 6) o anonimato dos dados dos participantes; 7) a demonstração contínua de respeito aos pesquisados, evitando coerção e manipulação dos relatos. Ressalta-se que todos esses princípios éticos foram observados nesta pesquisa, tanto na etapa de planejamento, quanto na coleta e análise dos dados, assim como na escrita deste texto.

Apresentados os procedimentos metodológicos e éticos utilizados na pesquisa, passa-se a partir do próximo capítulo à apresentação os resultados. Primeiramente, apresenta-se a análise dos dados referentes à versão piloto de Katri. Em seguida, analisa-se os resultados da versão definitiva do *chatbot* didático.

---

<sup>25</sup> CAAE: 19861219.5.0000.8507

## 6. RESULTADOS

Antes da apresentação dos resultados da versão definitiva de Katri, considera-se pertinente contextualizar como esse *chatbot* didático foi criado. Assim, no item 6.1 descreve-se o processo de produção da versão piloto de Katri, assim como a avaliação feita a partir do *Think Aloud Test* com alunos do curso e as decisões da equipe de produção relativas às melhorias a se implementar no *chatbot*. No item 6.2, apresenta-se os resultados desta pesquisa, com a validação da matriz de avaliação de *chatbots* didáticos por meio da avaliação da versão definitiva de Katri por alunos e equipe de produção, analisadas a partir do Sistema de Avaliatividade (MARTIN; WHITE, 2005).

### 6.1 Piloto: primeira versão de Katri

Katri “nasceu”, ou foi criada, porque o curso investigado era gerido por uma equipe de Inovação Acadêmica, da qual o pesquisador/coordenador do curso faz parte. Portanto, além da coordenação do curso o pesquisador tem como função, na instituição em que atuava, apresentar propostas inovadoras e testar a implementação de metodologias e instrumentos para inovação em educação. Após a aprovação da proposta de criação de um *chatbot* para o curso TLHE, iniciou-se o processo necessário para implementar o projeto. Para isso, foi contratada uma empresa desenvolvedora de *softwares*, especializada em *chatbots*. Parte dos produtores de Katri são funcionários dessa empresa: o desenvolvedor de *softwares*, o designer e o antropólogo. A outra parte da equipe de produção, o pesquisador/coordenador do curso e uma professora de linguística, especializada em tradução inglês-português, trabalharam no projeto como membros da instituição de ensino.

Quando o grupo se juntou pela primeira vez, não havia o nome Katri e os integrantes da equipe de produção não sabiam ainda como deveria ser o *chatbot*. A única certeza era que fariam um projeto prático para investigar como essa tecnologia digital poderia auxiliar o trabalho docente e os estudos dos alunos do curso. A primeira reunião da equipe foi realizada em 4 de setembro de 2018. Nesta reunião, chamada de *kick off* do projeto, os integrantes discutiram e definiram premissas para a criação do *chatbot* e organizaram o cronograma de trabalho.

Por questões de limites orçamentários, ficou definido que o *chatbot* poderia tratar de temas relacionados a até 75 materiais didáticos já existentes no curso (textos, livros completos, vídeos,

dentre outros) e que o sistema seria treinado para responder a até 120 diferentes questões para os alunos-usuários. Esse foi o escopo do projeto.

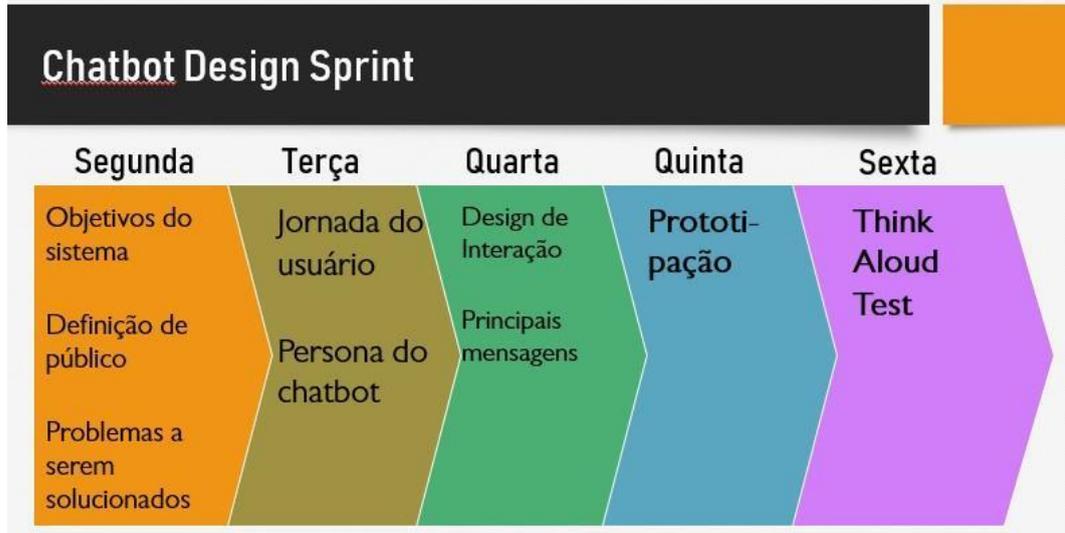
Também foi definido que uma versão piloto do *chatbot* precisava ficar pronta até o fim do mês de setembro de 2018 para ser testado com alunos que ainda estavam cursando a turma daquele ano do curso. A versão definitiva do *chatbot* deveria ser desenvolvida até o início de março de 2019, para ser usada a partir do primeiro dia de aula da turma seguinte do curso.

O desenvolvimento do *chatbot* didático foi dividido em três etapas. A primeira relativa ao treinamento do *chatbot* (processo onde a inteligência artificial é alimentada com o conteúdo da disciplina) para o módulo 1 do curso usando uma interface temporária para o sistema. Na segunda etapa, foi construída a versão final da interface do sistema. Na terceira, foi completado o treinamento do *chatbot* com os conteúdos dos três módulos restantes do curso.

O coordenador do curso e a professora-linguista ficaram responsáveis por preparar os conteúdos do *chatbot*, enquanto a equipe da empresa de desenvolvimento ficou responsável por definir questões relacionadas à programação e *layout* do sistema. Com todos os itens preparados pelos integrantes da equipe, foi realizada, entre os dias 17 e 21 de setembro de 2018, uma semana intensiva de trabalho para construção da primeira versão do *chatbot*, focada neste momento apenas no primeiro módulo do curso, usando o método *Design Sprint*.

O *Design Sprint* é uma metodologia prática e colaborativa centrada no usuário. Baseia-se em metodologias ágeis para que as equipes criem e prototipem rapidamente soluções para problemas. Ao aplicar o pensamento criativo e crítico para compreender, visualizar e descrever problemas, torna-se uma ferramenta muito útil para desenvolver abordagens práticas focadas na resolução de problemas (BROWN, 2010; SILVA, 2018). Foi desenvolvida por um braço da Google, a Google Ventures, e segue cinco etapas principais: “(I) entendimento e definição do problema, (II) divergência e proposição de diferentes ideias, (III) decisão por uma ideia através de votação, (IV) prototipação de média/alta fidelidade e (V) validação do produto com usuários potenciais dando *feedback* real sobre a experiência de uso” (SILVA, 2018, p. 194).

Durante a semana de *Design Sprint*, para prototipar o *chatbot*, a equipe passou por essas fases, dividindo os trabalhos da seguinte maneira:

FIGURA 14 - Semana de Design Sprint do *Chatbot*

Fonte: Dados da pesquisa

### Atividades realizadas no Design Sprint

**Segunda-feira:** foi discutido e definido que o objetivo<sup>26</sup> de uso do *chatbot* seria diminuir o tempo que professores e coordenadores dedicam a responder dúvidas da operação do curso (número das salas em que acontecem as aulas, datas de entrega de atividades etc.) e dúvidas de conteúdo previsíveis (que muitos alunos de turmas anteriores já tiveram e que, provavelmente, seriam pontos de dificuldade também para os novos alunos do curso).

Foi considerado que o perfil do público-alvo do *chatbot* seria o mesmo público-alvo do curso (perfil já definido anteriormente pela instituição ofertante): professores universitários, entre 30 e 50 anos, com mais de cinco anos de experiência docente, das Regiões Sul e Sudeste do Brasil (regiões em que a instituição que oferta o curso possuía unidades naquele ano), e com proficiência na língua inglesa (pois as professoras finlandesas não falam português e o curso não conta com o trabalho de tradutores). O *chatbot* poderia ser usado por pessoas que não fazem parte desta definição de público-alvo, mas por se considerar que seria usado majoritariamente com pessoas desse perfil, visto que é esse o perfil dos matriculados em todas as turmas do curso

<sup>26</sup> Não confundir o objetivo do uso do sistema com os objetivos desta pesquisa, já descritos na seção de Introdução. Esses eram os objetivos internos da equipe de criação do sistema. Os objetivos do pesquisador para construção desta tese de doutorado são exclusivamente os descritos na Introdução.

até aquele momento, os testes foram feitos com esse público.

Com base nas experiências anteriores do coordenador do curso com as outras turmas, foram definidos os principais problemas a se resolver. Foram elencados 68 temas relacionados ao curso sobre os quais o *chatbot* deveria saber responder. Esses temas estão relacionados aos conteúdos que devem ser estudados pelos alunos, às atividades a serem realizadas e também a questões administrativas do curso, como o calendário acadêmico e como acessar o Moodle.

FIGURA 15 - Exemplos de temas tratados pelo *chatbot*

O chatbot precisa dar resposta a questões relacionadas a	#		
Atividades por fazer	1	Cognitivismo	35
Calendário acadêmico	2	Behaviorismo	36
O que é o projeto pedagógico	3	Socioconstrutivismo	37
Planejamento de Proposta	4	Motivação	38
Implementação do Projeto Pedagógico	5	Engajamento	39
Projeto Final	6	Estratégias de Engajamento	40
Datas Projeto Pedagógico	7	Expertise	41
Exemplo Projeto Pedagógico	8	Alinhamento Construtivo	42
Quais são as tarefas individuais	9	Taxonomia Solo	43
Quais são as tarefas em grupo	10	Tem Plus Two	44
		KWL Strategy	45

Fonte: Dados da pesquisa

**Terça-feira:** neste dia também foi discutida a persona do *chatbot*, ou seja, que personagem seria criado para representar o papel de assistente virtual do professor, considerando questões sobre sua aparência física e sua personalidade. Decidiu-se que seria uma mulher, já que todas as professoras finlandesas do curso são mulheres, que deveria ter aparência e nomes que remetessem à Finlândia. Após conversarem sobre vários nomes possíveis, por maioria de votos a assistente virtual foi nomeada Katri, um nome comum naquele país.

Também ficou decidido que Katri assumiria o papel de monitora do curso, ou seja, uma aluna que auxilia os estudos de outros alunos. Sobre este papel, discutiu-se que se Katri fosse uma professora ou coordenadora do curso, o aluno poderia esperar que ela soubesse responder a todas as questões de conteúdo ou administrativas, mas como um *chatbot* vai se aprimorando e aumentando a base de dados/conhecimento com o tempo e com o número de interações realizadas, sabia-se que nas primeiras turmas em que fosse usado, muitas questões poderiam ainda não ter respostas. O descompasso entre o que se espera de um *chatbot* e o que ele é capaz

de responder é um fator que pode gerar desengajamento do usuário, segundo os produtores. Por isso, o papel de monitora foi pensado para tentar alinhar as expectativas dos alunos com a base de conhecimentos de Katri. Assim, a suposição da equipe foi de que a falta de algumas respostas poderia ser melhor compreendida caso Katri também estivesse estudando o curso e se assumisse como uma aprendiz, não como uma professora.

O designer que trabalhou no projeto apresentou possibilidades de desenho da imagem de Katri. A escolhida foi a imagem mostrada a seguir. Alguns aspectos de sua persona também foram definidos:

**FIGURA 16 - Identidade de Katri**



Fonte: Dados da pesquisa

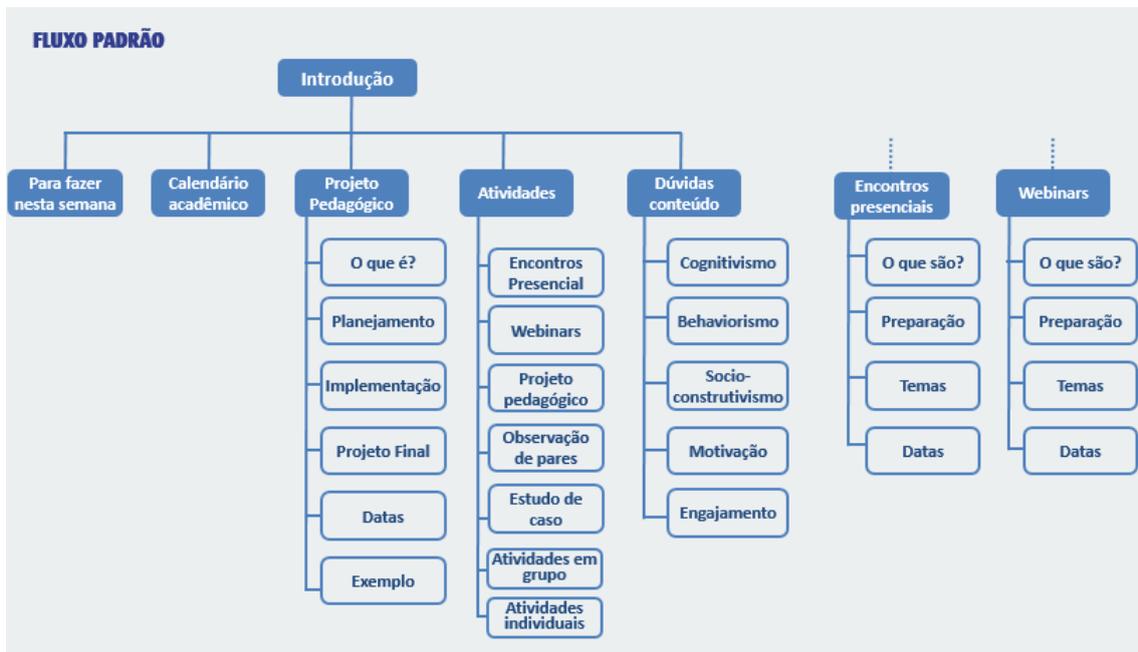
Além das questões relacionadas à persona, neste dia também foi discutida a jornada do usuário, que é a previsão de como os alunos irão interagir com o *chatbot*, considerando desde o momento em que eles descobrem que o *chatbot* existe até a maneira como eles usam o sistema (onde clicam, quais informações aparecem, o que eles podem fazer no sistema etc).

**Quarta-feira:** foi definida a hierarquia dos problemas a se resolver, ou seja, quais problemas eram os mais relevantes para que pudessem ter destaque no *chatbot*. Assim, na página inicial do *chatbot* haveria um menu já sugerindo os temas considerados mais relevantes para os alunos. O coordenador, com base nas experiências com as turmas anteriores, escolheu cinco temas que deveriam estar no menu inicial do *chatbot*: 1) Para fazer nesta semana; 2) Calendário

Acadêmico; 3) Projeto Pedagógico; 4) Atividades; 5) Dúvidas sobre Conteúdos.

Nesse dia, foram também definidas questões sobre o Design de Interação do sistema. Por exemplo, foi definido que os alunos/usuários poderiam digitar suas dúvidas, mas que as questões já previstas também estariam disponíveis por meio botões nos quais os alunos poderiam clicar para ver as respostas. Como muitos itens são complexos e longos, foram divididos em camadas de informação. Um fluxo (Figura 15) foi desenhado para representar como a informação estaria distribuída hierarquicamente no *chatbot*, sendo que para cada item (caixa) do fluxo haveria um botão que, se o aluno clicasse, veria um texto explicativo sobre o tema a ele correspondente.

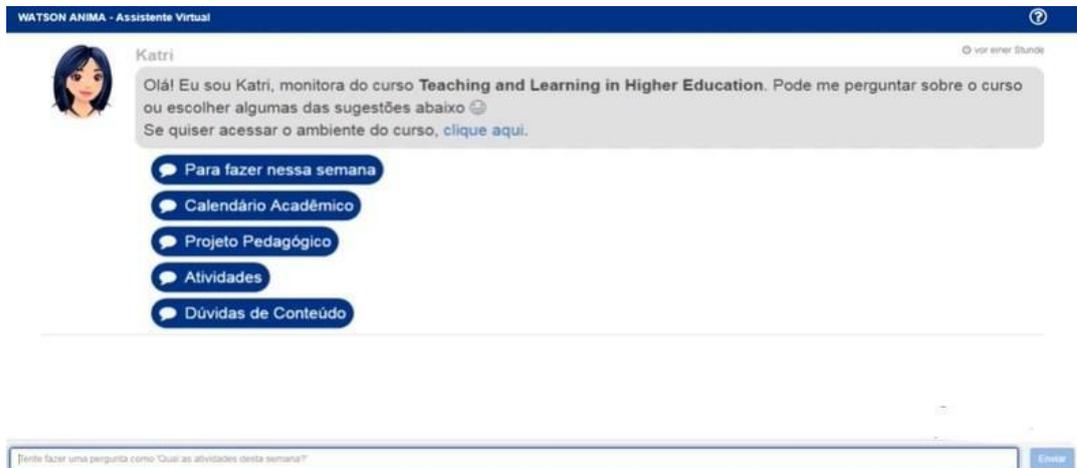
**FIGURA 17 - Fluxo para representar a hierarquia de informação do *chatbot***



Fonte: Dados da pesquisa

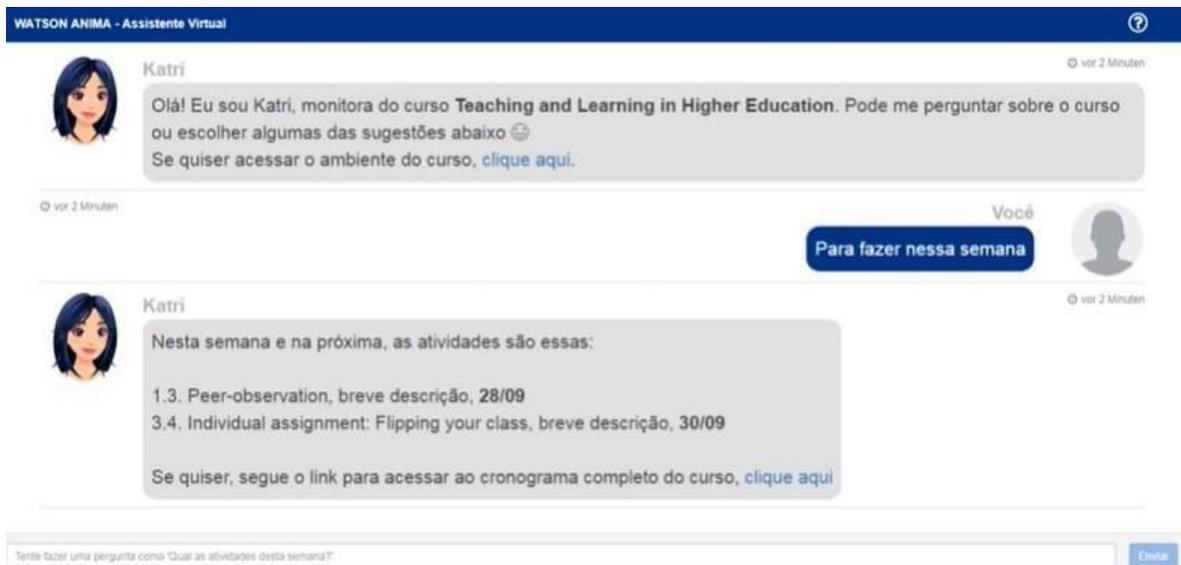
**Quinta-feira:** com base nas decisões tomadas nos dias anteriores, foi construído o primeiro protótipo do *chatbot*, uma versão piloto do sistema para ser testada com usuários reais.

**FIGURA 18 - Piloto da página inicial do chatbot**



Fonte: Dados da pesquisa

**FIGURA 19 - Tela do piloto de Katri respondendo à questão Para Fazer Nessa Semana**



Fonte: Dados da pesquisa

**Sexta-feira:** no último dia do *Design Sprint*, foi realizado o teste da versão piloto do chatbot com quatro usuários/alunos da turma do ano de 2018, que era formada por 32 alunos, usando a Metodologia *Think Aloud Test*. O teste e seus resultados serão descritos a seguir.

### 6.1.1 Teste e avaliação da versão piloto de Katri

No dia 21 de setembro de 2018, quatro alunos do curso TLHE testaram a versão piloto do *chatbot*. Este primeiro momento em que Katri foi mostrada ao seu público-alvo tinha o objetivo de validar as escolhas realizadas pelos produtores e coletar informações para melhoria do *chatbot* didático. Toda a turma foi convidada a participar do teste, que aconteceu presencialmente em Belo Horizonte na manhã de 21/09/2018. Os quatro participantes (12,5% da turma) foram os que se propuseram a ir à sede da produtora do *chatbot* para participar da experiência.

Neste teste, foi observada e analisada a verbalização de alunos ao interagir com Katri. Enquanto usavam o *chatbot*, eles deveriam falar em voz alta o que faziam ou gostariam de fazer na interação com o sistema. Por exemplo, poderiam dizer o que estavam buscando, o que digitaram ou onde clicaram, quais informações ou respostas conseguiram obter e quais são as suas opiniões sobre os resultados da interação. O pesquisador e o produtor-antropólogo também fizeram perguntas aos usuários sobre o processo de interação com o sistema. O conteúdo das falas dos participantes foi analisado de acordo com o Sistema de Avaliatividade.

Neste piloto, 11 descritores da matriz de avaliação de *chatbots* didáticos foram analisados, sendo que seis tiveram como base para avaliação qualitativa os discursos dos alunos no Think Aloud Test, com posterior validação/avaliação pelos produtores, e cinco descritores foram avaliados objetivamente por meio da verificação se o recurso ou conteúdo está implementado/presente ou ausente no *chatbot*.

**QUADRO 9 - Descritores analisados no piloto**

<b>Categoria</b>	<b>Descritores</b>
1 - Conteúdos	Antecipação às dúvidas mais comuns dos estudantes Conhecimentos especializados Conexão com outros conteúdos do curso
2: Discurso e linguagens	Clareza Suficiência de informações
3 - Navegação e Usabilidade	Opções de menu e de digitação simultâneas
4 - Aparência e Persona	Cumprimento Persona do <i>chatbot</i> Informações básicas sobre a identidade do <i>chatbot</i>
5 - Funcionalidades adicionais	Resposta a questões desconhecidas Opção para voltar ao menu inicial

Fonte: Dados da pesquisa

Os descritores analisados no piloto foram os relacionados às questões que a equipe de produção do *chatbot* decidiu priorizar, considerando quais tópicos poderiam trazer dados relevantes para apoiar na decisão de novos conteúdos a se produzir, funcionalidades a se adicionar e possíveis ajustes a se fazer no *chatbot*. A seguir, são apresentados os critérios e descritores avaliados no piloto, com a anotação das atitudes e gradação expressas pelos alunos no *Think Aloud Test*, assim como as transcrições de alguns trechos que exemplificam como os temas foram abordados por eles em seus discursos avaliativos. Apresenta-se também as decisões dos produtores para melhorias e desenvolvimento da versão definitiva de Katri.

### Categoria 1 - Conteúdos

Na categoria Conteúdos, foram avaliados três descritores: Antecipação às dúvidas mais comuns dos estudantes, Conhecimentos especializados e Conexão com outros conteúdos do curso. Primeiramente, será mostrado o resultado da avaliação, já de acordo com a matriz proposta nesta pesquisa. Em seguida, será detalhado como se chegou a esse resultado a partir da análise das avaliações do *chatbot*. O quadro 10 apresenta o resultado das avaliações dos 3 descritores da categoria Conteúdos.

**QUADRO 10 - Categoria: Aparência e persona - piloto de Katri**

Categoria: Conteúdos		
Antecipação às dúvidas mais comuns dos estudantes	As informações ofertadas pelo <i>chatbot</i> via opções de menu são capazes de antecipar e responder às possíveis dúvidas mais comuns dos estudantes?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Conhecimentos especializados	O quão completo é o conhecimento do <i>chatbot</i> sobre os conteúdos trabalhados no curso ou disciplina que ele atende, incluindo conteúdo (conceitos, teorias etc), atividades e agenda do curso. Os conteúdos são válidos e embasados cientificamente?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Conexão com outros conteúdos do curso	Os conteúdos apresentados no <i>chatbot</i> fazem menção e estão diretamente relacionados com outros conteúdos do curso?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Todos os participantes demonstraram concordar que as cinco opções que Katri apresenta no menu inicial são as questões mais relevantes para os alunos do curso, porém gostariam que houvesse uma maneira mais simples para voltar ao menu inicial durante a conversa, conforme indicações do Quadro 11.

**QUADRO 11 -Afeto e Apreciação: menu inicial**

Descritor: Antecipação às dúvidas mais comuns dos estudantes						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	P1	Positiva		Reação		
	P1	Positiva		Valoração	+	
	P4	Positiva		Composição/ complexidade		
	P2	Positiva		Composição/ equilíbrio		
	P1, P2, P3	Negativa		Composição/ complexidade		

Fonte: Dados da pesquisa

Os discursos avaliativos de todos os participantes trouxeram aspectos positivos de Apreciação sobre o menu inicial. A Participante 1 expressou reação e valoração: “Assim está ótimo. Você tem de início as coisas mais prioritárias”. O Participante 4 avaliou positivamente a Composição do menu do *chatbot*: “Desse jeito faz sentido. Não sei, eu não consigo pensar em nada para acrescentar aqui, não”.

Todos, porém, demonstraram ter dificuldades para saber como voltar ao menu inicial, quando já estavam no meio da interação sobre algum assunto específico e gostariam de ver as opções do menu novamente. Os discursos demonstraram avaliação negativa da Composição da organização do sistema. Enquanto conversava com Katri sobre uma atividade do curso, a Participante 1 quis acessar novamente o menu inicial, não conseguiu inferir a maneira de realizar essa tarefa e perguntou “Como é que eu volto?”. Em situação parecida, a Participante 2 questionou “Para ir no menu inicial tem que dar enter? Como que é?”. A Participante 3, além de questionar como voltar ao menu, fez uma sugestão para melhoria dessa funcionalidade. “Bem, depois se eu quiser voltar para a parte inicial com aquelas opções, eu não tenho isso, né? Para eu voltar agora, você sempre vai ter que ficar dando F5<sup>27</sup>? Talvez fosse bom ter um botão para voltar, não sei”.

<sup>27</sup> Comando no teclado para atualizar a página.

Os participantes também falaram sobre os Conteúdos Especializados, ou seja, a base de conhecimentos de Katri sobre os conteúdos trabalhados no curso, atividades, calendário acadêmico etc, conforme mostra o quadro 12.

**QUADRO 12 - Afeto e Apreciação: conhecimentos especializados**

Descritor: Conhecimentos Especializados						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	P1, P2, P3, P4	Negativa		Composição		
	P3	Positiva	Satisfação			+
	P3	Positiva		Reação		
	P3	Positiva	Felicidade			
	P2	Positiva	Felicidade			+
	P3	Positiva		Reação		+
	P3, P4	Positiva		Valoração		
	P4	Positiva		Composição		
	P2	Negativa	Satisfação			

Fonte: Dados da pesquisa

A Participante 2 gostaria que Katri soubesse dialogar sobre todos os conteúdos do curso. “Acho que todos os conteúdos trazem pra gente coisas novas, então eu acho que se tivesse o *chatbot* para tudo seria melhor”. Ao ponderar que de outra maneira “seria melhor”, a participante valorou negativamente o conteúdo do *chatbot*, mas com atenuação do foco ao não demonstrar certeza por meio do uso do modalizador “acho”.

Durante o tempo que interagiram com Katri, os participantes disseram que encontraram alguns conteúdos que buscavam, mas também apontaram a ausência de outros conteúdos. Os principais pontos que renderam avaliações positivas ao *chatbot* didático estão relacionados à facilidade de se encontrar o calendário acadêmico do curso e à funcionalidade Para Fazer Nessa Semana. “Eu entendo que assim, eu enquanto aluna, quando penso em um calendário acadêmico, a minha expectativa é de ver exatamente isso aqui que foi mostrado” (PARTICIPANTE 3).

Acho que Para Fazer Nessa Semana é muito bom, gostei bastante, porque sinceramente isso me incentivaria a vir toda semana conversar com ela. E eu tenho certeza que todas as minhas colegas também. A gente tem um grupo no Whatsapp que toda semana alguém fala assim: ‘gente, o que tem que entregar esta semana?’ Entendeu? Então, isso me incentivaria a vir toda semana. O que eu tenho que fazer? A gente se esquece mesmo. Eu gostaria de explorar isso aqui. Eu posso clicar (o participante clica no item Para Fazer Nessa Semana e lê a resposta de Katri)? Eu gostei muito dessa funcionalidade, é algo que eu usaria exaustivamente (PARTICIPANTE 4).

Todos os participantes fizeram comentários positivos sobre os conteúdos fornecidos por Katri sobre o Projeto Pedagógico a ser desenvolvido durante o curso. “Eu acho que muitas dúvidas surgiram no momento de criação do *Pedagogical project*, no sentido assim, de como eu poderia trabalhar esse conteúdo. Gosto de ver isso aqui” (PARTICIPANTE 2). Ao usarem o verbo “gostar”, tanto a Participante 2 quanto a 4 expressam Felicidade. Ou seja, em relação aos Conteúdos Especializados, expressões de Afeto estão presentes nas avaliações dos participantes, e não apenas Apreciação, como ocorreu com o descritor Antecipação às dúvidas mais comuns dos estudantes.

Entre as principais ausências de conteúdos sentidas pelos participantes estava a programação dos encontros presenciais do curso. A Participante 4 não encontrou essa informação e afirmou que a ausência de respostas após duas tentativas de busca poderia fazer com que ela desistisse de interagir com o *chatbot*. “Eu acho que eu ia desistir. Eu iria falar ‘então tá, deixa eu ir pesquisar na internet, no Google’”.

O descritor Conexão com Outros Conteúdos do curso é considerado um item da matriz para análise objetiva, ou seja, a proposta é verificar se essas conexões estão ou não implementadas, sem necessidade de análise qualitativa dos usuários. Entretanto, a Participante 3 trouxe uma contribuição relacionada a esse descritor. Ela sugeriu a inserção, nas atividades mostradas na função “Para Fazer Nessa Semana”, de um link que, após ser clicado, levasse o aluno diretamente à atividade no AVA do curso. “Uma coisa que eu esperaria encontrar aqui, talvez seria interessante incluir isso, é como aqui nessa atividade de *Peer Observation*. Poderia já ter um link para que eu pudesse clicar e já ir direto para a atividade (no AVA) que eu preciso fazer”.

O descritor Antecipação às dúvidas mais comuns dos estudantes foi considerado Adequado porque ao ver as opções do menu inicial apresentado por Katri, os usuários concordaram que os 5 itens continham as questões mais frequentes sobre as quais gostariam de conversar com ela. Assim, corroboraram com as inferências que tinham sido feitas pelos produtores quando planejaram os itens do menu. Como os usuários ficaram sem saber como voltar ao menu principal após o fim de uma interação, foi planejada uma alteração: na versão definitiva do *chatbot*, quando Katri envia uma resposta, em seguida uma das mensagens enviadas passa a ser uma orientação de como voltar ao menu principal.

Também foi validada a importância da funcionalidade “Para Fazer Nesta Semana”, pois os participantes demonstraram que esse é o item mais relevante para o uso do *chatbot* no dia a dia dos alunos. Assim, ele foi mantido como primeira opção do menu inicial.

O descritor Conhecimentos Especializados foi considerado Parcialmente Adequado porque, a partir da avaliação dos alunos, os produtores confirmaram a suposição de que Katri deveria saber conversar sobre todos os conteúdos do curso, o que ainda não estava programado no sistema. Os conteúdos sobre os quais ela já conseguia conversar foram validados. Por exemplo, o teste provou a importância de ter todas as etapas do Projeto Pedagógico descritas. Para a versão definitiva, os produtores se organizaram para treinar Katri para que ela pudesse responder questões sobre qualquer conteúdo presente na ementa do curso e nos materiais didáticos disponibilizados no AVA.

Já o descritor Conexão com outros conteúdos do curso foi avaliado como Ausente, pois não havia ainda menções, links ou outro tipo de conexão com o AVA e os outros materiais didáticos do curso. Essa conexão, entretanto, foi listada pelos produtores como um item para implementação na versão definitiva de Katri.

## **Categoria 2 - Discurso e Linguagens**

Na categoria Discurso e Linguagens, os descritores avaliados foram a Clareza (quadro 14) e a Suficiência de Informações (quadro 15). Conforme mostra o Quadro 13, ambos foram considerados Parcialmente Adequados, com base na análise das avaliações que serão detalhadas a seguir.

**QUADRO 13 - Categoria: Discurso e linguagens - piloto de Katri**

<b>Categoria: Discurso e linguagens</b>		
<b>Clareza</b>	Os textos são apresentados de maneira clara para entendimento dos alunos?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Suficiência de informações</b>	A quantidade de conteúdo é suficiente para dar apoio aos estudos autônomos dos alunos? A extensão/tamanho dos textos é adequada ao que os leitores querem ou estão dispostos a ler?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado

Fonte: Dados da pesquisa

Em relação ao quão claros são os textos para entendimento dos usuários, duas participantes deram sugestões para mudar a redação do item Para Fazer Nessa Semana, que faz parte do menu inicial. A Participante 2 reagiu positivamente, com atenuação de foco ao modalizar o discurso, ao dizer “eu acho que a linguagem está legal”, mas em seguida avaliou negativamente a composição de Katri, também com atenuação de foco, ao fazer uma sugestão de melhoria.

Suponho que aí são as atividades que eu tenha que fazer nos próximos dias, mas não fica muito claro para mim. Eu clicaria aqui (no item de menu supracitado) para saber se é isso que eu estou pesquisando. Acho que poderia ser Próximas Atividades, Atividades da Semana, alguma coisa assim. (PARTICIPANTE 2)

O Participante 4 ficou confuso com as opções apresentadas por Katri ao tentar encontrar o conteúdo que buscava.

Eu gostaria de saber exatamente como deve ser escrito, o que deve ser escrito no documento do projeto pedagógico e dentre as opções que tem aqui eu estava tentando entender qual delas poderia ser mais próxima. Eu fiquei entre esses dois aqui: projeto final ou implementação. (PARTICIPANTE 4)

**QUADRO 14 - Afeto e Apreciação: Clareza**

<b>Descritor: Clareza</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	P2	Positiva		Reação		-
	P2	Negativa		Composição		-
	P3	Negativa		Reação		-
	P4	Negativa		Composição		

Fonte: Dados da pesquisa

A participante 3, utilizando modalizadores como em “eu, particularmente” e “estou me questionando se”, expressou diferentes tipos de Apreciação ao afirmar que a quantidade de texto de algumas respostas estava grande para o que esperava de uma comunicação nesse tipo de sistema.

É que eu estou lendo esse texto que apareceu aqui no projeto pedagógico. Eu, particularmente, achei que veio muita informação. Eu estou me questionando até que ponto é importante ter tudo isso aqui. Eu sei que a intenção é de ajudar, é de esclarecer, mas quanto mais enxuta, quanto mais limpa for a informação que você estiver tentando transmitir, acho que fica melhor. Eu fico pensando que para não apresentar às vezes textos grandes, assim de uma vez, será que não poderia, por exemplo, permitir que a pessoa fizesse outras perguntas ou talvez subir sub opções dentro do planejamento de uma resposta, porque se a pessoa tiver interesse de saber, ela vai lá e clica ao invés de mostrar toda a informação de uma vez. Fica aí como sugestão.

**QUADRO 15 - Afeto e Apreciação: Suficiência de Informações**

<b>Descritor: Suficiência de Informações</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	P3	Negativa		Complexidade		-
				Valoração		-
				Reação		-

Fonte: Dados da pesquisa

Tanto os descritores Clareza quanto Suficiência de Informações receberam majoritariamente avaliações negativas dos alunos, quando falaram diretamente sobre eles. Entretanto, os produtores ponderaram que vários outros descritores, relacionados aos conteúdos e à organização textual do sistema, como a apresentação dos itens de menu, tiveram avaliações positivas e foram entendidos claramente pelos participantes. Dessa maneira, esses dois descritores se cruzam com as avaliações de descritores mostrados anteriormente. Por exemplo, se as informações ofertadas pelo *chatbot* via opções de menu são capazes de antecipar e responder às possíveis dúvidas mais comuns dos estudantes de maneira adequada, há ao menos relativa clareza e também informações suficientes tanto na redação desses itens do menu quanto nas respostas que Katri oferece após o usuário clicar em um desses itens. Assim, os dois itens foram considerados Parcialmente Adequados (quadro 16). Os produtores de Katri avaliaram alterações que precisam ser realizadas, mas consideraram que há textos já redigidos com clareza e extensão adequadas.

Foi decidido que para a versão definitiva de Katri, os produtores iriam reduzir o tamanho dos textos que consideram longos, seja resumindo as respostas ou dividindo-as em partes, usando para isso submenus. Por exemplo, se o aluno clicar em Atividades no menu inicial, serão abertas opções de um novo submenu para que ele escolha um tipo específico de Atividade, como Atividades em Grupo ou Atividades Individuais. Foi acatado pelos produtores que algumas mudanças de redação deveriam ser feitas, com novas etapas de revisão de texto para garantir mais clareza aos textos.

Outras questões de linguagem, não previstas previamente para coleta de informações no teste, também surgiram espontaneamente nas avaliações. A Participante 1 sugeriu que Katri se comunicasse com os alunos usando textos com mais marcas de personalidade.

Uma coisa que ajuda muito na didática é você sentir que está sendo guiado. Então, usar Suas Atividades em vez de Atividades. Seu Projeto Pedagógico, Nosso

Calendário Acadêmico. Isso aproxima no processo de ensino e aprendizado. O que você vai fazer essa semana? Sempre mais voltado para personalizar, para envolver, para sentir que está dizendo para você especificamente. Eu falo desse lugar que é um lugar que eu aprendi muito como valor em educação, como você considera o outro como uma pessoa e encontra estratégias linguísticas para mostrar que ele é respeitado, que ele é tratado com credibilidade e valor, que ele está junto com você, que ele está no mesmo barco que você. São pequenos detalhes.

Dois participantes tiveram dúvidas sobre qual língua usar ao interagir com Katri, pois os alunos são todos brasileiros, mas o curso é realizado em inglês. “Quando eu estiver lendo um texto e fazendo alguma atividade, com certeza eu poderia ter algum tipo de dúvida sobre o conteúdo e eu iria recorrer a ela (Katri). Os textos (no AVA) estão em inglês e as perguntas eu posso fazer em português?” (PARTICIPANTE 2). O Participante 4, antes de digitar sua primeira pergunta a Katri, perguntou: “Escrevo em português, né?”.

Diante do que foi observado nas interações dos participantes com Katri e em seus discursos de avaliação do piloto, foi decidido que os textos passaram a ter mais marcas de personalidade, como, por exemplo: “Se você quiser saber mais sobre (...)”, “Acho que vai te ajudar a organizar os seus estudos” e “Você quer saber sobre qual Atividade?”.

Os produtores também optaram por escrever “*Pedagogical project*”, em inglês, em vez de “Projeto Pedagógico”, pois foi assim que os participantes se referiram à atividade em parte dos diálogos e o produtor coordenador de curso reconheceu que é assim que se referem à atividade quando buscam informações com ele. “*Peer observation*” também foi um termo em inglês usado pelos participantes ao interagir com Katri em vez de “Observação por Pares”.

Inicialmente, os produtores imaginaram que os alunos escreveriam todas as perguntas com redação exclusivamente em português, mas na prática os alunos, no *Think Aloud Test*, falaram e escreveram como costumam fazer na interação com colegas no curso: apesar da construção dos discursos se dar em português, os nomes dos principais conceitos e atividades, repetidos e lidos diversas vezes em inglês ao longo do curso, são usados em inglês no meio de uma sentença construída em português. No cotidiano do curso, os alunos adotam espontaneamente essa hibridez de uso das duas línguas numa mesma frase, como em “Quando eu devo entregar o *Pedagogical Project*?” ou “Quem será a minha dupla no *Peer Observation*?”. Esses e outros termos normalmente usados em inglês pelos alunos passaram a ser usados também em inglês por Katri, já que a persona é a de uma aluna-monitora do curso e, por isso, os produtores decidiram que ela deve adotar um discurso similar ao dos alunos.

Validou-se também a necessidade de se treinar Katri para entender termos escritos em inglês e em português. Já que os alunos são brasileiros, mas têm acesso a outros materiais didáticos em inglês, eles poderiam fazer perguntas usando alguns termos da maneira como leram nos outros MD do curso. Katri, então, passou a ser bilingue.

#### Categoria 4 - Aparência e persona

Três descritores da categoria Aparência e Persona foram avaliados no piloto: Cumprimento, Persona do *chatbot* e Informações básicas sobre a identidade do *chatbot*. O quadro 16 apresenta o resultado das avaliações.

QUADRO 16 - Categoria: Aparência e persona – piloto de Katri

Categoria: Aparência e persona		
<b>Cumprimento</b>	Há uma mensagem para cumprimentar o usuário? Essa mensagem transmite algum aspecto da persona do <i>chatbot</i> ?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Persona do <i>chatbot</i></b>	Se o <i>chatbot</i> se apresenta como um personagem, ele apresenta informações (verbais ou não verbais) sobre questões de sua identidade, como idade, gênero, profissão, nome etc? Além de características demográficas, ele apresenta características comportamentais, como humor, curiosidade, gentileza etc, expressas por meio de sua imagem e discurso? Os alunos gostam dessa persona?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Informações básicas sobre a identidade do <i>chatbot</i></b>	O <i>chatbot</i> deixa claro qual papel desempenha na situação de interação com o usuário. Por exemplo, é uma colega que vai conversar, um atendente administrativo que vai ajudar a resolver questões de secretaria, uma professora que vai dirimir dúvidas pedagógicas etc? Os usuários aprovam esse papel?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado

Fonte: Dados da pesquisa

Logo no início do teste de cada participante, ao ver a tela inicial do *chatbot* e antes de começarem a buscar informações no sistema, foi perguntado o que eles pensavam sobre a aparência de Katri. Ao longo da interação, outras perguntas foram feitas para se entender como eles avaliam a persona criada e o papel desempenhado por Katri no curso. Foram identificadas expressões de afeto e de apreciação, algumas positivas e outras negativas.

O descritor Cumprimento foi analisado de maneira objetiva, conforme sugerido na matriz. Há um texto pelo qual Katri se apresenta ao usuário e o cumprimenta (ver Figura 15), por isso a avaliação deste critério foi considerada Implementada. Já os descritores Persona do *Chatbot* e Informações básicas sobre a identidade do *chatbot* passaram por avaliações qualitativas. O quadro 8 mostra o resumo das avaliações dos alunos em relação à Persona do *chatbot*. Em seguida ao quadro, são apresentados, à título de exemplificação, alguns trechos de avaliações que levaram à identificação das expressões de afeto e apreciação, assim como marcadores de foco e força.

**QUADRO 17 - Afeto e Apreciação: Persona do *chatbot***

Descritor: Persona do <i>chatbot</i>						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	P1	Negativa		Reação	+	
	P1	Negativa		Reação		
	P1	Negativa		Composição	-	
	P4	Positiva		Valoração		
	P4	Positiva	Felicidade			
	P3, P4	Positiva		Reação		

Fonte: Dados da pesquisa

Nesta versão piloto, Katri apresenta-se visualmente por meio de uma ilustração, conforme mostrado na FIGURA 18 - Piloto da página inicial do *chatbot* FIGURA 18. A Participante 1 expressou Reação negativa ao olhar para Katri, pois a achou “muito perfeita”. Segundo a participante, Katri “parece uma personagem de mangá<sup>28</sup> e elas são muito idealizadas”. O pesquisador perguntou se essa aparência mais real deveria ser representada por meio de desenho, foto ou outra linguagem. A Participante 1 respondeu que deveria ser “uma pessoa desenhada, como fazem em alguns filmes. Não fica nem uma foto e não fica também um desenho. Assim, fica mais real. Quando eu vejo só um desenho eu acho que ela é *fake*”.

A participante 2 disse que “ela me parece jovem, jovenzinha, não sei se jovem demais”. O antropólogo perguntou se “ser jovem demais é bom ou ruim”. A participante afirmou que “não faz muita diferença”. A participante 2 afirmou expressou Felicidade, ao usar o verbo “gostar” ao ler o nome dado à personalização do *chatbot*. “Eu gosto de Katri. Parece que ela é finlandesa, né? É um nome diferente, que não é brasileiro”. O participante 3 expressou Reação positiva:

<sup>28</sup> Gênero de história em quadrinhos japonesa.

“Katri é internacional, não inglesa. Ela é europeia, não sei de onde. É um nome que de fato acho legal. Parece uma abreviação de Catarina”.

Foi definido pelos produtores que Katri se apresentaria aos alunos do curso como uma monitora, ou seja, uma estudante que dá apoio aos colegas, como uma espécie de assistente das professoras. Os participantes leram a mensagem em que Katri se apresenta como monitora e fizeram as suas avaliações, que estão representadas no Quadro 9.

**QUADRO 18 - Afeto e Apreciação: Papel representado por Katri**

Descritor: Papel desempenhado pelo <i>chatbot</i>						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	P2	Positiva		Valoração	+	
	P2, P4	Positiva		Reação	+	
	P3	Positiva		Reação		

Fonte: Dados da pesquisa

A Participante 2 aprovou o papel de monitora e expressou Reação e Valoração, com amplificação de força por meio do uso do adverbio de intensidade “mais”, ao dizer que o papel de monitora é “mais legal” e “mais próximo” dos alunos. Ela se preocupou com a possibilidade de Katri exercer outros papeis, como professora ou coordenadora do curso. “Se fosse de outra forma, poderia passar pela cabeça da gente que podemos estar sendo avaliados pelo *chatbot* também, como se isso fosse um game que eu tivesse que fazer as perguntas certas para saber se eu estou acompanhando o curso ou não, entende?” (PARTICIPANTE 2).

A Participante 3 disse que “a ideia de monitora é algo que eu não vejo como problema, eu acho que assim está ok”. O papel de Katri foi bem avaliado pelo participante 4:

Bom, eu acho legal. Eu imagino que ela não deva saber tudo, ou seja, ser uma monitora é ser alguém que vai intermediar a conversa com os professores, é alguém que vai me ajudar numa parte mais de agenda, de procedimento, ajudar o aluno a entender o que está acontecendo no curso e quando ele deve ou não fazer as coisas. Então, me ajudar na preguiça que eu tenho de ter que ir lá no AVA. Saber que eu posso perguntar para ela e ela vai me dar a resposta certa. Mas eu acho que se você colocar como professora será uma grande responsabilidade. Principalmente quando você fala em dúvidas de conteúdo. Então, assim, se você pergunta para a professora e ela não sabe exatamente alguma coisa, deprecia bastante a ideia de professora. (PARTICIPANTE 4)

O descritor Persona do *chatbot* foi considerado Parcialmente Adequado, pois algumas

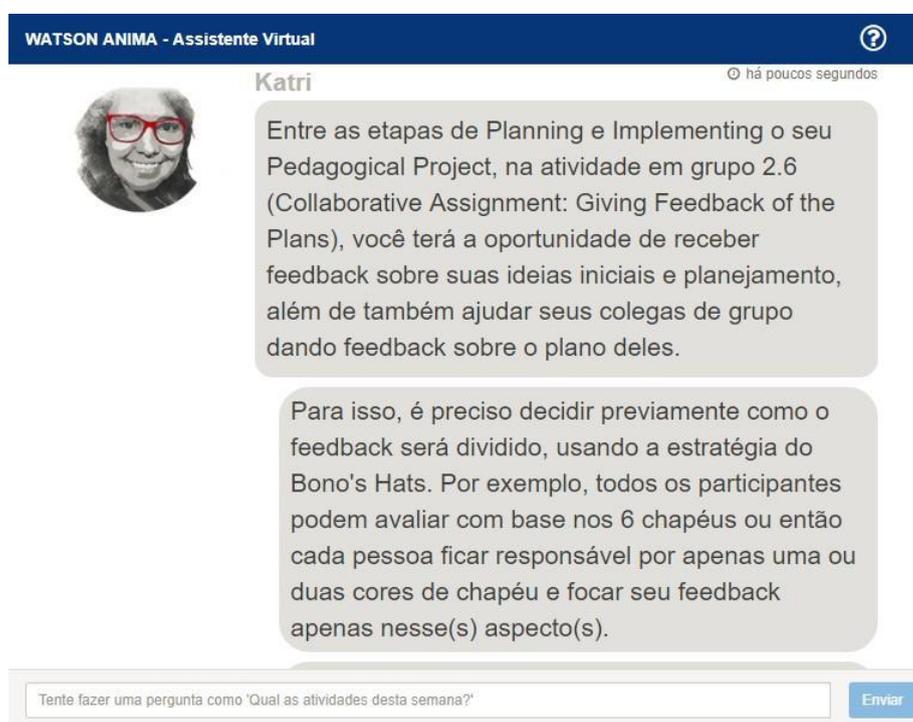
características da persona, como nome e profissão no curso foram validadas pelos alunos, e outras, como a aparência visual, foram criticadas negativamente. Já o critério Papel desempenhado pelo *chatbot* foi considerado Adequado devido à unânime avaliação positiva.

Observou-se que houve expressões de afeto e apreciação em relação à persona, mas apenas de apreciação em relação ao papel desempenhado por Katri, o que não é um fator negativo, mas um ponto a se observar em *chatbots* quando os produtores desejam estabelecer relações que envolvam afetividade com os usuários. Os produtores concordaram com as avaliações dos alunos e decidiram rever questões da imagem de Katri.

Foi validado o nome Katri e que o papel a ser exercido por ela seria mesmo o de monitora do curso, visto que os alunos ficaram confortáveis com esse papel e declararam que outros papéis, como o de professora ou coordenadora, poderiam inibir a interação por gerar dúvida se estariam ou não sendo avaliados no curso durante a interação.

Identificou-se, entretanto, a necessidade de promover mudanças significativas na imagem de Katri, deixando-a mais real, humanizada e adulta. Foi decidido utilizar uma imagem estilizada da produtora-linguista, por considerar que ela representa exatamente o público-alvo do curso, conforme mostra a FIGURA 20.

**FIGURA 20 - Nova imagem de Katri**



Fonte: *chatbot* Katri

A imagem de Katri é mostrada na versão *web*, quando os usuários navegam em um computador para acessar o *chatbot*. Quando o usuário acessa o *chatbot* por celular, a versão *mobile* não mostra a imagem de Katri, devido ao tamanho reduzido da tela.

### **Categorias 3 e 5 - Navegação e Usabilidade e Funcionalidades Adicionais**

Algumas funcionalidades do *chatbot* que necessitavam exclusivamente da conferência da implementação ou não implementação no sistema foram também alvo dos diálogos com os usuários no teste para que pudesse ser entendida a relevância dessas funcionalidades para o público-alvo de Katri. Por não se tratarem de descritores da matriz com avaliação qualitativa, eles não foram analisados a partir do Sistema de Avaliatividade.

O teste buscou coletar impressões dos participantes sobre a maneira como a interação humano-computador é realizada, ou seja, como cada um dos participantes interage com o *chatbot*.

Verificou-se que apesar de acharem positiva a possibilidade de digitar suas dúvidas, há uma preferência por clicar em botões com opções de temas pré-definidos no sistema. “Prefiro clicar nesse balão (do que digitar perguntas). Quanto mais visual for o formato, mais esquematizado, eu acho que fica mais rápido” (PARTICIPANTE 1). Esta avaliação está diretamente relacionada ao descritor Opções de menu e de digitação simultâneas (quadro 17), da Categoria Navegação e Usabilidade. A opção de deixar que os usuários escolham entre digitar suas mensagens ou clicar em botões com opções de menu, seja na tela inicial ou em telas subsequentes, já estava devidamente implementada, mas foi relevante saber a preferência dos usuários. Assim, foi confirmada também a importância de ter opções de botões clicáveis (menus e submenus) para os conteúdos e, simultaneamente, a opção de digitação, ou seja, todos os conteúdos do *chatbot* poderiam ser encontrados por meio da escrita e envio de perguntas usando a caixa de diálogo do *chatbot*.

QUADRO 19 - Categoria: Navegação e Usabilidade - piloto de Katri

Categoria: Navegação e Usabilidade		
Opções de menu e de digitação simultâneas	Ao interagir com o <i>chatbot</i> , os usuários podem escolher entre digitar ou clicar em botões com opções de menu, seja na tela inicial ou em telas subsequentes?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Os participantes deram sugestões sobre como gostariam de interagir com o *chatbot* para sanar problemas que encontraram relacionados à dinâmica da interação. Esses problemas, por vezes, resultavam em não saber qual o próximo passo para continuar a interação, o que interrompia o fluxo do diálogo. Os participantes demonstraram incômodo por não receberem respostas quando faziam perguntas sobre assuntos que Katri não dominava e também por não saberem como voltar ao menu inicial para começar um novo diálogo. “Eu acho que se ela não me responder, primeiro ela tem que depois de um tempo dizer assim ‘olha, infelizmente eu não consegui te ajudar’ e me levar de volta ao menu principal, mas informe que não conseguiu, principalmente” (PARTICIPANTE 4).

Uma coisa que dá um pouquinho de desconforto é que a gente não sabe se acabou ou não acabou a interação, sabe? Quando ela (Katri) responde ‘ok, ok, ok,’ (texto que aparecia na tela enquanto Katri tentava buscar informações em seu banco de dados), se essa for uma resposta padrão para dizer que ela está pesquisando, eu acho que poderia ser um pouquinho diferente o texto. Tipo ‘entendi, vou pesquisar e já te dou um retorno’ para eu saber que vai ter mais alguma interação (PARTICIPANTE 2).

As avaliações das duas participantes estão diretamente ligadas a dois descritores da Categoria Funcionalidades Adicionais (quadro 18) que ainda não tinham sido implementados no sistema: Resposta a questões desconhecidas e Opção para voltar ao menu inicial.

QUADRO 20 - Categoria: Funcionalidades adicionais - piloto de Katri

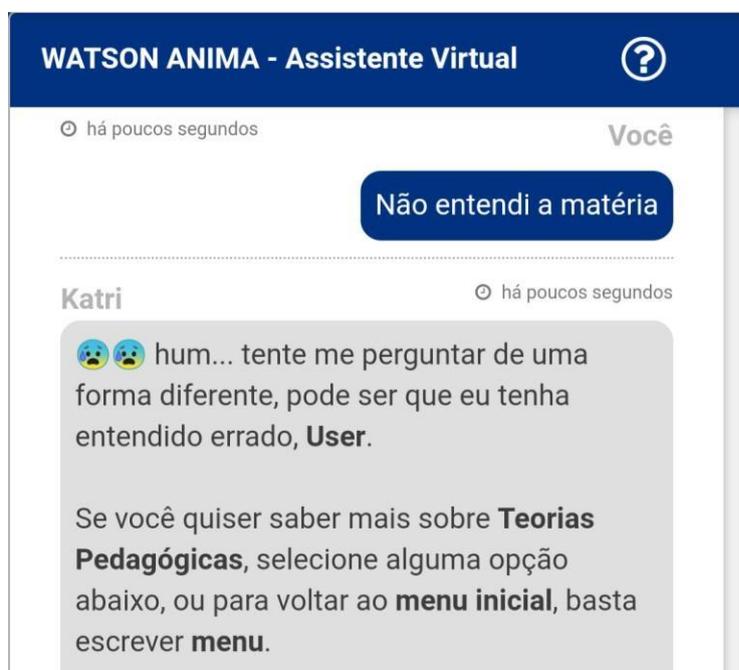
Categoria: Funcionalidades adicionais		
Resposta a questões desconhecidas	Existe uma ou mais respostas-padrão para os casos em que usuário faz uma pergunta para a qual o <i>chatbot</i> não tem resposta?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Opção para voltar ao menu inicial	Há algum comando ou opção de onde clicar para voltar ao menu inicial do sistema?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Como os usuários gostariam de saber se Katri ainda estaria tentando encontrar uma resposta para a pergunta realizada, os produtores decidiram que na versão definitiva do *chatbot*, enquanto a informação é buscada no banco de dados, Katri mostraria reticências (...) na tela do sistema, sinal comumente utilizado na internet para mostrar que um sistema ainda está “pensando”.

Devido ao fato de os usuários apresentarem dúvidas sobre como voltar ao menu principal após o fim de uma interação, na nova versão do *chatbot*, quando Katri envia uma resposta, em seguida uma das mensagens enviadas passou a ser uma orientação de como voltar ao menu principal: “Para voltar ao **menu inicial**, basta escrever **menu**”.

**FIGURA 21 - Trecho de diálogo com orientação sobre como voltar ao menu inicial**



Fonte: *Chatbot Katri*

Essas foram as decisões tomadas pelos produtores após ouvir as avaliações dos participantes sobre a versão piloto do *chatbot*. O quadro 21 compila as avaliações dos 11 descritores contemplados neste teste piloto.

**Quadro 19:** Matriz de avaliação da versão piloto de Katri

QUADRO 21 - Matriz de avaliação da versão piloto de Katri

<b>Categoria: Conteúdos</b>		
Descritor	Questões norteadoras	Avaliação
Antecipação às dúvidas mais comuns dos estudantes	As informações ofertadas pelo <i>chatbot</i> via opções de menu são capazes de antecipar e responder às possíveis dúvidas mais comuns dos estudantes?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Conhecimentos especializados	O quão completo é o conhecimento do <i>chatbot</i> sobre os conteúdos trabalhados no curso ou disciplina que ele atende, incluindo conteúdo (conceitos, teorias etc), atividades e agenda do curso. Os conteúdos são válidos e embasados cientificamente?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Conexão com outros conteúdos do curso</b>	Os conteúdos apresentados no <i>chatbot</i> fazem menção e estão diretamente relacionados com outros conteúdos do curso?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
<b>Categoria: Discurso e linguagens</b>		
Clareza	Os textos são apresentados de maneira clara para entendimento dos alunos?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Suficiência de informações</b>	A quantidade de conteúdo é suficiente para dar apoio aos estudos autônomos dos alunos? A extensão/tamanho dos textos é adequada ao que os leitores querem ou estão dispostos a ler?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Categoria: Navegação e Usabilidade</b>		
<b>Opções de menu e de digitação simultâneas</b>		<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Categoria: Aparência e persona</b>		

Cumprimento	Há uma mensagem para cumprimentar o usuário? Essa mensagem transmite algum aspecto da persona do <i>chatbot</i> ?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Persona do <i>chatbot</i>	Se o <i>chatbot</i> se apresenta como um personagem, ele tem nome, profissão, gênero, humor ou outras características demográficas ou comportamentais expressas por meio de sua imagem e discurso? Os alunos gostam dessa persona?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Informações básicas sobre a identidade do <i>chatbot</i>	O <i>chatbot</i> deixa claro qual papel desempenha na situação de interação com o usuário. Por exemplo, é uma colega que vai conversar com o aluno, um atendente administrativo que vai ajudar a resolver questões de secretaria, uma professora que vai dirimir dúvidas pedagógicas etc? Os usuários aprovam esse papel?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Categoria: Funcionalidades adicionais</b>		
Resposta a questões desconhecidas	Existe uma ou mais respostas-padrão para os casos em que usuário faz uma pergunta para a qual o <i>chatbot</i> não tem resposta?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Opção para voltar ao menu inicial	Há algum comando ou opção de onde clicar para voltar ao menu inicial do sistema?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente

Fonte: dados da pesquisa

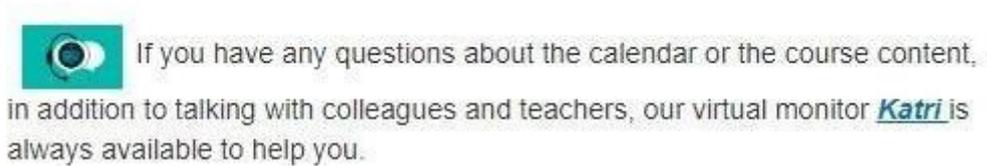
As adaptações e melhorias no sistema foram feitas a tempo do início das aulas da nova turma do curso. Essa nova versão de Katri começou a ser usada no dia 12 de março de 2019, primeiro dia de aula da turma daquele ano do curso *Teaching and Learning in Higher Education*. Os alunos tiveram acesso ao *chatbot* até 27 de setembro de 2019, último dia letivo do curso. Katri poderia ser acessada por meio de um link, enviado por *Whatsapp* e por email aos alunos do curso, ou também por meio de link disponível no AVA, conforme mostra a Figura 22<sup>29</sup>.

<sup>29</sup> Tradução do inglês para o português:

Se você tiver qualquer pergunta sobre o calendário ou o conteúdo do curso, além de falar com os colegas e

### FIGURA 22 - Introdução a Katri no Moodle

\*\*\*\*\*



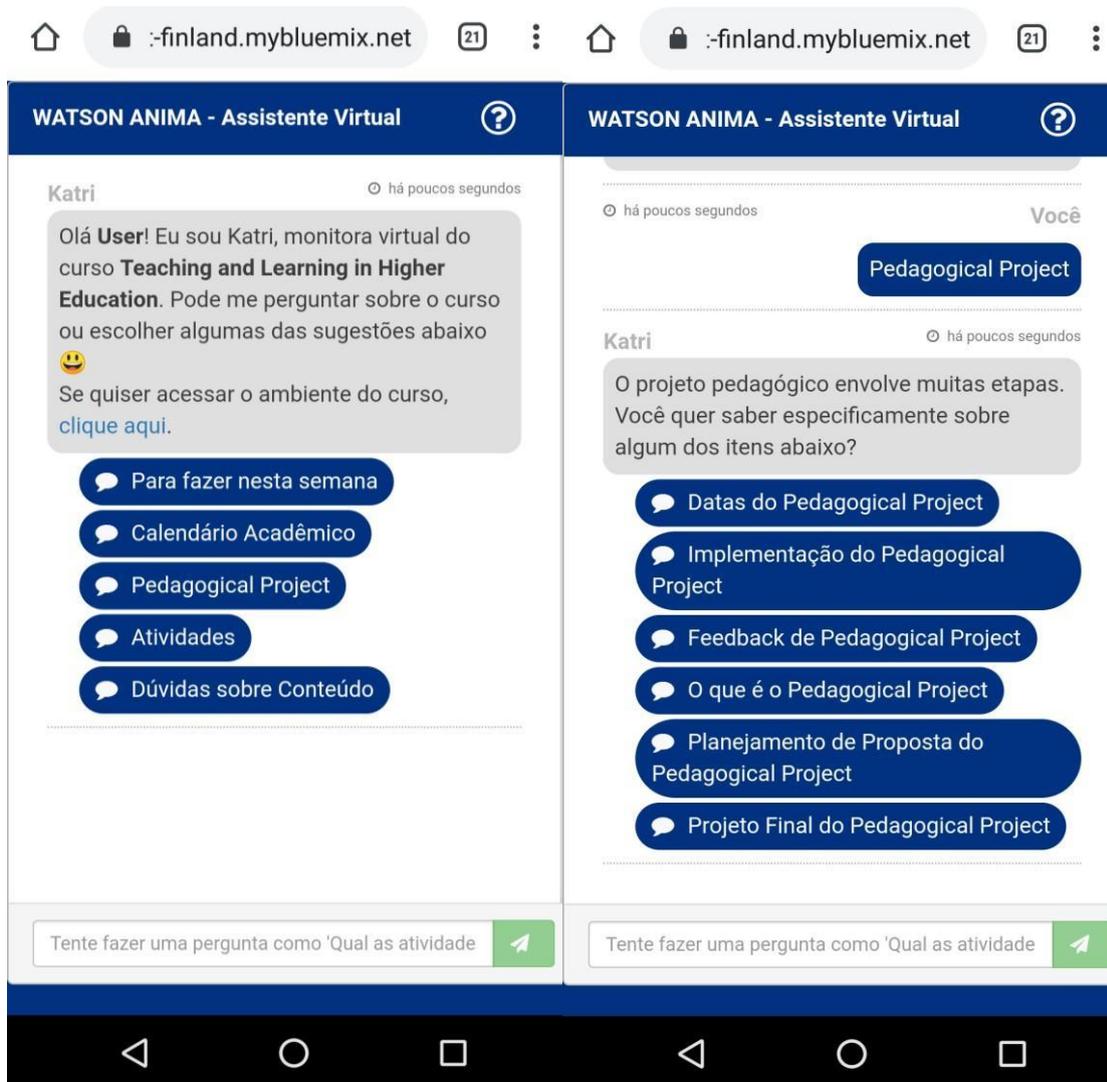
Fonte: Moodle - primeira página do AVA do curso.

Katri foi treinada para conversar sobre todos os conteúdos da ementa do curso, além de responder a questões administrativas, como prazos para entregas de atividades e como acessar o AVA do curso. As Figuras 23 e 24 exemplificam a interação de um aluno com Katri por meio de cliques em botões do menu inicial e submenus.

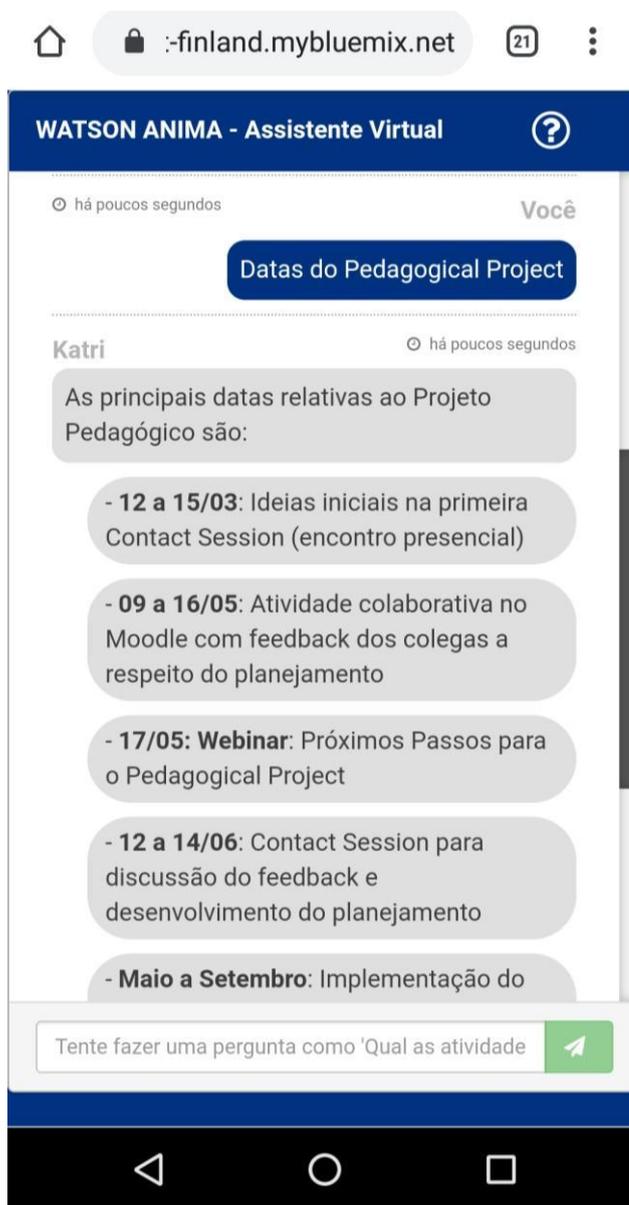
---

professoras, nossa monitora virtual Katri está sempre disponível para lhe ajudar.

**FIGURA 23 - Menu inicial e primeira mensagem sobre *Pedagogical Project***



Fonte: *Chatbot Katri*

FIGURA 24 - Continuidade de diálogo sobre o *Pedagogical Project*

Fonte: *Chatbot Katri*

Assim, chegou-se no que se considera a “versão definitiva de Katri”, ou no que seria a sua versão quase-definitiva, visto que este tipo de sistema está aberto a revisões e alterações durante o tempo que está em uso, sempre que os produtores acharem necessário. Os resultados de interação e avaliação dessa nova versão de Katri são apresentados a seguir.

## 6.2 Resultados: versão definitiva de Katri

O principal objetivo desta seção é **analisar os dados e discutir os resultados da pesquisa** para que se possa validar o uso da **matriz de avaliação de chatbots didáticos**. Primeiramente, serão apresentados dados quantitativos do uso de *chatbot* durante o curso. Em seguida, os descritores da matriz proposta serão avaliados considerando, além dos dados quantitativos sobre a interação com os usuários, as avaliações (dados qualitativos) de alunos sobre a interação com o *chatbot* ao longo do período letivo e as autoavaliações dos produtores do MD.

### 6.2.1 Dados de interação

Ao longo dos 7 meses de duração do curso em 2019, foram realizadas 341 interações entre alunos e Katri. As interações foram agrupadas em temas, de acordo com a proximidade semântica entre elas. Os temas, por sua vez, foram organizados em categorias. O quadro 20 apresenta uma síntese das interações, divididas por mês durante o período de realização do curso.

**QUADRO 22 - Síntese das interações entre usuários e Katri**

Categoria	Tema	Número de mensagens por mês							Total
		Mar	Abr	Mai	Jun	Jul	Ago	Set	
Calendário	Datas de atividades	16	3	24	2	3	5	5	58
	Calendário Acadêmico <sup>30</sup>	3		4	1	1	1	1	11
Conteúdos do curso	Dúvidas sobre conceitos teóricos	11	2	6	6	2	2		29
	Dúvidas sobre outros tipos de conteúdo	9	4	6	5				24
	Mensagens de confirmação de conteúdo encontrado	1						1	2

<sup>30</sup> Calendário completo com todas as atividades e eventos do curso.

Realização de atividades	Dúvidas sobre a atividade Projeto Pedagógico	14	2	10	5	2	4	43	80
	Dúvidas sobre a atividade Estudo de Caso		2	1					3
	Outras Atividades	9	1	15	5	3	1	3	37
Encontros síncronos	Encontros presenciais			2					2
	Webinários <sup>31</sup>	3		2					5
Avaliação das respostas obtidas	<i>Feedbacks</i> positivos para a interação com a Katri	20	4	12	7	1	5	11	60
	<i>Feedbacks</i> negativos para a interação com a Katri	3	1	1				2	7
Questões do sistema	Solicitação de voltar ao menu inicial	2		1					3
	Acesso ao AVA	1		2					3
Outros	Outras questões <sup>32</sup>	5	1	2	6		1	2	17
	<b>Total Geral</b>	<b>97</b>	<b>20</b>	<b>88</b>	<b>37</b>	<b>12</b>	<b>19</b>	<b>68</b>	<b>341</b>

Fonte: Dados da pesquisa

Os dados de interação confirmam alguns pontos que o teste piloto já havia demonstrado: questões relativas às datas e organização do curso e a realização da Atividade Projeto Pedagógico foram as mais demandadas pelos alunos.

<sup>31</sup> Seminários virtuais realizados via plataforma de videoconferência.

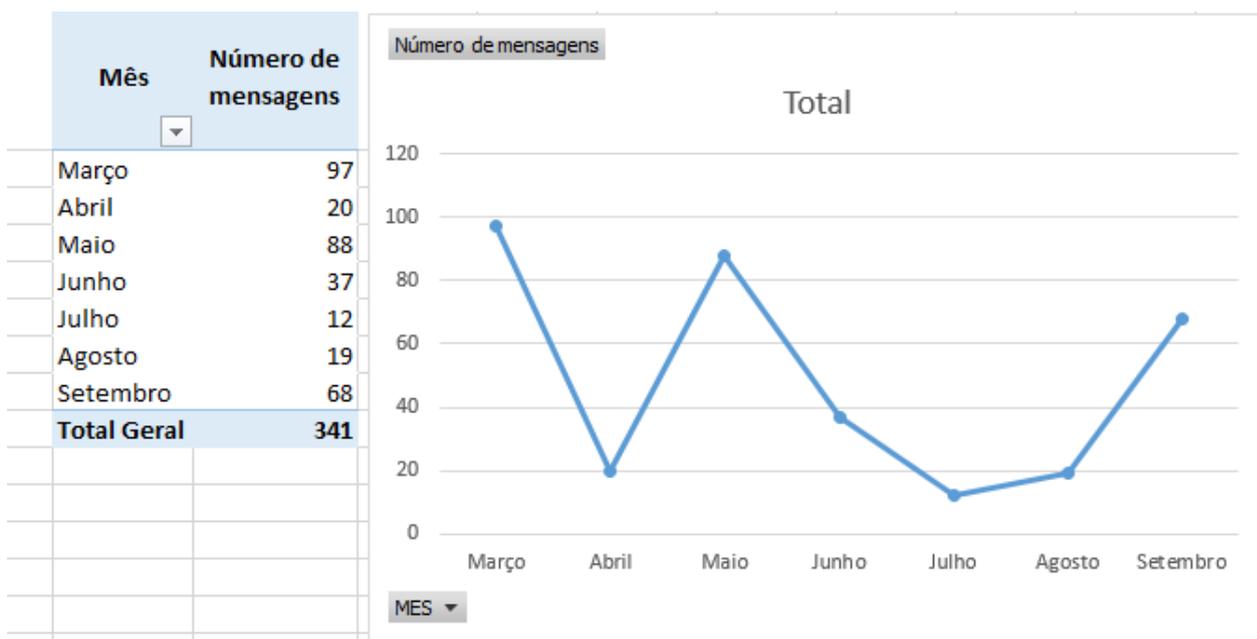
<sup>32</sup> Perguntas que não se repetiram e não tem proximidade semântica com outras mensagens enviadas à Katri foram classificadas como Outras. Exemplo: “Qual é o email da professora?”.

Das 341 interações, em 77 os usuários optaram por digitar perguntas. Nas outras 264 interações, os usuários se guiaram exclusivamente clicando nos botões de menu que Katri apresenta. Ou seja, 77,4% das interações foram realizadas apenas por cliques em opções pré-definidas pelos produtores do *chatbot*, enquanto em 22,6% das interações os usuários preferiram criar suas próprias estratégias textuais para dialogar com Katri.

Ao final das interações, Katri pedia um *feedback* aos alunos, perguntando se a resposta foi útil. Essa avaliação não era obrigatória. No total, 67 interações foram avaliadas pelos usuários, que poderiam optar apenas por duas alternativas. A opção “Sim, você me ajudou” foi marcada 60 vezes, enquanto a opção “Não, você não me ajudou” foi marcada 7 vezes, representando um índice de satisfação de 89,5%.

Ao longo do curso, houve uma grande variação do número de interações dos alunos com Katri, conforme mostra o Gráfico 1.

**GRÁFICO 1 - Variação de interações durante os meses do curso**



Fonte: Dados da pesquisa

Os meses de março, maio e setembro tiveram os picos de uso do *chatbot*. Em entrevista ao pesquisador, os outros produtores disseram que isso pode ser explicado pela dinâmica do próprio curso, pois em março o *chatbot* foi apresentado aos alunos, em maio as primeiras atividades avaliativas foram entregues e em setembro o Projeto Pedagógico, atividade final do



Aluno 15			3					3
Aluno 16	5	5		2			4	16
Aluno 17				5			4	9
Aluno 18		2						2
Aluno 19	4			4				8
Aluno 20	4							4
Aluno 21	2	6						8
Aluno 22			18	2		1	1	22
Aluno 23			4					4
Aluno 24	1		1	1	2			5
Aluno 25			3					3
Aluno 26		4	4					8
Aluno 27	8	3	12				15	38
Não identificado	38		8	4	4	5	15	74
Total	97	20	88	37	12	19	68	341

Fonte: Dados da pesquisa

Houve, também, entre os próprios alunos, uma grande diferença no número de interações que cada um realizou com Katri. Dos 32 alunos da turma, cinco não interagiram nenhuma vez com Katri. Os outros 27 variaram entre uma e quarenta interações cada ao longo do curso. Os alunos tinham a opção de se identificar ou não em cada acesso ao *chatbot*. Do total, 74 interações foram realizadas sem identificação.

### 6.2.2 Avaliação da versão definitiva de Katri: análise e discussão

A versão definitiva de Katri desenvolveu os 68 temas, já listados na fase piloto, sobre o qual o sistema deveria responder. Esses temas foram desdobrados em 120 respostas diferentes. Ou seja, em sua versão definitiva Katri sabe fornecer ao usuário, em média, aproximadamente dois tipos de respostas por tema. Como exemplo, Katri sabe indicar duas maneiras diferentes de acessar o Moodle. Também foram mapeadas as possíveis ambiguidades que as respostas poderiam gerar. O intuito desse mapeamento foi construir um banco de dados que permita à

Katri dar a resposta que os usuários de fato buscam. A estratégia de desambiguação utilizada foi a análise de intenções (INBOT, 2019), com o cadastro de sinônimos para as entidades e intenções.

Como padrão, buscou-se ter ao menos três entidades e sete intenções para cada uma das respostas que Katri saber dar ao usuário. Quando um aluno escreve uma pergunta à Katri, por meio das palavras usadas na pergunta ela consegue entender a intenção e a entidade. Assim, ela busca em seu banco de dados a resposta que esteja simultaneamente relacionada com a entidade e intenção identificada.

No mapeamento de ambiguidades, por meio da conversa com estudantes e professores, verificou-se que ao buscar o link para acesso ao curso, o aluno geralmente deseja receber o link do Moodle, e não de outro tipo de sistema também usado no curso, como o Adobe Connections, que era usado como sistema de videoconferência para aulas ao vivo a distância (ver Figura 25). Por meio do cadastro de entidades e intenções pela equipe de produção de Katri, foi possível trabalhar a desambiguação, direcionando a mensagens consideradas genéricas como “link do curso” e “como eu posso acessar o portal do curso” respostas que levam ao acesso ao Moodle. Já o link do Adobe Connections era fornecido apenas quando havia menção explícita a este sistema.

**FIGURA 25 - Exemplos de conteúdos criados para treinar a Katri**

Tema	Resposta	Entidades	Intenções
Acesso ao Moodle	O link do moodle é: <a href="http://bit.ly/finlandia3">http://bit.ly/finlandia3</a> Você vai precisar inserir o seu login e senha para acessar.	Moodle	como eu posso acessar
		Portal do curso	qual é o link
		link do moodle	qual link
		pagina web do curso	como entrar
		Link	como acesso ao
		Link do curso	como posso entrar
			quero entrar
	não encontro o		
	onde posso acessar		
	onde posso entrar		
	tenho problemas para		
Tema	Resposta	Entidades	Intenções
Acesso ao Adobe Connections	O Adobe Connect é um software de conferência via Web que estamos usando como um ambiente de webinar neste curso. Vamos configurar salas para os webinars e designar uma sala de reunião individual para cada grupo de estudo. Além disso, cada aluno pode configurar e gerenciar as salas.	adobe connections	como eu posso acessar
		adobe	qual é o link
		adobe connect	qual link
		AC	como entrar
		AC meeting rooms	como acesso ao
		adobe meeting rooms	como posso entrar
meeting rooms	quero entrar		

Fonte: dados da pesquisa

Essa estratégia de desambiguação foi essencial para que os alunos conseguissem ter um fluxo

satisfatório de interação com Katri. Em seu banco de dados há 120 respostas e Katri foi treinada para entender 840 maneiras diferentes que os usuários poderiam fazer perguntas sobre os conteúdos de seu banco de dados. A inteligência artificial do sistema consegue combinar palavras existentes nessas diferentes maneiras de perguntar e também consegue identificar padrões textuais que os alunos usam na interação (mesmo as não mapeadas pela equipe), então outras centenas de combinações de maneiras de se fazer perguntas podiam ser entendidas por Katri, segundo os produtores.

Após o uso da versão definitiva de Katri em contexto real pelos alunos do curso, buscou-se entender como eles avaliam o *chatbot*. Para coletar as impressões dos alunos da turma de 2019 (turma 2), foram aplicados questionários dividindo os 27 respondentes em 4 perfis: 1) alunos que interagiram com o *chatbot* em diferentes momentos ao longo do curso (do aluno 1 ao 11); 2) alunos que interagiram com o *chatbot*, mas que não voltaram a ter interações ao longo do curso (do aluno 12 ao 16); 3) alunos que só interagiram com o *chatbot* na etapa final do curso (do aluno 17 ao 19); 4) alunos que não interagiram com o *chatbot* (do aluno 20 ao 27). Cada aluno recebeu um questionário adaptado às especificidades do perfil em que foi agrupado, sendo que os perfis 1, 2 e 3 compartilham muitas questões (ver Apêndice 3). São esses 19 respondentes dos três primeiros grupos que de fato avaliaram o *chatbot*. Os outros oito respondentes contribuíram para que os produtores pudessem entender os motivos do não engajamento para uso desse MD, o que não faz parte do objetivo da pesquisa, por isso as respostas desse grupo não serão analisadas. Entrevistas com os produtores também foram realizadas. Assim, avaliou-se o *chatbot* de acordo com as avaliações dos alunos-usuários e a autoavaliação dos produtores do MD, seguindo as orientações de Cabero (2001).

Reiterando uma explicação do capítulo de Metodologia, alguns descritores não necessitam de avaliação qualitativa dos usuários e produtores, por tratar-se de uma avaliação objetiva: basta verificar se a funcionalidade existe ou não no *chatbot*. Para todos os outros descritores, em que a avaliação qualitativa é necessária, foram considerados relevantes para os dados desta pesquisa excertos das respostas dos alunos aos questionários e dos produtores durante as entrevistas, desde que tivessem relação com os critérios para avaliação de *chatbots* propostos nesta tese. Considerações realizadas, mas que não se relacionam com os objetivos desta pesquisa, não foram consideradas parte do corpus para análise.

São apresentados a seguir para avaliação os 30 descritores da matriz proposta (ver página 59).

Todas as 5 categorias propostas foram avaliadas. São apresentadas individualmente as avaliações dos descritores da matriz, assim como a análise do pesquisador sobre os resultados da avaliação. Em seguida, o resultado da avaliação deste conjunto de descritores será apresentado em um só quadro para melhor visualização do processo avaliativo e da validação da matriz. Conforme indicado na página 62, preencheu-se a coluna de Avaliação com os indicadores Adequado ou Implementado, no caso de avaliação satisfatória, Parcialmente Adequado ou Parcialmente Implementado, no caso de avaliação considerada intermediária; e Inadequado ou Ausente, no caso de avaliação considerada insatisfatória.

### **Categoria 1: Conteúdos**

A categoria Conteúdos conta com 5 descritores. Dois deles foram analisados qualitativamente e três tiveram avaliação objetiva: verificou-se se os descritores estavam ou não implementados. Segue abaixo o quadro de avaliação dessa categoria, já com a marcação dos resultados. Em seguida ao quadro, apresenta-se como o pesquisador chegou a esses resultados de avaliação.

**QUADRO 24 - Avaliação da categoria Conteúdos**

<b>Categoria 1: Conteúdos</b>		
<b>Descritor</b>	<b>Questões norteadoras</b>	<b>Avaliação</b>
Conteúdo de apresentação	O <i>chatbot</i> é capaz de se apresentar, dizendo quem é e para que serve?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Conhecimentos especializados	O quão completo é o conhecimento do <i>chatbot</i> sobre os conteúdos trabalhados no curso ou disciplina que ele atende, incluindo conteúdo (conceitos, teorias etc), atividades e agenda do curso. Os conteúdos são válidos e embasados cientificamente?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Atualização dos conteúdos	Os conteúdos são atualizados em relação à versão do curso ofertado, caso o <i>chatbot</i> esteja sendo ofertado em períodos letivos consecutivos? Por exemplo, o calendário do curso foi atualizado para a nova turma?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Antecipação às dúvidas mais comuns dos estudantes	As informações ofertadas pelo <i>chatbot</i> via opções de menu são capazes de antecipar e responder às possíveis dúvidas mais comuns dos estudantes?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Conexão com outros conteúdos do curso	Os conteúdos apresentados no <i>chatbot</i> fazem menção e estão diretamente relacionados com outros conteúdos do curso?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Para avaliar o **descriptor Conteúdo de Apresentação**, verificou-se que Katri, no primeiro contato com um aluno-usuário, envia uma mensagem (ver Figura 23, página 111) na qual se apresenta como monitora do curso e explica que pode ajudar o aluno respondendo a dúvidas sobre o curso, além de apresentar um menu inicial com dúvidas frequentes. Considerou-se, objetivamente, que o descriptor está **Implementado**. Pondera-se que não foi avaliada a qualidade e os efeitos do Conteúdo de Apresentação, por exemplo, no engajamento dos usuários.

Já o descriptor Conteúdos Especializados foi avaliado qualitativamente. Por meio dos questionários respondidos pelos 19 alunos que interagiram com Katri, especialmente nas perguntas em que responderam sobre qual tipo de conteúdo foi pesquisado com mais frequência e se os conteúdos contribuíram para a aprendizagem, foi possível encontrar excertos de avaliações de 13 alunos sobre o descriptor.

Nove alunos avaliaram como positiva a experiência e os conteúdos apresentados. Dentre esses 9 participantes, 2 trouxeram aspectos da Apreciação e valoraram positivamente o impacto de Katri ao afirmar que a ferramenta era uma maneira “prática” e “importante” para estudar o conteúdo do curso. Apenas uma respondente (aluna 18) expressou Afeto, indicando explicitamente que os conteúdos trazidos por Katri a fizeram sentir-se “segura” (Segurança). Houve ainda 4 respondentes que valoraram negativamente a experiência, sendo que três disseram que os conteúdos de Katri não contribuíram para a aprendizagem. O aluno 1 expressou Valoração negativa do impacto do conteúdo ao afirmar que “não eram novidades”. Ainda entre os respondentes, 6 alunos não souberam ou não quiseram opinar.

Os três produtores entrevistados foram unânimes na autoavaliação positiva sobre os conteúdos especializados, conforme mostra o quadro 25. O produtor-antropólogo, por exemplo, expressou Apreciação positiva da Composição de Katri, ao avaliar que o conteúdo é “rico”, com aumento

da Força por meio dos advérbios de intensidade “bem” e “muito”. Segundo ele, “o conteúdo em si do nosso bot é muito rico. E é bem especializado, relacionado com os trabalhos em grupos. É um conteúdo bem rico que a gente tem lá dentro.”

**QUADRO 25 - Avaliação do descritor Conteúdos Especializados**

<b>Categoria 1: Conteúdos</b>						
<b>Descritor: Conteúdos especializados</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	1	Negativo		Valoração		
	2,3,5,7,9, 11	Positivo				
	8	Positivo			+	
	10	Positivo				-
	14, 15, 17	Negativo				-
	16	Negativo				
	18	Positivo	Segurança			
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Antropólogo	Positivo		Valoração	+	
	Linguista Desenvolvedor	Positivo				

**Fonte:** Dados da pesquisa

A partir da avaliação majoritariamente positiva pelos alunos e unânime pelos produtores, **avaliou-se o descritor Conteúdos Especializados como Adequado**. Apesar da avaliação positiva, observa-se que um elemento ainda precisa de atenção dos produtores. Ao dizer que os conteúdos “não eram novidades”, o aluno 1 traz um importante questionamento: deve um *chatbot* didático conversar com os alunos exclusivamente sobre conteúdos que já estão disponíveis por meio de outros MD e sistemas usados no curso, como no caso de Katri, ou é papel dele também trazer conteúdos complementares e adicionais? Como essa função de trazer novos conteúdos não fazia parte do propósito de implementação de Katri no curso, essa ponderação não influenciou negativamente a avaliação, visto que o MD cumpre bem a função para o qual foi projetado, mas considera-se importante trazer o questionamento para reflexão em novas produções.

O **descritor Atualização dos Conteúdos foi considerado Implementado**, pois verificou-se que todos os conteúdos estavam atualizados para oferta de 2019 do curso TLHE. Todas as mudanças realizadas pelas professoras finlandesas no curso, entre a oferta de 2018 (piloto) e a oferta de 2019 (versão definitiva do *chatbot*), como inclusão de novos textos, assim como as

datas de encontros presenciais, *webinars*, entre outras datas relevantes no calendário do curso, foram atualizados para que Katri soubesse responder especificamente às questões dos alunos da turma vigente.

Para avaliação do **descriptor Antecipação às Dúvidas Mais Comuns dos Estudantes**, analisou-se as respostas dos produtores às entrevistas e dos estudantes aos questionários. Onze alunos dos grupos de estudantes que usaram o *chatbots* ao longo do curso expressaram avaliações sobre esse descriptor, principalmente nas respostas à pergunta “Há algum tipo de informação ou conteúdo do curso que você gostaria que estivesse no menu ou que o *chatbot* soubesse lhe responder e que você não conseguiu encontrar?”. Foram dez avaliações positivas e uma negativa, sendo nove delas valorações positivas. Alguns alunos usaram adjetivos como “adequadas” e “interessantes” para avaliar como os itens do menu inicial de Katri antecipam e respondem às suas dúvidas. O aluno 5 afirmou que os itens que estão no menu “são os mais relevantes”, expressando sua valoração positiva por meio do adjetivo “relevantes”, com aumento da Força com o uso do advérbio “mais”. O aluno 18 afirmou: “pelo que lembro, o tema que eu buscava constava no menu inicial, o que facilitou minha navegação”. A avaliação positiva tem o Foco atenuado pela modalização do discurso com o uso da expressão “pelo que lembro”, que demonstra incerteza. A única avaliação negativa foi uma Apreciação da Composição de Katri pela aluna 2, que afirmou que os temas do menu eram “protocolares” e que o menu poderia ter “maior complexidade”.

Os produtores também avaliaram positivamente o descriptor. O produtor-desenvolvedor expressou a sua Satisfação com o fato de um dos itens do menu (Para Fazer Esta Semana) ter sido usado com muita frequência pelos alunos. “Acho que as atividades da semana, ou próximas atividades, foi uma surpresa nas interações” (PRODUTOR-DESENVOLVEDOR). A produtora-linguista demonstrou sua Apreciação positiva da Composição do *chatbot*, se referindo a esse mesmo item de menu ao usar a expressão popular “acertar na mosca”: “eu acho que essa a gente acertou na mosca que essa seria de longe a maior interação, e foi mesmo” (PRODUTORA-LINGUISTA).

QUADRO 26 - Avaliação do descritor Clareza

Categoria 1: Conteúdos						
Descritor: Antecipação às dúvidas mais comuns dos estudantes						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	3, 4, 10, 17	Positiva		Valoração		
	5, 7, 11	Positiva			+	
	6, 18	Positiva				-
	19	Positiva		Composição		
	2	Negativa				
Avaliação por produtores	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	Desenvolvedor	Positivo	Satisfação			
	Antropólogo	Positivo		Composição	+	
	Linguista	Positivo				+

Fonte: Dados da pesquisa

O descritor foi considerado **Adequado** nesta avaliação. Além de se considerar a avaliação majoritariamente positiva do descritor pelos alunos e pelos produtores, também se levou em consideração na análise a verificação, via dados do sistema, de que os alunos preferiram clicar nos itens do menu (77,4% das interações) do que digitar as suas próprias perguntas (22,6% das interações). Como a maior parte dos alunos que respondeu ao *feedback* solicitado por Katri disse que as respostas foram úteis (89,5%), entende-se que eles não apenas optaram por clicar nas opções de menu, como obtiveram as respostas que procuravam. Infere-se, assim, que de fato os itens do menu ajudavam a antecipar e/ou responder às principais dúvidas dos estudantes.

Já o descritor **Conexão com Outros Conteúdos do Curso** foi avaliado como **Parcialmente Adequado**, pois verificou-se que na versão definitiva há links para acesso ao Ambiente Virtual de Aprendizagem do curso, inclusive direcionamento a links diretos para acesso a vídeos e textos disponibilizados pelas professoras. Entretanto, nem todos os materiais didáticos do curso são mencionados e conectados ao *chatbot* e não há conexão de Katri com as atividades avaliativas e fóruns de discussão do curso.

Apesar da avaliação como Parcialmente Implementado, considera-se que a ausência de conexões com todos os conteúdos do curso não é necessariamente um problema de um *chatbot* didático. É necessário se perguntar a que o *chatbot* se propõe, ou seja, qual é o objetivo desse material didático no curso. A ausência de conexões com outros materiais didáticos que

poderiam contribuir para o alcance dos objetivos pedagógicos do *chatbot* seria considerada um problema, mas a ausência de outras conexões, não ligadas aos objetivos pedagógicos do *chatbot*, poderia ser simplesmente considerada como parte das limitações de escopo de um MD, característica presente em livros didáticos (pela limitação do número de páginas), em videoaulas (pela limitação do tempo do vídeo) ou qualquer outro MD. Aqui, reforça-se o posicionamento desta pesquisa em olhar para os *chatbots didáticos* considerando suas semelhanças com outros tipos de MD (sem esquecer de suas especificidades), o que se sugere que seja feito também como novos tipos de MD que estão sendo criados e que ainda serão criados com o uso de inteligência artificial.

No caso específico de Katri, há algumas melhorias a se fazer que não se justificam pela limitação de escopo. Há, por exemplo, textos por ela mencionados que não possuem links para acesso direto no AVA, enquanto outros textos possuem links para acesso. Há, então, conexões a se adicionar, relacionadas diretamente aos temas abordados.

A categoria Conteúdos teve em sua avaliação dois descritores considerados Adequados, dois Implementados e um Parcialmente Adequado. Apesar de ter apontamentos para melhorias, que merecem atenção, verifica-se que o trabalho de desenvolvimento de conteúdos, de maneira geral, foi considerado satisfatório.

### **Categoria 2: Discurso e Linguagens**

A categoria Discurso e Linguagens tem 6 descritores, sendo 4 analisados qualitativamente, com base nas avaliações de alunos e produtores do *chatbot* didático, e dois avaliados objetivamente, com a verificação da implementação ou ausência do descritor. O quadro 27 apresenta os resultados da avaliação dessa categoria.

**QUADRO 27 - Avaliação da categoria Discurso e Linguagens**

<b>Categoria 2: Discurso e linguagens</b>		
Modalidades oral e escrita de uso da língua	O <i>chatbot</i> aceita <i>inputs</i> por texto oral (áudio) e escrito e dá respostas também com texto oral e escrito?	( ) Implementado (X) Parcialmente Implementado ( ) Ausente

Clareza	Os textos são apresentados de maneira clara e coerente para entendimento dos alunos?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Lembrança do nome do usuário	O <i>chatbot</i> consegue identificar o nome do usuário e se referir a ele pelo nome durante a conversa?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>.Suficiência de informações</b>	A quantidade de conteúdo é suficiente para dar apoio aos estudos autônomos dos alunos? A extensão/tamanho dos textos é adequada ao que os leitores querem ou estão dispostos a ler?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Adequação do vocabulário ao curso.</b>	O vocabulário usado pelo <i>chatbot</i> está coerente e em sintonia com o vocabulário usado nos outros materiais didáticos do curso? Os diálogos são realizados em linguagem natural com registros próximos aos usados pelos usuários/alunos no curso?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Formalidade	O grau de formalidade está adequado ao contexto, ao público-alvo e à persona do <i>chatbot</i> ?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado

Fonte: Dados da pesquisa

O descritor Modalidades Oral e Escrita de Uso da Língua foi avaliado como Parcialmente Adequado, pois Katri aceita *inputs* exclusivamente por texto escrito e dá respostas exclusivamente também por texto escrito. Por questões orçamentárias, não foram implementados recursos para que Katri recebesse dos alunos *inputs* por áudio ou vídeo ou respondesse os alunos também por áudio ou vídeo. Essas modalidades de uso da língua, entretanto, são tecnologicamente viáveis e já há *chatbots* que as utilizam. Há, por exemplo, a possibilidade de um usuário fazer uma pergunta por voz em seu aparelho celular e receber uma resposta em áudio para ouvi-la.

Entende-se que a ausência de modalidade oral em Katri é uma questão técnico-orçamentária, que prejudica a interação, visto que não é fornecido ao usuário a possibilidade de escolha da maneira como prefere interagir. As questões orçamentárias, evidentemente, não podem ser

desconsideradas no processo de produção de um MD, por isso recomenda-se que, em caso de disponibilidade de recursos em possível futura versão de Katri, o uso da modalidade oral seja implementado.

Já em relação ao descritor Clareza, 16 alunos expressaram opiniões positivas sobre Katri ao avaliar positivamente a Composição da Complexidade ao dizerem, por exemplo, que foi “fácil”, “muito fácil” e “bem fácil” compreender os textos de Katri e que a forma como Katri escreve era “tranquila” de se entender. A maior parte dos alunos, utilizou estratégias linguísticas que levaram à Gradação do valor da Apreciação. Dez alunos usaram estratégias como o uso de advérbios de intensidade, que ampliaram o Foco ou aumentaram a Força da Apreciação. Dois deles atenuaram o foco da Apreciação da Composição, como a Aluna 10, que afirmou que Katri “tirou dúvidas pontuais, de forma prática”. A aluna escolhe suavizar o tom de sua Apreciação, ao dizer a ação de Katri foi realizada apenas em “dúvidas pontuais”. Três alunos não souberam ou não quiseram opinar e não houve nenhuma avaliação negativa sobre o **descritor Clareza**, que **foi considerado Adequado**.

Os produtores antropólogo e linguista também expressaram Apreciações positivas sobre esse descritor, ao dizerem que Katri tem uma linguagem de “fácil entendimento”. O produtor-antropólogo respondeu “acho que conseguimos”, ao ser perguntado se a construção dos textos resultou em diálogos claros no *chatbot*. Apesar de positiva a autoavaliação, o produtor-antropólogo atenua o Foco da Apreciação ao demonstrar sua incerteza quando diz que “acha” que a linguagem estava clara.

**QUADRO 28 - Avaliação do descritor Clareza**

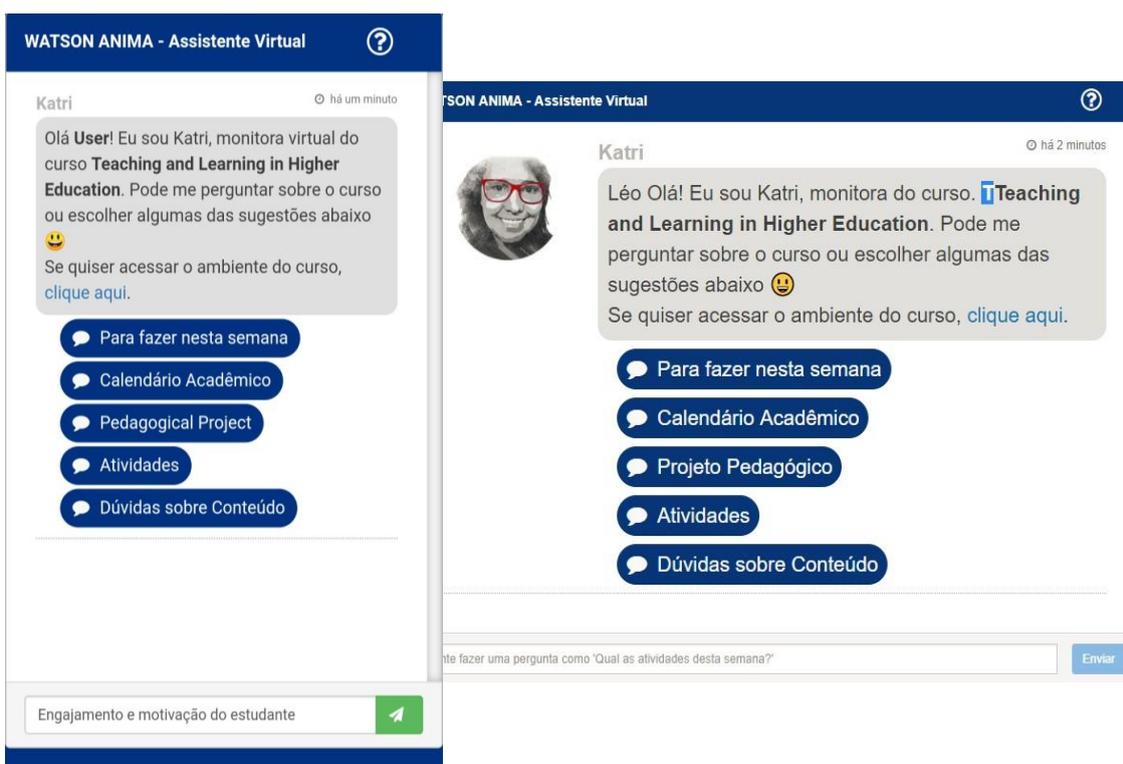
<b>Categoria 2: Discurso e linguagens</b>						
<b>Descritor: Clareza</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	1	Positivo		Composição	+	+
	2, 3, 4, 7, 10, 16, 17, 18, 19	Positivo			+	
	5, 6, 8, 12, 15	Positivo				
	9, 10	Positivo				-
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Antropólogo			Composição		-
	Linguista					

Fonte: Dados da pesquisa

Sobre o **descritor Lembrança do Nome do Usuário**, verifica-se que assim que o usuário

acessa o sistema, Katri pergunta o nome do aluno. A partir de então, em várias mensagens utiliza o nome do usuário durante os diálogos, conforme mostra a imagem X. **Avalia-se, então, este descritor como Implementado**, pois Katri é capaz de recuperar o *input* fornecido pelo usuário na pergunta sobre o seu nome e usá-lo durante a interação. Esse recurso é usado para conferir um tom de personalidade ao discurso de Katri e para demonstrar ao aluno que ele está participando de uma experiência individualizada de aprendizagem, o que está também diretamente relacionado ao descritor Adaptação/individualização do processo de interação e aprendizagem, da categoria Navegação e Usabilidade.

**FIGURA 26 - Katri chama o usuário pelo nome “Léo”, fornecido por ele na primeira interação**



Fonte: *Chatbot Katri*

Houve grande variação nos tipos de Apreciação expressos pelos alunos em relação ao **descritor Suficiência de Informações**. Verificou-se que há expressões positivas de Reação, Valoração e sobre a Composição de Katri, além de avaliações negativas da Composição. Houve, ainda, expressão de Afeto/Felicidade por uma aluna e uma produtora. Foi observada considerável diferença entre as avaliações dos grupos de alunos, de acordo com a frequência de interação com Katri.

Dentre os 11 alunos que usaram o *chatbot* durante todo o curso, 8 fizeram comentários positivos sobre a suficiência da quantidade de informações que Katri provia para dar apoio aos estudos, enquanto 2 acreditaram ser uma quantidade “mediana” e um aluno não soube responder. Dos 5 alunos que interagiram com Katri apenas no início do curso, 3 disseram ter tido a impressão que o conteúdo era insuficiente, 1 afirmou que era suficiente e 1 não soube opinar. Todos os 3 respondentes que usaram o *chatbot* apenas no final do curso tiveram opiniões positivas sobre este descritor. Ou seja, dos 19 respondentes que usaram o *chatbot*, 12 acreditam avaliam positivamente o item, 3 negativamente, 2 de maneira parcialmente positiva e 2 não sabem como avaliar.

Os alunos que interagiram com Katri apenas nos primeiros dias de aula majoritariamente avaliaram negativamente este descritor, enquanto os alunos que interagiram mais vezes com Katri ao longo do curso em sua maioria expressaram opiniões positivas. Isso pode significar que os alunos que interagiram apenas no início do curso não quiseram acessar mais o *chatbot* mais porque acharam que havia pouca informação relevante ou que eles sequer conhecem o conteúdo do *chatbot* porque não interagiram o suficiente ao longo do curso. Essas são inferências não confirmadas pela coleta de dados e que podem suscitar novas questões a se fazer para futuras turmas que usem o *chatbot* didático.

A aluna 18 afirmou que a quantidade de informações das respostas de Katri era “excelente”, valorando positivamente o descritor, e completou a sua avaliação afirmando que o conjunto de informações era “simples e consistente”, o que expressa avaliação positiva da Composição. A aluna 8 disse que o conteúdo provido por Katri é “bem sucinto e direto” (avaliação positiva da Composição) e expressou Felicidade (Afeto) ao finalizar a avaliação deste item escrevendo “eu gostei”. A aluna 2 declarou que achou a quantidade de informações “mediana”, mas modalizou o discurso dizendo: “entendo que ela estava aprendendo”. Essa avaliação da aluna 2 foi considerada de polaridade neutra, por não haver indícios de que essa opinião foi positiva ou negativa.

Já os produtores acreditam que há informação suficiente e que os textos, na verdade, poderiam ainda ser mais sucintos, mantendo quantidade de informação significativa, mas diminuindo a extensão das redações. A produtora-linguista acredita que há textos ainda muito longos nas respostas de Katri, e afirmou:

Eu gosto bastante. Alguns textos, na última vez que olhei, eu falei assim ‘nossa, aparecem 4 balõezinhos (texto dividido em quatro partes). A gente que é de Letras adora (textos longos), mas nem todo mundo gosta, então... Mas isso é autocrítica mesmo, é achar que nunca está bom, sempre tem coisa para melhorar. (PRODUTORA-LINGUISTA)

**QUADRO 29 - Avaliação do descritor Suficiência de informações**

<b>Categoria 2: Discurso e linguagens</b>						
<b>Descritor: Suficiência de informações</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	5	Positiva		Reação		
	3	Positiva		Reação	+	
	8, 17, 18, 19	Positiva		Composição		
	2, 4	Neutra		Composição		
	8	Positiva	Felicidade			
	9	Positiva		Valoração		
	10, 18	Positiva		Valoração		-
	6, 11, 12	Positiva		Reação		
	14, 15	Negativa		Composição		-
	16	Negativa		Composição		
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Linguista	Positiva	Felicidade		+	
	Desenvolvedor Antropólogo	Positiva		Composição		-
	Linguista	Negativa		Composição		-

Fonte: dados da pesquisa

Considerando a autoavaliação positiva dos produtores, mas com ressalvas à extensão dos textos, e diante de um número minoritário, porém significativo de alunos que consideram a quantidade de informações mediana ou insuficiente, este descritor foi avaliado como Parcialmente Adequado. Ao contrário da inferência dos produtores de que o conteúdo estava muito extenso, alguns alunos afirmam que o conteúdo era insuficiente, o que pode significar que gostariam de um conteúdo ainda mais extenso que o apresentado por Katri. Assim, sugere-se que em uma nova rodada de revisão do *chatbot*, os produtores avaliem possíveis necessidades de ajustes na quantidade e extensão dos textos. Uma nova rodada de pesquisa dos produtores com os alunos, por meio de grupo focal ou de novo *Think Aloud Test*, poderia auxiliar nas decisões sobre novas adaptações nos textos fornecidos por Katri aos alunos.

O descritor **Adequação ao Vocabulário do Curso** foi avaliado positivamente pelos produtores do curso. O produtor-antropólogo afirmou que foi um desafio manter a sintonia com o

vocabulário do curso ao mesmo tempo em que deveria ser respeitada à persona de aluna-monitora criada para Katri. “Ela tinha que ter conhecimento do curso, mas não era professora, tinha que ser mais próxima dos alunos. Esse foi um desafio no estilo dos textos, que em certa medida eu acho que a gente conseguiu”. Ao dizer que “conseguiu”, ele expressa sua autoavaliação positiva da Composição de Katri, mas atenua o Foco da apreciação ao modalizar o discurso, demonstrando incerteza ao dizer que “em certa medida”, ele “acha” que conseguiu.

A produtora-linguista também avaliou positivamente Katri, afirmando que a decisão de manter no *chatbot* os títulos de conteúdos e alguns termos redigidos exatamente como eram usados nos outros MD do curso e no Moodle foi “uma decisão necessária” porque se fosse diferente “isso poderia gerar confusão para o aluno saber se era a mesma coisa ou não”. Indiretamente, o discurso da produtora-linguista demonstra que ela acredita que houve consistência no vocabulário usado por Katri, por isso considera-se que ela apreciou positivamente a Composição de Katri.

Por outro lado, ela questiona a consistência do vocabulário de Katri devido a um fator específico: o uso de termos em inglês, em meio a textos escritos em português.

Eu acho que foi um passo necessário, mesmo que não tenha ficado muito consistente essa parte do menu em inglês e depois as respostas em português. Eu acho que as respostas têm que ser em português mesmo, e eu acho que manter partes do menu em inglês, usando exatamente os termos como são disponibilizados nos títulos dos materiais no Moodle, vamos dizer, se foi um mal, foi um mal necessário. Porque eles (os alunos) têm a proficiência para fazer as aulas em inglês, mas sempre aparecem dúvidas sobre os termos em inglês. (PRODUTORA-LINGUISTA)

Apesar da ponderação sobre a consistência, inclusive utilizando o termo “mal necessário” (que atenua o foco da apreciação), entende-se que a produtora-linguista valorou positivamente a maneira como os diálogos são realizados, em linguagem próxima ao que é verificado no uso cotidiano dos alunos no curso e nos outros materiais didáticos. Infere-se que em sua autoavaliação, ela acredita que “valeu a pena” usar termos em inglês no *chatbot*.

O produtor-antropólogo concordou com a autoavaliação de sua colega:

É, realmente eu acho que fazia parte da lógica do bot (*chatbot*). Foi uma decisão que gerou mais trabalho para gente no treinamento porque o bot tem que entender as palavras em duas línguas, mas termina sendo importante porque as pessoas no curso usam essas referências em inglês. Tem algumas palavras do curso que para mim tinham que estar em inglês no bot. A gente não tinha como fugir disso, senão provavelmente a capacidade de resposta Katri iria ser menor. (PRODUTOR-

ANTROPÓLOGO)

QUADRO 30 - Avaliação do descritor Adequação do vocabulário

Categoria 2: Discurso e linguagens						
Descritor: Adequação do vocabulário ao curso.						
Avaliação por produtores	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	Linguista	Positiva		Valoração		-
	Antropólogo	Positiva		Composição		-
	Linguista	Positiva		Composição		

Fonte: Dados da pesquisa

Nas respostas dos alunos aos questionários, não havia excertos que expressavam avaliações diretas sobre esse descritor, porém considerou-se na análise o fato de, em outras questões, os alunos avaliarem positivamente o discurso de Katri, como ao avaliar a Clareza, e também afirmarem, majoritariamente, que conseguiram encontrar os conteúdos do curso nas respostas fornecidas. Isso permite inferir que há conexão e coerência do vocabulário de Katri com o vocabulário do curso e também que a maneira como Katri escreve seus textos tem registro próximo aos usados pelos alunos no curso, visto que ela costuma reconhecer as perguntas que os alunos escrevem e eles dão feedbacks positivos sobre as respostas por ela fornecidas. São inferências que, em novas rodadas de avaliação de possíveis futuras versões de Katri, merecem investigação.

**O descritor, foi, então, considerado Adequado**, pois percebe-se que há **Adequação ao Vocabulário ao Curso** e à maneira como os alunos falam e escrevem no curso. Entretanto, mesmo com a avaliação positiva, sugere-se que os produtores revisem a maneira como usam os termos em inglês nos textos do *chatbot*. Quando a produtora-linguista afirma que há um problema de consistência, e o uso das palavras e expressões em língua inglesa foi “um mal necessário”, isso indica que uma revisão para melhoria da consistência/composição precisa ser realizada pelos produtores caso exista uma nova versão do *chatbot*.

Dos 19 respondentes que interagiram com o *chatbot* didático, 15 apreciaram positivamente a **Formalidade** do discurso de Katri, último **descritor da Categoria Discurso e Linguagens**. Dez alunos utilizaram em seus discursos avaliativos adjetivos como “adequado”, “acessível”,

“efetiva”, “apropriada”, “simples”, e “tranquila”, que demonstram a avaliação positiva da Composição de Katri. A aluna 18 disse que Katri é “bem simples, clara, efetiva, direta ao ponto”. Essa sequência de adjetivos em uma frase expressa a gradação da avaliação da Composição do *chatbot*, por meio da ampliação do Foco dessa avaliação positiva da formalidade do discurso de Katri.

Cinco alunos reagiram positivamente, afirmando que a formalidade de Katri é “boa” ou “muito boa”. O aluno 15 disse que o *chatbot* é “Formal, com extensão similar ao de páginas oficiais”. Esta opinião foi considerada neutra, pois não há dados para inferir se o aluno gostou ou não da interação. Três alunos não souberam opinar sobre o grau de formalidade do discurso usado nas interações.

**QUADRO 31 - Avaliação do descritor Formalidade**

<b>Categoria 2: Discurso e linguagens</b>						
<b>Descritor: Formalidade</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	1, 2, 6, 7, 8, 10, 16, 17, 19			Composição		
	4, 11, 12			Reação		
	3, 5			Reação	+	
	15	Neutra		Composição		
	18	Positiva		Composição		+
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Linguista Desenvolvedor	Positiva		Composição		-

Fonte: Dados da pesquisa

A produtora-linguista afirmou que há uma diferença no grau de formalidade de Katri em diferentes momentos. Segundo a produtora, Katri é menos formal como quando ela está cumprimentando o usuário, introduzindo um tema, tentando entender o que ele quer, “situações em que ela até usa emoticons”. E há situações em que ela trata os conteúdos do curso com mais profundidade “porque aí ela é mais formal”. A produtora afirma que, para a persona definida para Katri, o grau de formalidade é adequado, inclusive com essas variações.

Porque ao iniciar essas conversas, perguntando o que os alunos gostariam de saber, como ela pode ajudá-los, são frases para estabelecer a conexão, então estar mais informal nessa parte é mais importante para que não haja distância entre a Katri e os alunos. Mas quando chega a parte da Katri dar conceitos, conteúdos completos, os textos não são tão densos, mas não mais formais, o que acho que é bom porque é assim que faço com meus alunos normalmente em sala de aula. Enquanto (sic) professora, quando estou naquela primeira parte da aula, conversando com os alunos, o canal está

mais aberto convidando os alunos para interagir e até fazendo piadas. Quando eu entro nas explicações, é natural que o tom de voz mude, as construções mudem. Quando lidamos com conceitos teóricos em sala de aula a gente muda o tom até para mostrar a mudança da importância do conteúdo. Essa diferença na formalidade acontece não só com a Katri, mas na vida real. Isso foi planejado e eu acredito que executamos bem. (PRODUTORA-LINGUISTA)

Ao dizer que “executamos bem”, a produtora-linguista aprecia positivamente a Composição de Katri, mas atenua o foco por não demonstrar certeza ao dizer que isso é o que ela “acha”. Diante das avaliações quase em sua totalidade positivas dos alunos e da autoavaliação também positiva dos produtores, **avalia-se o descritor Formalidade como Adequado.**

A avaliação da categoria Discurso e Linguagens culminou em três descritores considerados Adequados, um Implementado e dois Parcialmente Adequados. Apesar da avaliação geral da categoria poder se considerada satisfatória, verifica-se uma quantidade considerável de itens a se desenvolver ou revisar para futuras versões do *chatbot* didático.

### Categoria 3: Navegação e Usabilidade

A categoria Navegação e Usabilidade tem nove descritores, sendo quatro analisados qualitativamente e cinco avaliados objetivamente. O quadro 30 apresenta os resultados da avaliação da categoria.

QUADRO 32 - Avaliação do Categoria Navegação e Usabilidade

Categoria 3: Navegação e Usabilidade		
<i>Layout</i>	Formato dos textos, imagens, cores, disposição de elementos, entre outras questões estéticas se apresentam de forma agradável e funcional?	( ) Adequado (X) Parcialmente Adequado ( ) Inadequado
<i>Adaptação/individualização do processo de interação e aprendizagem</i>	Cada experiência de interação é única, diferente das dos colegas de curso? Há a possibilidade de os estudantes intervirem no processo de construção do conhecimento ou de estabelecerem um processo único de aprendizagem?	( ) Adequado (X) Parcialmente Adequado ( ) Inadequado
<b>Reprodução de sons e imagens</b>	Sons e imagens (por exemplo, mensagens de voz, músicas, sons diferenciados, desenhos, fotos e vídeos) são reproduzidos	( ) Implementado ( ) Parcialmente Implementado

	pelo <i>chatbot</i> ?	<input checked="" type="checkbox"/> Ausente
<b>Responsividade</b>	A disposição dos elementos visuais se adequa ao tamanho da tela em que usuário acessa o <i>chatbot</i> ?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Acessibilidade</b>	Há recursos tecnológicos para que pessoas com deficiência consigam utilizar o <i>chatbot</i> ?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Uso de hipertextos	São usados hipertextos nas mensagens enviadas pelo <i>chatbot</i> aos alunos para promoção da não linearidade?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Facilidade de uso do sistema	Os alunos têm facilidade em usar o <i>chatbot</i> , identificando rapidamente onde clicar, escrever e outras maneiras de usar o sistema para atingir os objetivos de interação?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Opções de menu e de digitação simultâneas</b>	Ao interagir com o <i>chatbot</i> , os usuários podem escolher entre digitar ou clicar em botões com opções de menu, seja na tela inicial ou em telas subsequentes?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Dos 19 alunos respondentes que interagiram com Katri, 11 apreciaram positivamente o **Layout** do *chatbot*, três fizeram avaliações negativas e cinco não souberam responder sobre este descritor. Não houve expressões de Afeto. As Apreciações foram, em sua maioria, da Composição do *chatbot*, mas houve também Reações e uma Valoração positivas. A maior parte dos alunos usou as palavras simples (Composição) e bom (Reação) para expressar suas opiniões positivas. A aluna 15 disse que o layout de Katri era “simples e bem construído”, indicando aprovação de aspectos da Composição do sistema. O aluno 19 avaliou negativamente a Composição, indicando que “caso seja inserido em ambientes virtuais, seria interessante repensar aspectos de UX (Experiência do Usuário, do inglês *User Experience*)”. A Gradação da avaliação se manifesta por meio da modalização do discurso (“caso seja”, “seria interessante”) e suaviza o tom (atenuação do Foco) da avaliação negativa expressa pelo aluno.

A produtora-linguista usou termos como “leve” e “simples”, que denotam a sua avaliação positiva da Composição de Katri. O produto-antropólogo também avaliou positivamente a

Composição ao dizer que o layout “foi muito bem pensado”. Ele suavizou (Atenuação do Foco) a avaliação positiva ao explicar que ainda há o que melhorar no sistema. “Eu diria que de 1 a 10, provavelmente (eu daria) um 7. É uma boa nota, mesmo. Por que eu não dou 10? Porque eu acho que a gente de algum jeito conseguiria que (Katri) fosse mais interativa”.

**QUADRO 33 - Avaliação do descritor Layout**

<b>Categoria 3: Navegação e Usabilidade</b>						
<b>Descritor: Layout</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	1, 3, 5, 18	Positivo		Reação		
	2	Positivo		Valoração		
	4, 8, 11, 17	Positivo		Composição		
	7	Negativa		Reação		
	10	Positivo		Composição		-
	15	Positivo		Composição		-
	18	Positivo		Composição		+
	19	Negativo		Composição		-
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Linguista Antropólogo	Positiva		Composição		-

Fonte: Dados da pesquisa

Diante da avaliação majoritariamente positiva dos estudantes, mas com ponderações sobre adaptações a se realizar, e da concordância dos produtores sobre a necessidade de implementação de melhorias, mesmo com um layout já “simples”, “leve” e “bem construído”, a análise do descritor culminou na **avaliação como Parcialmente Adequado**. A observação do aluno 19 sobre a necessidade de melhorias na experiência do usuário tem relação direta com a autoavaliação do produtor-antropólogo sobre a necessidade de melhorar as possibilidades de interação por meio da reorganização do *layout* do *chatbot*. Verifica-se que essas adaptações, entretanto, podem ser feitas mantendo um *layout* leve e simples, visto que essas características foram repetidamente apontadas como qualidades de Katri.

O descritor **Adaptação/individualização do processo de interação e aprendizagem** foi avaliado pelos alunos no questionário principalmente por meio da questão “Você sentia que estava participando de uma experiência individualizada de aprendizagem ou que todos os alunos tinham acesso ao mesmo conteúdo? O que te fez entender isso?”. Treze alunos

responderam esta pergunta e seis disseram não saber responder. Dez alunos disseram que a experiência não é percebida como individualizada e três acreditam que há uma experiência de aprendizagem individualizada e/ou adaptativa. Houve, exclusivamente, avaliações da Composição do *chatbot* didático. O aluno 7 afirmou que “ela (Katri) busca uma personalização, mas ainda não é capaz de passar este atendimento individualizado”. Esta avaliação está em linha com a maior parte das avaliações dos outros alunos, que afirmam acreditar que “todos tinha acesso ao mesmo conteúdo (aluna 4)”. A aluna 18, por sua, entendeu que “era personalizada porque respondeu à minha questão específica. Mas a resposta que obtive me pareceu padronizada, o conteúdo serviria a outros alunos com dúvidas semelhantes”. Essa ressalva feita pela aluna atenua o Foco de sua avaliação positiva.

A percepção dos alunos é oposta à intenção e à autoavaliação dos produtores, que demonstram certeza de que o processo de aprendizagem é adaptativo e individual na interação com Katri. A produtora linguista afirmou que

que cada interação é única e personalizada, sim. Ainda que a interação parta de um conjunto finito de possibilidades de interações e também de respostas, a experiência de interação inclui também a jornada do aluno até ali, o objetivo da interação - ‘porque o aluno buscou a Katri?’ -, a formulação da pergunta e a aceitação ou não das sugestões iniciais. Penso que não só o aluno tem poder de intervir como de realmente ‘guiar’ seu processo de aprendizagem, fazendo as escolhas de seguir ou não as opções predefinidas ou perguntar diretamente o que gostaria de saber. (PRODUTORA-LINGUISTA)

**QUADRO 34 - Avaliação do descritor Adaptação/individualização do processo**

<b>Categoria 3: Navegação e Usabilidade</b>						
<b>Descritor: Adaptação/individualização do processo de interação e aprendizagem</b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	3, 8, 9, 12, 15, 16, 18	Negativa		Composição		
	5, 19	Positiva		Composição		
	4, 6, 10	Negativa		Composição		-
	18	Positiva		Composição		-
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Linguista Antropólogo	Positiva		Composição		

Fonte: Dados da pesquisa

Diante das opiniões opostas entre produtores (positiva) e alunos (majoritariamente negativas) **avaliou-se o descritor como Parcialmente Adequado**. Ao analisar os dados de interação do sistema, inclusive a diversidade das interações que diferentes usuários tiveram ao longo do

tempo com Katri (ver Quadros 22 e 23), percebe-se que de fato as experiências são diversas entre diferentes estudantes, o que caracteriza uma experiência adaptativa/individualizada. Entretanto, a não percepção dessa experiência única demonstra que as estratégias discursivas de Katri podem soar “genéricas” ou “padronizadas” (palavras usadas por alguns alunos), o que demonstra que revisões precisam ser feitas para que exista a percepção de personalização da experiência de aprendizagem pretendida pelos produtores. Uso de mais marcas de personalidade no discurso, uso mais frequente do nome do usuário nas trocas de mensagens, entre outras estratégias linguísticas podem ser experimentadas a fim de se obter mais percepção de que a experiência é individual.

Já o **descriptor Reprodução de sons e Imagens foi avaliado como Ausente**, pois Katri se comunica exclusivamente por mensagens de texto escrito, sem uso de mensagens de voz, músicas, desenhos, fotos e outros recursos de som e imagem. O potencial multimodal das plataformas digitais é pouco explorado no *chatbot*, o que poderia trazer mais diversificação de conteúdos e uma experiência estética e pedagógica mais rica. Aqui novamente ressalta-se que o uso dessas diferentes mídias e modos é tecnicamente viável com a tecnologia atual para produção de *chatbots*. Sem descartar a relevância do orçamento disponível para a produção do MD, recomenda-se que uma possível futura versão de Katri explore o seu potencial multimodal, o que pode favorecer a experiência de estudantes com diferentes estilos de aprendizagem.

O *chatbot* é responsivo, ou seja, a disposição dos elementos visuais se adequa ao tamanho da tela em que o usuário acessa o *chatbot*. Por exemplo, na versão mobile, para acesso em *smatphones*, a foto de Katri não é exibida para economizar espaço de tela. A diferença entre o acesso em diferentes *gadgets* é apresentada nas Figuras 25 e 26. Assim, **o descriptor Responsividade foi avaliado como Implementado**.

FIGURA 27 - Versão web de Katri

WATSON ANIMA - Assistente Virtual

há poucos segundos

Você

Para fazer nesta semana

Katri

As atividades são:

- 4.7 Contact sessions in Brazil, Descrição: , encerra em 8/11

**Atenção**, depois dessa, a próxima atividade no calendario é:

- 4.8 Feedback of the programme, Descrição: , encerra em 18/11

Tente fazer uma pergunta como 'Qual as atividades desta semana?'

Fonte: *Chatbot Katri*

FIGURA 28 - Versão mobile de Katri

WATSON ANIMA - Assistente Virtual

Katri

há poucos segundos

Segue abaixo o menu inicial 😊  
Se quiser acessar o ambiente do curso,  
[clique aqui.](#)

há poucos segundos

Você

Calendário Acadêmico

Katri

há poucos segundos

Ei. Dá uma olhadinha no calendário completo do curso. Acho que vai te ajudar a organizar os seus estudos.

[Clique aqui.](#)

Está resposta foi útil pra você?

Sim, me ajudou

Não, você não me ajudou

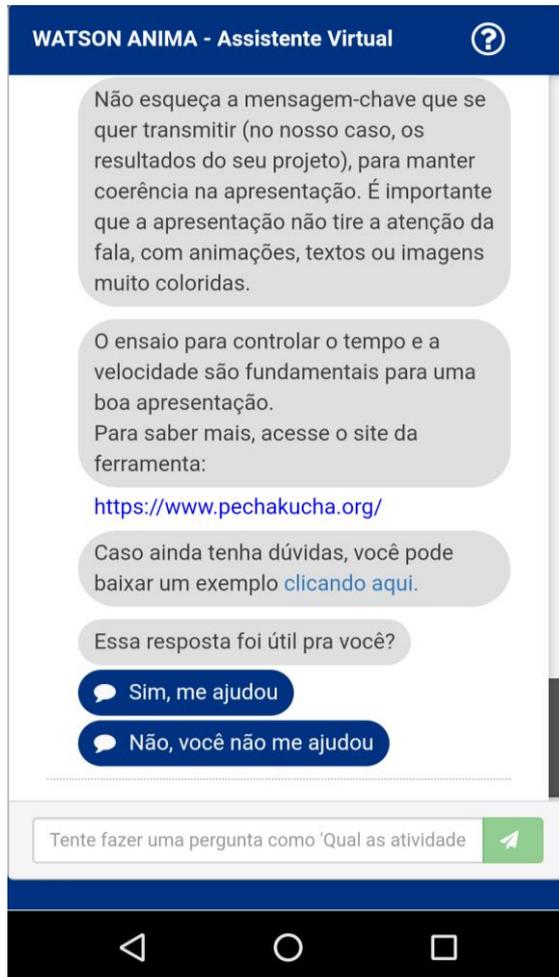
Tente fazer uma pergunta como 'Qual as atividade'

Fonte: *Chatbot Katri*

O *chatbot* não foi produzido considerando questões de acessibilidade e não há funcionalidades específicas para atendimento a alunos com deficiências. Não há, por exemplo, informações em LIBRAS ou em áudio. Na turma em que foi usado, isso não gerou dificuldades de interação, pois não havia alunos com deficiência. Entretanto, considera-se que este é um problema do MD, que merece atenção e desenvolvimento de novas funcionalidades para que possa ser ofertado a novas turmas e a um público mais abrangente. Por isso, **o descritor Acessibilidade foi avaliado como Ausente**. Acessibilidade é um direito e o pleno uso do sistema deve ser garantido a todas as pessoas com deficiência que eventualmente se matriculem no curso. Recomenda-se que *chatbots didáticos* sejam sempre desenvolvidos já com o uso de tecnologias assistivas para que a acessibilidade esteja prontamente disponível quando um aluno/usuário precisar de algum recurso para ter acesso à informação e à experiência de aprendizagem.

Katri envia hiperlinks para acesso a textos e vídeos disponibilizados pelas professoras do curso no Moodle e em outros repositórios de textos e materiais didáticos, e também oferece arquivos para download, promovendo a não linearidade da leitura. Além disso, a própria arquitetura de informação de Katri é hipertextual e não linear. A medida em que interagem com Katri, os alunos podem escolher clicar em diferentes opções de menu ou digitar perguntas, que levam à caminhos de leitura individualizados e não lineares. Assim, **avaliou-se como Implementado o descritor Uso de Hipertextos**. A imagem Y traz um exemplo de resposta que Katri forneceu a um usuário, na qual ela indica um link externo e disponibiliza um material para download.

**FIGURA 29 - Katri fornece link e material para download**



Fonte: *Chatbot Katri*

A Facilidade de Uso do Sistema foi bem avaliada por todos os informantes que abordaram o tema em suas respostas. Todos expressaram avaliações positivas sobre a Composição do *chatbot* e não houve outros tipos de Apreciação ou expressões de Afeto relacionados ao descritor. Adjetivos como “simples” e “prático” foram citados. A maior parte dos alunos afirmou que era “fácil” ou “muito (aumento da Força) fácil” usar o *chatbot*. A aluna 8 afirmou que Katri era “simples e fácil de mexer”. O aluno 10 disse “não lembro de ter dificuldades ao utilizar a plataforma”, avaliação que foi considerada positiva, pois se não houve dificuldades, infere-se que foi fácil usar o sistema, porém considerou-se que houve atenuação do Foco da Apreciação devido à maneira indireta e suavizada de demonstrar essa facilidade.

Os três produtores entrevistados também avaliaram positivamente a Composição do *chatbot* em suas autoavaliações. Os produtores desenvolvedor e antropólogo disseram que o sistema é “objetivo” e por isso fácil de usar. “Por ser um *chatbot* bem objetivo, oferece um conjunto de

opções sem ter que navegar em várias árvores para chegar a um destino. Desta forma, tem-se pequenas interações até chegar ao objetivo do usuário” (PRODUTOR-DESENVOLVEDOR). Entende-se que esta avaliação trata da Composição de Katri, por descrever o quanto é simples e fácil para o aluno conseguir interagir com o sistema e chegar à resposta que procura. A produtora-linguista concorda e completa o raciocínio de seu colega.

Até onde eu acompanhei, os alunos pareceram ter muita facilidade em interagir com o *chatbot*. O uso de *chatbots* já está bem popularizado em outras situações. Muitos sites de venda e atendimento têm *chatbots*, então acho que isso não foi problema. Também acho que a arquitetura inicial com as sugestões por temas predefinidos para clicar ou a possibilidade de fazer uma pergunta diretamente ficou bem clara e bem explicada.

Nos discursos dos produtores, verificou-se o uso de advérbios antecedendo os adjetivos para aumentar a Força das Apreciações da Composição de Katri, como em “bem objetivo” e “muita facilidade”. A produtora-linguista atenuou o Foco da Apreciação em trechos de seu discurso, demonstrando incerteza em excertos como “acho que não foi um problema”.

QUADRO 35 - Avaliação do descritor Facilidade de Uso do Sistema

Categoria 3: Navegação e Usabilidade						
Descritor: Facilidade de uso do sistema						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	5, 8, 11, 12, 17, 19	Positiva		Composição		
	10	Positiva		Composição		-
	16	Positiva		Composição	+	
Avaliação por produtores	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	Linguista Antropólogo Desenvolvedor	Positiva.		Composição	+	
	Linguista	Positiva		Composição		-

Fonte: Dados da pesquisa

Diante da unanimidade nas avaliações positivas, **avaliou-se o descritor Facilidade de uso do sistema como Adequado**. Não se verificou, nas avaliações de alunos e produtores, assim como na análise dos dados de uso do sistema, necessidades de adaptações que tornem o *chatbot* mais fácil de ser usado.

O último descritor da Categoria Navegação e Usabilidade, **Opções de menu e de digitação simultâneas, foi avaliado como Implementado**, pois os usuários podem decidir entre interagir

exclusivamente por meio de cliques nos botões das opções de menu pré-definidas e apresentadas por Katri ou por meio da digitação de perguntas livremente criadas pelos próprios usuários. Podem, também, variar entre o uso das duas mecânicas de interação.

A avaliação da categoria Navegação e Usabilidade teve um descritor considerado Adequado, dois Parcialmente Adequados, três Implementados e dois Ausentes. É uma categoria na qual metade dos descritores tiveram apontamentos para melhorias significativas em possíveis novas versões do *chatbot* didático. Recomenda-se aos produtores especial atenção a esta categoria, a partir da avaliação já apresentada, pois nela há diferentes possibilidades de desenvolvimento para versões aprimoradas de Katri.

#### **Categoria 4: Aparência e Persona**

Aparência e Persona é uma categoria com quatro descritores, sendo dois analisados partir dos discursos avaliativos de alunos e produtores e dois avaliados por meio da verificação da implementação ou ausência do descritor. O resultado das avaliações da categoria é apresentado a seguir no quadro 36.

**QUADRO 36 - Avaliação do Categoria Aparência e Persona**

<b>Categoria 4: Aparência e persona</b>		
Persona do <i>chatbot</i>	Se o <i>chatbot</i> se apresenta como um personagem, ele tem nome, profissão, gênero, humor ou outras características demográficas ou comportamentais expressas por meio de sua imagem e discurso? Os alunos gostam dessa persona e da aparência do <i>chatbot</i> ?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Papel desempenhado pelo <i>chatbot</i>	O <i>chatbot</i> deixa claro qual papel desempenha na situação de interação com o usuário. Por exemplo, é uma colega que vai conversar com o aluno, um atendente administrativo que vai ajudar a resolver questões de secretaria, uma professora que vai dirimir dúvidas pedagógicas etc? Os usuários aprovam esse papel?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado

Revelação de sua identidade de <i>chatbot</i>	O <i>chatbot</i> , em algum momento inicial de interação, revela ao usuário que não é um ser humano respondendo às perguntas, mas uma máquina?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Cumprimento	Há uma mensagem para cumprimentar o usuário? Essa mensagem transmite algum aspecto da persona do <i>chatbot</i> ?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Doze alunos avaliaram a persona do *chatbot*, principalmente por meio de respostas no questionário à questão “Qual é a sua opinião sobre a imagem e personalidade da Katri?”. Alguns alunos, entretanto, avaliaram a persona em suas respostas à questão “Quais são os principais pontos positivos do uso desse *chatbot* no curso?”. Algumas respostas trouxeram Reações positivas expressas ao qualificar as características de Katri como “adequadas” ou “*good* (boas)” e ao dizer que Katri tem uma personalidade “acolhedora”. A aluna 9 reagiu positivamente ao dizer que “Katri apresenta simpatia”. O aluno 5 também expressou Reação positiva, com atenuação do Foco, ao dizer que acredita (incerteza) que a aparência e personalidade de Katri são “adequadas ao curso”. O aluno 7 apreciou o *chatbot* negativamente: “Não gostei (Reação), foi vinculado a foto de uma pessoa. Achei (atenuação do Foco) fora do contexto de um produto de IA (Composição)”. A aluna 20 afirmou “talvez eu mudasse a imagem”, uma Reação negativa, com atenuação do Foco pelo modalizador “talvez”.

QUADRO 37 - Avaliação do descritor Persona do *chatbot*

Categoria 4: Aparência e persona						
Descritor: Persona do <i>chatbot</i>						
Avaliação por alunos	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	1, 15	Positiva		Reação	+	
	3, 4, 9, 11, 17	Positiva		Reação		
	5	Positiva		Reação		-
	7	Negativa		Reação		
	10	Negativa		Reação		-
	2	Positiva		Valoração		
	7, 19	Negativa		Composição		-
Avaliação por produtores	Avaliadores	Polaridade	Afeto	Apreciação	Força	Foco
	Antropólogo	Positiva		Composição	+	
	Antropólogo	Positiva		Composição		-

	Linguista	Positiva		Valoração		-
--	-----------	----------	--	-----------	--	---

Fonte: Dados da pesquisa

Os produtores avaliaram a Aparência e Persona de Katri positivamente, mas com algumas ressalvas que atenuaram o Foco da apreciação. O produtor-antropólogo apreciou positivamente a Composição de Katri “A gente definiu em certa medida (atenuação do Foco) uma personalidade bem (aumento da Força) clara (Composição) para ela. A personagem como tal também para mim foi super bem (aumento da Força) planejada (Composição)”. O produtor-antropólogo explicou que apesar de avaliar positivamente a persona de Katri, gostaria de implementar futuramente algumas mudanças na imagem da monitora virtual, por meio da exibição de diferentes expressões faciais de acordo com o texto que ela envia ao aluno. “Eu gostaria de experimentar mudando a cara da Katri em diferentes faces. Por exemplo, se ela falar “não entendi”, a cara fica triste. Que ela tivesse tipo 5 caras”. (PRODUTOR-ANTROPÓLOGO).

A produtora-linguista valorou positivamente a persona, ressaltando o resultado que algumas características comportamentais de Katri trazem para a interação com os alunos. “Eu acho que ela é alegre o suficiente para não deixar o aluno intimidado. Eu acho que se eu fosse aluna, teria usado bastante, teria interagido, porque teria sido interessante ter conversas com essa persona que é alguém assim, amigável”. (PRODUTORA-LINGUISTA). A modalização do discurso, repetida por meio da palavra “acho”, atenua o Foco da avaliação por não demonstrar certeza e segurança sobre a apreciação realizada.

Com base nas avaliações majoritariamente positivas, mas considerando que parte significativa dos discursos positivos tiveram o Foco atenuado, assim como a ocorrência de algumas avaliações negativas pelos alunos, **o descritor Aparência e Persona foi considerado Parcialmente Adequado.**

Não havia no questionário uma questão específica sobre o Papel Desempenhado pelo *Chatbot*, mas 5 alunos avaliaram o descritor em suas respostas nas perguntas sobre a persona do *chatbot* e sobre os pontos positivos de Katri. Quatro expressaram avaliações positivas. A aluna 9 expressou sua Satisfação ao que dizer que vivenciou “um sentimento positivo ao utilizá-la (Katri) como ferramenta de busca!”. O aluno 19 disse que Katri “facilita a interação e

otimização de suporte técnico”, demonstrando a sua valoração positiva, pois fica implícito que essas características fazem valer a pena interagir com o *chatbot*. O aluno 16, entretanto, expressou uma avaliação negativa da composição de Katri ao sugerir a mudança de gênero da monitora virtual. “Existe hoje uma discussão sobre por que esses auxiliares têm nome sempre femininos, a ideia de que é papel da mulher cuidar, ajudar. Acho que vale a pena colocar um nome com gênero neutro”.

**QUADRO 38 - Avaliação do descritor Papel desempenhado pelo *chatbot***

<b>Categoria 4: Aparência e persona</b>						
<b>Descritor: Papel desempenhado pelo <i>chatbot</i></b>						
<b>Avaliação por alunos</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	9	Positivo	Satisfação			
	15	Positivo		Composição	+	
	16	Negativa		Composição		-
	18, 19	Positiva		Valoração		
<b>Avaliação por produtores</b>	<b>Avaliadores</b>	<b>Polaridade</b>	<b>Afeto</b>	<b>Apreciação</b>	<b>Força</b>	<b>Foco</b>
	Antropólogo	Positiva		Composição		+
	Linguista	Positiva		Composição		-

Fonte: Dados da pesquisa

Os produtores demonstraram estar plenamente satisfeitos com o papel desempenhado por Katri. “Eu não mudaria nada, nem de tom, nem que ela seja monitora, eu não mudaria nada.” (PRODUTOR-ANTROPÓLOGO). A repetição da declaração “eu não mudaria nada” amplia o FOCO da avaliação positiva da Composição.

A produtora-linguista concordou com o seu colega e explicou como a linguagem utilizada por Katri está coerente com o seu papel de monitora virtual do curso, demonstrando sua apreciação positiva da Composição de Katri.

Ela (Katri) não está sentada na mesa do bar conversando com os outros alunos. Ela está numa posição que tem que passar confiança? Por ser monitora, que também, é uma colega, em alguns momentos a interação, os textos, são muito mais leves, muito mais engraçados. São mais leves e depois entra num tom muito mais sério, professoral. Eu acho que se ficasse muito “youtubezinho” assim, muito galerinha, ela perderia um pouquinho da confiança deles, né? Então a gente pensou muito claramente na posição da monitora justamente por isso. Não é uma autoridade, mas é alguém que está no nível deles (professores universitários que são alunos do curso de formação docente) e que pode fornecer uma informação confiável. Se a gente tiver a oportunidade de mexer em alguma coisa antes do próximo grupo, com certeza ainda vai ter uma correção ou outra, mas acho que a Katri como monitora funcionou muito bem.

A afirmação da produtora de que poderia mudar algo em Katri, que “com certeza vai ter uma correção ou outra”, não foi considerada uma atenuação do Foco da avaliação de descritor, pois pelo contexto da avaliação inferiu-se que as mudanças não se dariam no papel desempenhado por Katri, mas em outras funcionalidades e nos textos enviados por Katri aos alunos, pois o discurso avaliativo enfatiza o quanto a produtora acredita que o papel de monitora está aprovado, como “funcionou muito bem”, em linha com o que ela e seus colegas já tinham também avaliado na versão piloto de Katri.

O descritor **Papel Desempenhado pelo *chatbot* foi considerado Adequado**. Todas as avaliações que versaram diretamente sobre o papel de monitora virtual de Katri foram positivas. A avaliação negativa do aluno 16 aponta mais para a questão do gênero feminino atribuído à monitora virtual que ao papel de monitora que ela exerce.

**Os descritores Cumprimento e Revelação de sua identidade de *chatbot*** foram avaliados de maneira objetiva, conforme sugerido na matriz. Os dois descritores **foram considerados Implementados**, pois a primeira mensagem que Katri envia ao usuário, na primeira interação que estabelecem entre si, é: “Olá, (nome do usuário)! Eu sou a Katri, monitora virtual do curso *Teaching and Learning in Higher Education*. Pode me perguntar sobre o curso ou escolher uma das opções abaixo”.

A primeira frase da mensagem é um cumprimento padrão, comum nas interações entre brasileiros para se cumprimentar por mensagens de texto. Já na frase seguinte, Katri se apresenta como uma “monitora virtual” e apresenta qual é a sua função no curso. Considera-se que esta foi uma clara revelação de sua identidade de *chatbot*, pois *chatbots* comumente são apresentados como “assistentes virtuais” e a troca de função para “monitora”, mantendo os padrões textuais e estéticos de um *chatbot*, não prejudicou esse entendimento pelos usuários. Todos os alunos do curso, como fica evidente em suas avaliações sobre Katri, tiveram pleno entendimento que interagiam com uma máquina e não com uma pessoa.

A avaliação da categoria Aparência e Persona teve dois descritores avaliados como Implementados, um Adequado e um Parcialmente Adequado. Essa foi uma categoria com poucas questões a se revisar, mas destaca-se para futuras discussões e possíveis alterações

alguns pontos levantados por alunos, como o uso de gênero neutro ou não binário para que não se caracterize Katri como homem ou mulher e também a possibilidade de que Katri não seja um ser humano. Esta não é uma recomendação de alterações, visto que o *chatbot* já parece funcionar bem e ser bem avaliado com a aparência e persona atual, mas uma questão levantada pelos respondentes e que se considera válido trazer para discussão, visto que tratam de questões sociais contemporâneas relacionadas à gênero e sexualidade e também à tecnologia.

### **Categoria 5: Funcionalidades Adicionais**

Funcionalidades adicionais, última categoria da matriz, tem sete descritores e uma especificidade: todos com avaliação objetiva a partir da verificação da Implementação ou Ausência do descritor.

**QUADRO 39 - Avaliação do Categoria Funcionalidades Adicionais**

<b>Categoria 5: Funcionalidades adicionais</b>		
Resposta a questões desconhecidas	Existe uma ou mais respostas-padrão para os casos em que usuário faz uma pergunta para a qual o <i>chatbot</i> não tem resposta?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Conexão com base de dados externas	O <i>chatbot</i> utiliza bases de dados externas, como o Google, Wikipedia ou outras fontes disponíveis <i>online</i> , e não apenas sua própria base de dados para coletar informações e gerar textos de respostas?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Visualização do histórico de mensagens	O usuário pode usar rolagem de tela ou outro recurso para reler o histórico de mensagens trocadas ao longo do diálogo?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Opção de ajuda	Há algum botão ou opção de ajuda para que o usuário leia informações sobre como utilizar o sistema?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Opção para voltar ao menu inicial	Há algum comando ou opção de onde clicar para voltar ao menu inicial do sistema?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Coleta de <i>feedbacks</i>	Existe algum mecanismo de coleta de <i>feedbacks</i> para que o usuário avalie a qualidade das respostas/interação?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Direcionamento de dúvidas	Dúvidas não respondidas ou demandas não solucionadas pelo <i>chatbot</i> são direcionadas ao docente, coordenador do curso ou alguém que possa dirimir as dúvidas do usuário?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

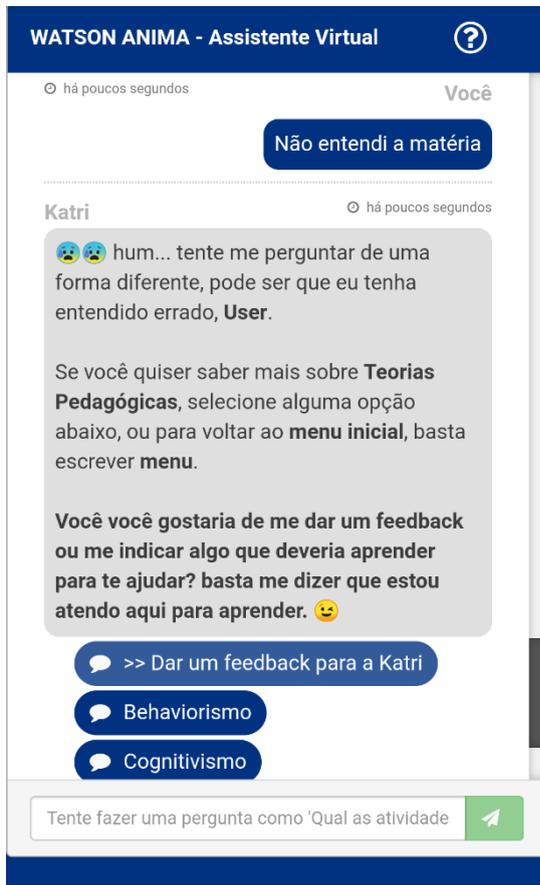
Fonte: Dados da pesquisa

Quando o aluno faz uma pergunta sobre tema que Katri desconhece ou que não identifica como responder, existem mensagens já programadas como resposta. Katri responde, por exemplo: “Humm, não entendi. Tente me perguntas de uma maneira diferente”. Quando não reconhece o assunto abordado, Katri explica que não conhece o tema e fornece o endereço e-mail do coordenador do curso, para que o aluno possa direcionar a ele a sua dúvida, caso assim deseje. O direcionamento é opcional e o aluno pode seguir ou não a recomendação de Katri.

Quando a monitora virtual reconhece em seu banco de dados alguma palavra-chave na pergunta do aluno, mas não tem certeza sobre o que ele quer falar, Katri explicita a sua incerteza e oferece opções de resposta relacionadas ao que ela reconhece como provável tema da questão, além de pedir um feedback ao usuário para que possa aprimorar o seu entendimento do tema e a sua maneira de responder, conforme mostra a Figura 28. Assim, **os descritores Resposta a questões desconhecidas e Direcionamento de dúvidas foram considerados Implementados**, pois fazem parte dos padrões de resposta que Katri consegue enviar aos alunos.

Figura 28: Resposta a questões desconhecidas

**FIGURA 30 - Resposta a questões desconhecidas**



Fonte: *Chatbot Katri*

Katri busca informações para enviar mensagens aos alunos apenas em sua própria base de dados. Não houve desenvolvimento de integração do sistema com outras fontes de dados, como o Google, Wikipedia ou o próprio Ambiente Virtual de Aprendizagem do curso para que Katri busque a informação externamente, em bancos de dados que não sejam o próprio banco criado para o sistema do *chatbot*. A implementação dessa busca externa poderia trazer ganhos na qualidade das respostas fornecida por Katri a usuários, inclusive fornecendo múltiplas de fontes de informação. Assim, o **descriptor Conexão com base de dados externas foi avaliado como Ausente.**

À direita da tela, tanto no celular quanto no computador, o usuário pode usar uma barra de rolagem na janela do navegador para reler o histórico de mensagens trocadas ao longo do diálogo em curso. Entretanto, não há recursos para que o usuário leia as mensagens trocadas em diálogos de interações anteriores com Katri. A cada vez que a janela é fechada, o histórico

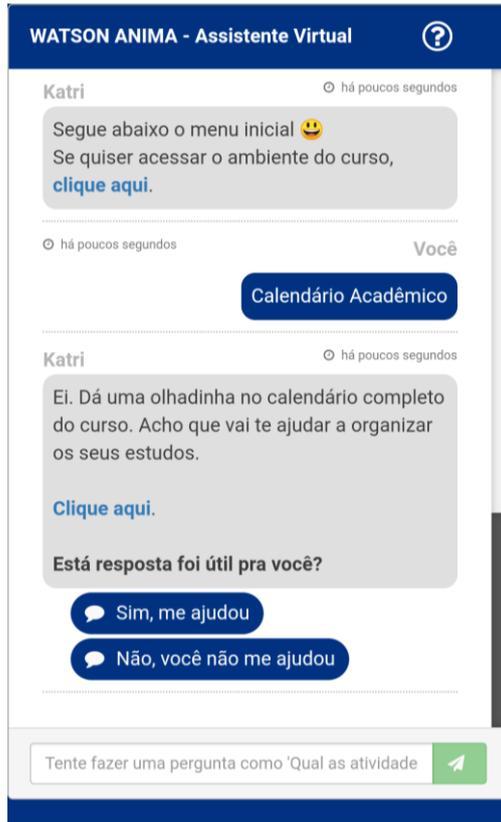
se apaga. Ou seja, não há visualização do histórico completo, mas apenas do histórico da conversa atual. Diante da maneira como essa funcionalidade está presente no *chatbot*, **avaliou-se o descritor Visualização do Histórico de Mensagens como Parcialmente Implementado**. Recomenda-se, para possíveis futuras versões de Katri, que algum recurso de recuperação de diálogos anteriores seja implementado.

Não há nenhuma opção de Ajuda para que o usuário possa tirar dúvidas sobre o uso do sistema ou mesmo acessar algum tutorial que o apoie na aprendizagem sobre como interagir com Katri. Assim, **o descritor Opção de Ajuda foi considerado Ausente**. Nenhum aluno, entretanto, informou ter tido dificuldades para interagir com Katri, o que pode significar que a mesmo com a avaliação Ausente, não é necessário implementar a funcionalidade. Essa avaliação chama atenção para o fato de que nem todos os descritores da matriz precisam ser implementados ou considerados adequados para que o *chatbot* promova uma boa experiência de interação e aprendizagem.

Em algumas mensagens, Katri explica ao usuário que para voltar ao menu inicial, a fim de reiniciar a interação, ele deve escrever a palavra “menu”, conforme mostra a Figura 28. Não há, entretanto, nenhum botão para voltar ao menu inicial. Essa opção fica disponível exclusivamente por meio de digitação, o que exige que o aluno tenha aprendido o comando que deve utilizar por escrito para que possa fazer uso deles. A existência de um botão de “voltar ao menu” ou “voltar ao início” poderia ser mais intuitiva. Portanto, há um comando, mas não uma opção clicável para voltar ao menu inicial do sistema, motivo pelo qual o **descritor “Opção para voltar ao menu inicial” foi avaliado como Parcialmente Implementado**.

Katri também pede feedbacks aos estudantes. Em algumas mensagens (ver Figura 29), ela explica o motivo: ainda está aprendendo e os feedbacks contribuem para o seu desenvolvimento. Quando Katri supõe ter fornecido ao usuário a informação solicitada, ela pergunta se foi útil e dá duas opções de resposta: “sim, me ajudou” e “não, você não me ajudou”. O usuário pode, também, digitar feedbacks se desejar. Assim, **considerou-se que o descritor Coleta de feedbacks está Implementado**.

**FIGURA 31 - Pedido de feedback**



Fonte: *Chatbot Katri*

A categoria Funcionalidades Adicionais teve 3 descritores Implementados, 2 Parcialmente Implementados e 2 Ausentes. Na matriz, considerou-se que esses descritores teriam avaliação objetiva, a partir da verificação da Implementação ou Ausência do descritor, sem necessidade de coletar opiniões de usuários de produtores. Entretanto, apesar de não ser obrigatória, essa coleta pode ser feita em caráter optativo, caso seja desejável e viável para os avaliadores/educadores, pois pode trazer pontos de vista relevantes para a análise e melhorias do *chatbot* pedagógico.

Para se ter uma melhor visualização da validação da matriz avaliação de *chatbots* didáticos, apresenta-se no Quadro 38 a matriz completa com a avaliação de suas cinco categorias e 30 descritores. Destes, 18 foram considerados Implementados ou Adequados, 8 Parcialmente Implementados ou Parcialmente Adequados, e 4 Ausentes.

QUADRO 40 - Avaliação Completa da Matriz

<b>Categoria 1: Conteúdos</b>		
Descritor	Questões norteadoras	Avaliação
Conteúdo de apresentação	O <i>chatbot</i> é capaz de se apresentar, dizendo quem é e para que serve?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Conhecimentos especializados	O quão completo é o conhecimento do <i>chatbot</i> sobre os conteúdos trabalhados no curso ou disciplina que ele atende, incluindo conteúdo (conceitos, teorias etc), atividades e agenda do curso. Os conteúdos são válidos e embasados cientificamente?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Atualização dos conteúdos	Os conteúdos são atualizados em relação à versão do curso ofertado, caso o <i>chatbot</i> esteja sendo ofertado em períodos letivos consecutivos? Por exemplo, o calendário do curso foi atualizado para a nova turma?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Antecipação às dúvidas mais comuns dos estudantes	As informações ofertadas pelo <i>chatbot</i> via opções de menu são capazes de antecipar e responder às possíveis dúvidas mais comuns dos estudantes?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Conexão com outros conteúdos do curso</b>	Os conteúdos apresentados no <i>chatbot</i> fazem menção e estão diretamente relacionados com outros conteúdos do curso?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Categoria 2: Discurso e linguagens</b>		
Modalidades oral e escrita de uso da língua	O <i>chatbot</i> aceita <i>inputs</i> por texto oral (áudio) e escrito e dá respostas também com texto oral e escrito?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Clareza	Os textos são apresentados de maneira clara e coerente para entendimento dos alunos?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Lembrança do nome do usuário	O <i>chatbot</i> consegue identificar o nome do usuário e se referir a ele pelo nome durante a conversa?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

<b>Suficiência de informações</b>	A quantidade de conteúdo é suficiente para dar apoio aos estudos autônomos dos alunos? A extensão/tamanho dos textos é adequada ao que os leitores querem ou estão dispostos a ler?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Adequação do vocabulário ao curso.</b>	O vocabulário usado pelo <i>chatbot</i> está coerente e em sintonia com o vocabulário usado nos outros materiais didáticos do curso? Os diálogos são realizados em linguagem natural com registros próximos aos usados pelos usuários/alunos no curso?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Formalidade	O grau de formalidade está adequado ao contexto, ao público-alvo e à persona do <i>chatbot</i> ?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Categoria 3: Navegação e Usabilidade</b>		
<i>Layout</i>	Formato dos textos, imagens, cores, disposição de elementos, entre outras questões estéticas se apresentam de forma agradável e funcional?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Adaptação/individualização do processo de interação e aprendizagem</b>	Cada experiência de interação é única, diferente das dos colegas de curso? Há a possibilidade de os estudantes intervirem no processo de construção do conhecimento ou de estabelecerem um processo único de aprendizagem?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<b>Reprodução de sons e imagens</b>	Sons e imagens (por exemplo, mensagens de voz, músicas, sons diferenciados, desenhos, fotos e vídeos) são reproduzidos pelo <i>chatbot</i> ?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
<b>Responsividade</b>	A disposição dos elementos visuais se adequa ao tamanho da tela em que usuário acessa o <i>chatbot</i> ?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Acessibilidade</b>	Há recursos tecnológicos para que pessoas com deficiência consigam utilizar o <i>chatbot</i> ?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Uso de hipertextos	São usados hipertextos nas mensagens enviadas pelo <i>chatbot</i> aos alunos para promoção da não linearidade?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Facilidade de uso do sistema	Os alunos têm facilidade em usar o <i>chatbot</i> , identificando rapidamente onde clicar, escrever e outras maneiras de usar o sistema para atingir os objetivos de interação?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
<i>Opções de menu e de digitação simultâneas</i>	Ao interagir com o <i>chatbot</i> , os usuários podem escolher entre digitar ou clicar em botões com opções de menu, seja na tela inicial ou em telas subsequentes?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Categoria 4: Aparência e persona</b>		
Persona do <i>chatbot</i>	Se o <i>chatbot</i> se apresenta como um personagem, ele tem nome, profissão, gênero, humor ou outras características demográficas ou comportamentais expressas por meio de sua imagem e discurso? Os alunos gostam dessa persona e da aparência do <i>chatbot</i> ?	<input type="checkbox"/> Adequado <input checked="" type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Papel desempenhado pelo <i>chatbot</i>	O <i>chatbot</i> deixa claro qual papel desempenha na situação de interação com o usuário. Por exemplo, é uma colega que vai conversar com o aluno, um atendente administrativo que vai ajudar a resolver questões de secretaria, uma professora que vai dirimir dúvidas pedagógicas etc? Os usuários aprovam esse papel?	<input checked="" type="checkbox"/> Adequado <input type="checkbox"/> Parcialmente Adequado <input type="checkbox"/> Inadequado
Revelação de sua identidade de <i>chatbot</i>	O <i>chatbot</i> , em algum momento inicial de interação, revela ao usuário que não é um ser humano respondendo às perguntas, mas uma máquina?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Cumprimento	Há uma mensagem para cumprimentar o usuário? Essa mensagem transmite algum aspecto da persona do <i>chatbot</i> ?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
<b>Categoria 5: Funcionalidades adicionais</b>		
Resposta a questões desconhecidas	Existe uma ou mais respostas-padrão para os casos em que usuário faz uma pergunta para a qual o <i>chatbot</i> não tem resposta?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Conexão com base de dados externas	O <i>chatbot</i> utiliza bases de dados externas, como o Google, Wikipedia ou outras fontes disponíveis <i>online</i> , e não apenas sua própria base de dados para coletar informações e gerar textos de respostas?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Visualização do histórico de mensagens	O usuário pode usar rolagem de tela ou outro recurso para reler o histórico de mensagens trocadas ao longo do diálogo?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Opção de ajuda	Há algum botão ou opção de ajuda para que o usuário leia informações sobre como utilizar o sistema?	<input type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input checked="" type="checkbox"/> Ausente
Opção para voltar ao menu inicial	Há algum comando ou opção de onde clicar para voltar ao menu inicial do sistema?	<input type="checkbox"/> Implementado <input checked="" type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Coleta de <i>feedbacks</i>	Existe algum mecanismo de coleta de <i>feedbacks</i> para que o usuário avalie a qualidade das respostas/interação?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente
Direcionamento de dúvidas	Dúvidas não respondidas ou demandas não solucionadas pelo <i>chatbot</i> são direcionadas ao docente, coordenador do curso ou alguém que possa dirimir as dúvidas do usuário?	<input checked="" type="checkbox"/> Implementado <input type="checkbox"/> Parcialmente Implementado <input type="checkbox"/> Ausente

Fonte: Dados da pesquisa

Katri teve uma avaliação majoritariamente positiva, mas conforme os dados apresentados, há ainda muitos espaços para implementação de melhorias e adequações, o que é considerado algo normal nesse tipo de sistema, que pode ter versões aprimoradas ao longo do tempo. Os itens considerados Ausentes ou para os quais foram apresentadas críticas negativas ou oportunidades de melhorias não representam um problema sob o ponto de vista do processo de avaliação de MD. Ao contrário, trata-se do resultado esperado de um amadurecimento do processo de produção. Conforme Cabero (2001), o ciclo de avaliação de MD deve se repetir sucessivamente durante o tempo em que o material é usado. E a cada vez que a avaliação é feita, espera-se que se encontre (novamente) pontos de melhoria, pois somente com esses apontamentos haverá oportunidades de desenvolvimento para uma versão do MD que ajude a promover mais aprendizagens que a versão anterior.

No caso de Katri, a cada vez que for usada com uma nova turma, poderia haver uma nova versão do *chatbot* didático. Assim, ela se aproxima (ou se distancia menos) também de outros tipos de materiais didáticos, como os livros didáticos, que podem ter novas edições lançadas, com atualizações, novos conteúdos e aprimoramentos, e as videoaulas, que podem ser reeditadas ou regravadas para promover melhor experiência de aprendizagem.

Mesmo sem solucionar todos os problemas do *chatbot* didático usado no curso, os processos de avaliação e produção de MD do caso estudado foram positivos. A maior parte dos problemas verificados na versão piloto de Katri foram solucionados, mostrando que a equipe de produção não ignorou os resultados da avaliação ou insistiu em procedimentos considerados equivocados na primeira etapa de avaliação. Dos 11 descritores avaliados no piloto, 4 já haviam recebido a melhor avaliação possível segundo a matriz proposta. Dos 7 descritores em que foram evidenciados espaços de melhoria, 5 receberam avaliação mais positiva na versão definitiva de Katri e 2 mantiveram o mesmo resultado de avaliação nas duas versões. Ou seja, houve percepção de melhoria do MD.

Na versão definitiva, foram apontados novos problemas no *chatbot* didático, inclusive relacionados àqueles descritores que já haviam sido avaliados na versão piloto. O fato de novos argumentos serem apresentados sobre o mesmo item é interpretado aqui como um amadurecimento do processo de avaliação, pois novos alunos trazem novas perspectivas e os produtores, ao repetirem o processo de avaliação, podem aprimorar o seu olhar sobre o MD e dar atenção a questões antes não percebidas. A expectativa é que, havendo novos ciclos de avaliação de Katri, mais pontos de melhoria sejam apresentados e que isso contribua para que os alunos e professores que interajam com o *chatbot* tenham em cada oferta do curso uma melhor experiência.

O aprimoramento de Katri, quando comparadas as duas versões, e as indicações de novas melhorias no segundo ciclo avaliativo, com base nos critérios e indicadores apresentados, são indicativos da viabilidade e eficiência do uso da matriz, assim como do uso do Sistema de Avaliatividade para a análise dos dados coletados, pois esse é um resultado esperado do processo de avaliação de MD.

A organização das categorias de avaliação e a proposta de análise foi feita durante a pesquisa, enquanto Katri era produzida. Seria muito diferente produzir o *chatbot* didático já com uma

matriz como a proposta nesta pesquisa como base para se referenciar durante o processo de produção. Uma matriz como a apresentada nesta tese traria novas perspectivas e poderia ser um guia (não um recurso de uso mandatório) para auxiliar a produção e trazer de partida recomendações de funcionalidades e perspectivas que merecem atenção na produção deste tipo de MD.

Considera-se que Katri implementa as quatro possibilidades que, segundo Lima (2014), deveriam ser entregues por esse tipo de sistema: 1) Katri proporciona uma interface de conversação natural e que pareceu ser interessante para produtores e usuários; 2) Katri fornece conteúdo especializado sobre o curso/contexto de estudos dos usuários e é adaptável, pois os alunos de fato conseguem ter diferentes experiências de conversação de acordo com os *inputs* que dão ao *chatbot*; 3) ela dá liberdade de diálogo aos alunos, pois aqueles que por timidez ou outro motivo não querem expor suas perguntas em sala de aula, podem o fazer de maneira individual e até anônima na interação via sistema; e 4) Katri realiza o atendimento aos alunos a qualquer horário, em qualquer dia da semana, por meio do acesso on-line.

A matriz proposta mostrou-se eficaz para a avaliação do *chatbot* Katri. Foi possível avaliar todos os descritores, de todas as categorias, e essa avaliação trouxe reflexões sobre o atual estágio do *chatbot*, reconhecimento de pontos positivos e funcionalidades implementadas, assim como apontou problemas a se resolver e revelou fragilidades do MD. Essa avaliação ainda proporcionou *insights* para promover melhorias em futuras versões de Katri. Esse movimento de avaliação segue a proposta de Cabero (2001), que afirma que os resultados da avaliação de um MD devem indicar as necessidades de revisão.

Richter (2006) explica que a avaliação pós-uso de um MD, estratégia utilizada nesta pesquisa, permite identificar vantagens ou desvantagens após um período de uso do MD para que sejam tomadas decisões sobre a continuidade ou não do uso do material. No caso de Katri, recomenda-se a continuidade do uso, principalmente se os pontos de melhoria apontados forem considerados, como a implementação de tecnologias assistivas para tornar o *chatbot* acessível para pessoas com deficiência, por exemplo.

O ciclo de avaliação de MD apresentado por Cabero (2001) foi utilizado duas vezes: no piloto e na versão definitiva de Katri. Entretanto, pode-se pensar que se o curso for novamente ofertado, Katri pode não ter tido ainda a sua versão definitiva (apesar de sua segunda versão ter

sido tratada como a definitiva), pois com a implementação de melhorias, uma nova versão do *chatbot* poderia entrar no ar. Assim, seria relevante realizar um terceiro ciclo de avaliação do MD para identificar os possíveis avanços e até mesmo desvantagens da possível nova versão. Se outras edições do curso forem lançadas, novos ciclos de avaliação e novas versões de Katri podem se mostrar necessários.

A utilização do Sistema de Avaliabilidade (MARTIN; WHITE, 2005), mostrou-se uma base adequada para analisar os textos avaliativos de produtores e alunos, servindo de base para avaliação do *chatbot* e preenchimento da matriz. A identificação das Atitudes e da Gradação revelou subjetividades dos textos avaliativos e permitiu a classificação e organização dos dados para tomada de decisão no momento de preenchimento da matriz.

Foi possível perceber como na maior parte dos textos avaliativos as Atitudes de alunos e produtores de Katri estão mais ligadas à Apreciação e menos ao Afeto. Essa análise da prevalência de um tipo de Atitude é relevante para que os produtores possam refletir se os objetivos traçados estão de fato sendo alcançados. A análise da Gradação envolvida nos textos avaliativos também revelou o que existe de mais impactante no *chatbot*, positiva ou negativamente. Por meio da análise da Gradação, com muitos textos que ampliaram o Foco ou aumentaram a Força da Apreciação, foi possível ver o quanto a Clareza dos textos é um ponto de destaque positivo de Katri. Já o descritor Persona do *Chatbot* teve um considerável número de textos com marcas de atenuação do Foco, o que sugere aos produtores darem atenção a possíveis pontos de melhoria relacionados a esse descritor.

Desta maneira, o Sistema de Avaliabilidade deu base para que houvesse uma clara visualização do que significavam os textos avaliativos proferidos pelos informantes da pesquisa. Por meio dessa estratégia de análise e avaliação, considera-se validada a matriz de avaliação de *chatbots* didáticos e sugere-se o seu uso para avaliação de outros *chatbots* usados para fins de ensino e aprendizagem.

Recomenda-se o uso do Sistema de Avaliabilidade como base para análise e preenchimento da matriz, mas faz necessário ressaltar que os avaliadores que fizerem uso da matriz devem ter liberdade para lançar mão de outras teorias ou *frameworks* que julguem servir como base para sua própria análise e avaliação.

## 7. CONSIDERAÇÕES FINAIS

A motivação principal para realização desta pesquisa foi a observação do pesquisador, ao trabalhar na produção de um *chatbot* didático, de que as pessoas que compunham a equipe de produção, inclusive ele próprio, não tinham nenhum material-referência, muito menos uma matriz de avaliação, para guiar os processos de produção e de avaliação do MD. Ao se deparar com essa situação em sua prática profissional, chegou-se à proposta desta pesquisa que teve como objetivo geral criar e validar o uso de uma matriz com critérios e descritores para avaliação de *chatbots* didáticos. Considera-se que este objetivo foi alcançado, pois uma matriz com 5 categorias e 30 descritores foi criada e testada em contexto real com alunos e produtores.

Esta matriz agora pode ser disponibilizada por meio desta tese para que profissionais que passem por situação similar à vivida por este pesquisador e a equipe de produção de Katri, possam ter um material-guia ou, no mínimo, uma matriz de referência que de alguma maneira contribua na decisão sobre funcionalidades, organização textual, estética, interação e contribuições didáticas deste tipo de MD para os cursos em que os *chatbots* forem usados. Espera-se que os resultados desta pesquisa sejam úteis na criação de novos *chatbots* didáticos.

O objetivo geral da pesquisa foi alcançado para se cumprir o que foi proposto como objetivos específicos. O primeiro, mapear critérios para produção e avaliação de *chatbots* didáticos, foi alcançado por meio da conexão entre critérios de avaliação de materiais didáticos, especialmente os livros didáticos e os materiais interativos online, e critérios de avaliação de *chatbots* usados para fins não pedagógicos. Foi com base no mapeamento de diversas referências, como Cabero (2001), Belisário (2003), Tomlinson e Masuhara (2004), Silva, Parreiras e Fernandes (2015), Kuligowska (2015), Radziwill e Benton (2017), que foi possível pensar em uma matriz inovadora para um tipo de MD tão recente.

Observar a produção e o uso de um *chatbot* didático em um curso de formação de professores foi um objetivo específico alcançado por meio de um estudo de caso com observação participante, na perspectiva da autonetnografia (KOZINETZ, 2014): a experiência de produção do *chatbot* Katri como MD do curso de formação docente Teaching and Learning in Higher Education.

O objetivo de se utilizar a matriz proposta, com seus critérios e descritores, para avaliar o *chatbot* do curso TLHE considerando as percepções de produtores deste *chatbot* e alunos do curso foi alcançado com a aplicação da matriz com duas turmas do curso: primeiramente, 11 descritores foram usados na versão piloto do *chatbot* com a turma que iniciou o curso em 2018. Em seguida, com a turma do ano de 2019, a matriz completa, com seus cinco critérios e 30 descritores, foram usados para avaliar Katri. Nas duas versões, alunos do curso e produtores do *chatbot* participaram do processo de avaliação do MD, como sugere Cabero (2001).

A partir das reflexões sobre os resultados dos dois ciclos de avaliação realizados, foi possível alcançar o último objetivo específico da pesquisa: sugerir melhorias e revisões no *chatbot* com base na avaliação realizada com o uso da matriz proposta. Essas melhorias foram sugeridas na discussão dos resultados que se fez a medida em que a análise dos dados coletados era apresentada no capítulo 6 desta tese. Foram muitos os acertos na produção de Katri, mas também muitos os pontos de melhoria apresentados. Ressalta-se que não era o objetivo desta pesquisa construir um *chatbot* didático perfeito, mas estabelecer uma matriz para avaliá-lo, justamente para que tais pontos de melhoria pudessem ser apontados, proporcionando o espaço de discussão para desenvolvimento de melhores versões do MD. Com o cumprimento destes objetivos, validou-se a viabilidade e eficiência do uso da matriz em um contexto real.

Validou-se, também, a hipótese de que era possível criar uma matriz com critérios para avaliação de *chatbots* usados como material didático, tendo como base critérios para avaliação de materiais didáticos de outras naturezas e de critérios para avaliação de *chatbots* usados para fins não didáticos. Esse movimento, de combinação e diálogo entre diferentes fontes, foi essencial para que a matriz pudesse ser criada, devido à escassez de referências que apoiassem a produção deste tipo de MD que se utiliza de inteligência artificial.

Acredita que, assim como esta tese está sendo publicada agora em 2022, outros pesquisadores no mundo também já passaram pela mesma dificuldade na produção sentida pelos produtores de Katri, principalmente após o início da pandemia de COVID-19, quando acelerou-se a produção de *chatbots* para uso em contextos educacionais. Assim, espera-se que em breve se possa encontrar mais publicações que possam dialogar com esta tese, também propondo critérios para produção e avaliação de *chatbots* didáticos. Pesquisas futuras poderiam estabelecer o diálogo entre as propostas de autores de diferentes países e contextos.

A proposta de critérios e descritores para avaliação de *chatbots* didáticos apresentadas nesta pesquisa é uma primeira proposta para que esse tipo de MD possa ser constantemente avaliado e reelaborado. É uma contribuição inicial para o preenchimento de uma lacuna existente nesta área.

Espera-se que o relato do percurso desta pesquisa seja útil para educadores e produtores de *chatbots* didáticos, guiando-os por meio da matriz ou inspirando seus processos de produção, uso e avaliação dos *chatbots*. Como parte da contribuição à comunidade acadêmica, educadores e produtores podem chegar com mais facilidade a acertos na produção de *chatbots* didáticos ou evitar estratégias apresentadas nesta pesquisa que resultaram em avaliações negativas.

Reconhece-se a limitação desta pesquisa por aplicar a matriz em um único caso e sugere-se, também, que outros pesquisadores a apliquem e divulguem os resultados obtidos, assim como sugestões de melhorias na matriz ou mesmo novas sugestões de matrizes, utilizando a matriz proposta nesta tese como uma referência. Futuras pesquisas poderiam também contribuir para ampliar a relação de categorias e descritores, além da criação de detalhadas grades de avaliação para os descritores a serem avaliados. Também sugere-se que futuras pesquisas possam reafirmar o uso do Sistema de Avaliatividade (MARTIN; WHITE, 2005) para análise dos dados obtidos na avaliação ou sugerir novas propostas de análise dos dados coletados no uso da matriz.

Espera-se a contribuição da comunidade científica para revalidar a matriz, refutá-la ou sugerir aprimoramentos, pois a construção coletiva do conhecimento é bem-vinda e a expectativa é que essa seja a primeira versão da matriz de avaliação de *chatbots* didáticos, e não a única.

## REFERÊNCIAS BIBLIOGRÁFICAS

AGRAWAL, S. et al. Are word embedding and dialogue act class-based features useful for coreference resolution in dialogue. In: **Proceedings of PACLING**. 2017.

Disponível em: <https://www.cse.iitb.ac.in/~pb/papers/pacling17-dialogue-coref.pdf>. Acesso em 16 nov 2019.

ALLEN, J. F.; FERGUSON, G.; STENT, A. An architecture for more realistic conversational systems. In: **Proceedings of Intelligent User Interfaces 2001** (IUI-01), pp. 1 - 8, Santa Fe, NM. 2001. Disponível em: <https://www.cs.rochester.edu/research/cisd/pubs/2001/allen-ferguson-stent-iui2001.pdf>. Acesso em: 15 nov 2019.

ALMEIDA JUNIOR, O.A. **Beck: Um Chatbot Baseado na Terapia Cognitivo-Comportamental para Apoiar Adolescentes com Depressão**. Mestrado em Ciência da Computação. Centro de Informática, Programa de Pós-graduação em Ciência da Computação. Universidade Federal de Pernambuco. Recife, 2017. Disponível em <https://repositorio.ufpe.br/bitstream/123456789/26790/1/DISSERTA%C3%87%C3%83O%20Oberdan%20Alves%20de%20Almeida%20Junior.pdf>

AMARAL, A. Autonetnografia e inserção *Online* : o Papel do Pesquisador-Insider nas Práticas Comunicacionais das Subculturas da Web. **Revista Fronteiras – Estudos Midiáticos**. Unisinos. V. 14-24, 2009. Disponível em <http://revistas.unisinos.br/index.php/fronteiras/article/view/5037>>. Acesso em: 22/05/2019.

ANDRÉ, M. E. D. A. **Estudo de caso em pesquisa e avaliação educacional**. Brasília: Liberlivros, 2005.

ARAUJO, M. A criação de *chatbots* é o futuro do trabalho do escritor? Gerando narrativas de ficção em sociedades algorítmicas. In: A personagem nos mundos possíveis do insólito ficcional. **III Congresso Internacional Vertentes do Insólito Ficcional**. Rio de Janeiro: UERJ/ Instituto de Letras, 2016, p. 208.

BAHJA, M.; HAMMAD, R. ; BUTT, G. A User-Centric Framework for Educational *Chatbots* Design and Development. **22nd International Conference on Human Computer Interaction, HCII 2020**; 12424 LNCS:32-43, 2020.

BANDEIRA, D. Material didático: conceito, classificação geral e aspectos da elaboração. In: CIFFONE, H. (Org.). **Curso de MD para smartphone e tablet**. Curitiba, IESDE, 2009, p. 13-33. Disponível em: <http://www2.videolivrraria.com.br/pdfs/24136.pdf>. Acesso em: 18 nov. 2019.

BANDEIRA, A.; STANGE, C.; SANTOS, J. M. T. **Uma proposta de critérios para análise de livros didáticos de Ciências Naturais na Educação Básica**. III Simpósio Nacional de Ensino de Ciência e Tecnologia, Ponta Grossa – PR, 2012.

BARDIN, L. **Análise de conteúdo**. Lisboa. Portugal. Edições 70, 2011.

BELISÁRIO, A. O material didático na educação a distância e a constituição de propostas interativas. In: SILVA, M. (org.) **Educação online: teoria, práticas, legislação e formação corporativa**. São Paulo: Loyola, 2003. p. 137-148

CABERO, J. C. Avaliar para melhorar: meios e materiais de ensino. In: SANCHO, M. (Org.). **Para uma tecnologia educacional**. Trad. Beatriz Affonso Neves. Porto Alegre: Artes Médicas, 2001.

CABRAL, V. **Macrotendências na educação de hoje**. Congresso Bett Educar, 2019. Disponível em <https://www.bettbrasileducar.com.br/bettblog/macrotend%C3%A2ncias-na-educa%C3%A7%C3%A3o-de-hoje>>. Acesso em: 26 mai 2019.

CÂMARA, R. H. Análise de conteúdo: da teoria à prática em pesquisas sociais aplicadas às organizações. **Gerais: Revista Interinstitucional de Psicologia**, v. 6, n. 2, p. 179-191, 2013.

CAMARGO, L. D. V. **Produção e Avaliação de Materiais Didáticos Audiovisuais para Ensino de LIBRAS a Distância**. **Dissertação**. Mestrado em Estudos de Linguagens. Centro

Federal de Educação Tecnológica de Minas Gerais, 2013.

CASAGRANDE, R. **Inteligência artificial e a educação além da curva**. Curitiba: Alleanza Educação, 2019.

CORRÊA, M. V.; ROZADOS, H. B. F. A netnografia como método de pesquisa em Ciência da Informação. **Encontros Bibli: revista eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 22, n. 49, 2017. p. 1-18. Disponível em <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n49p1/34047>>.

Acesso em: 20 Abr. 2019.

CHO, W. et al. **Disambiguating Speech Intention via Audio-Text Co-attention Framework: A Case of Prosody-semantics Interface**. Seoul National University, 2019. Disponível em <<https://arxiv.org/pdf/1910.09275.pdf>>. Acesso em: 09 NOV.

2019.

CRUZ, D. M. **O professor midiático: a formação docente para a educação a distância no ambiente virtual da videoconferência**. Tese (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina, UFSC, Florianópolis, 2001. Disponível em <<https://repositorio.ufsc.br/handle/123456789/81518>>. Acesso em: 29 de dez. 2018.

CUNNINGSWORTH, A. **Choosing your coursebook**. Oxford: Heineman, 1995.

DEMARQUES, E. A. **Estudo exploratório sobre compatibilização semântica de vocabulário utilizando ontologia: fundamentação teórica- metodológica visando o aperfeiçoamento do MHTX**. Dissertação de Mestrado. Programa de Pós-Graduação em Ciência da Informação. Universidade Federal de Minas Gerais. Belo Horizonte, 2011.

DERYUGINA, O. V. **Chatterbots**. *Scientific and Technical Information Processing*, v. 37, n. 2, p. 143-147, 2010.

DIBITONTO, Massimiliano et al. *Chatbot in a campus environment: design of LiSA, a virtual assistant to help students in their university life*. In: **International Conference on Human-Computer Interaction**. Springer, Cham, 2018. p. 103-116.

DUNCAN, K. **Global Review of Commercial AI: Artificial Intelligence and Machine Learning Report for Business.** The Digital Finance Institute. Vancouver, 2018.

EGGINS, Suzanne. **Introduction to systemic functional linguistics.** A&c Black, 2004.

EICHER, B.; POLEPEDDI, L.; GOEL, A. Jill Watson doesn't care if you're pregnant: Grounding AI ethics in empirical studies. In: **Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society.** 2018. p. 88-94.

FERREIRA, L. P.; UCHÔA, J. Q. Desenvolvimento de um *Chatbot* para auxiliar o ensino de Espanhol como Língua Estrangeira. **Revista Bazar: Software e Conhecimento Livres**, 2008. Disponível em [http://repositorio.ufla.br/bitstream/1/9629/1/ARTIGO\\_Desenvolvimento%20de%20um%20chatbot%20para%20auxiliar%20o%20ensino%20de%20espanhol%20como%20l%C3%ADngua%20estrangeira.pdf](http://repositorio.ufla.br/bitstream/1/9629/1/ARTIGO_Desenvolvimento%20de%20um%20chatbot%20para%20auxiliar%20o%20ensino%20de%20espanhol%20como%20l%C3%ADngua%20estrangeira.pdf)> Acesso em: 12 fev 2019.

GOMES, D. S. Inteligência Artificial: conceitos e aplicações. **Revista Olhar Científico – Faculdades Associadas de Arquimedes** – V. 01, n.2, Ago-Dez. 2010. Disponível em <http://www.olharcientifico.kinghost.net/index.php/olhar/article/view/49/37>>. Acesso em: 25 Mai. 2019.

GONDA, D. E. et al. Evaluation of developing educational *chatbots* based on the seven principles for good teaching. In: **2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE).** IEEE, 2018. p. 446-453.

GRAELLS, P. M. **Los Medios Didácticos.** Departamento de Pedagogía Aplicada, Facultad de Educación, Universitat Autònoma de Barcelona, 2000. Disponível em <http://www.peremarques.net/medios2.htm>>

HALLIDAY; M.; MATTIESSEN, C. Halliday's Introduction to Functional Grammar. 4ª ed. London: Routledge, 2014.

INBOT. **Glossário de termos relacionados à Inteligência Artificial para Assistentes**

**Virtuais.** 2019. Disponível em: <<https://in.bot/glossario-de-inteligencia-artificial.php>>. Acesso em: 20 de dez. de 2019.

JOHNSON, K. **Why Bill Gates would like Prepathon's new education chatbots.** 2016. Disponível em <<https://venturebeat.com/2016/08/15/why-bill-gates-would-like-prepathons-education-chatbots/>> Acesso em 19 jun. 2018.

JOSEPH, S. **The Need for Real-time Contextual Awareness in Enterprise Chatbots.** 2018. Disponível em: <https://chatbotslife.com/the-need-for-real-time-contextual-awareness-in-enterprise-chatbots-e9bcaefa3b35>. Acesso em: 16 NOV 2019

KATCHAPAKIRIN, K; ANUTARIYA, C. An architectural design of scratchthai: A conversational agent for computational thinking development using scratch. In: **Proceedings of the 10th International Conference on Advances in Information Technology.** 2018. p. 1-7.

KOZINETS, R. V. **Netnografia: realizando pesquisa etnográfica online.** Porto Alegre: Penso, 2014.

KULIGOWSKA, K. **Commercial Chatbot: Performance Evaluation, Usability Metrics and Quality Standards of Embodied Conversational Agents.** 1-16. 2015. Disponível em: <https://www.researchgate.net/publication/282474589> Commercial Chatbot Performance Evaluation Usability Metrics and Quality Standards of Embodied Conversational Agents. Acesso em: 08 MAI 2019

LANKSHEAR, C. KNOBEL, M. **Pesquisa Pedagógica – do projeto à implementação.** Porto Alegre: Artmed, 2008, cap. 6 A ética na pesquisa.

LEON-HENRI, D.P. Pedagogical engineering and ESP: The ICD and active approach. **Journal of Teaching English for Specific and Academic Purposes**, v. 4, n. 2, p. 241-252, 2016.

LEONHARDT, M. D. **Doroty: um Chatterbot para Treinamento de Profissionais Atuantes no Gerenciamento de Redes de Computadores.** Dissertação, Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, 2005.

LIMA, L. A. **Estudo de implementação de um robô de conversação em Curso de língua estrangeira em ambiente virtual: um caso de estabilização do Sistema Adaptativo Complexo**. Tese de Doutorado. Belo Horizonte: Universidade Federal de Minas Gerais, 2014.

MAINGUENEAU, D. **Análise de textos de comunicação**. Tradução de Maria Cecília Silva e Décio Rocha. 6ª edição. São Paulo: Cortez, 2013.

MAKADIA, M. **5 Reasons Why Your Chatbot Needs Natural Language Processing**. 2019. Disponível em: <https://towardsdatascience.com/5-reasons-why-your-chatbot-needs-natural-language-processing-ed20fb0a3655>. Acesso em: 16 NOV 2019

MARTIN, J.; WHITE, P. R. R. **The language of evaluation: Appraisal in English**. New York, London, UK: Palgrave Macmillan, 2005.

MARINONI, G.; VAN'T LAND, H.; JENSEN, T. **The Impact of Covid-19 on Higher Education Around The World**. International Association of Universities Survey Report: Paris, France, 2020.

MEDEIROS, L.; JUNIOR, A.; MOSER, A. Uma Inteligência Artificial Ensinando sobre Inteligência Artificial: relato de experiência. In: **23º CIAED - Congresso Internacional ABED de Educação a Distância**, 2017, Curitiba. Anais do 23º CIAED, 2017. Disponível em <http://www.abed.org.br/congresso2017/trabalhos/pdf/395.pdf>. Acesso em: 05 Mar 2019.

MEYER, M. A. Didactics, sense making, and educational experience. **European Educational Research Journal**, v. 6, n. 2, p. 161-173, 2007.

MORAES, R. Análise de conteúdo. **Educação**, Porto Alegre, v. 22, n.37, p. 7-32, mar. 1999.

MORAES, T. C. **Resolução de correferência e categorias de entidades Nomeadas**. Dissertação. Faculdades de Informática, PUC-RS. Porto Alegre, 2010. Disponível em: <http://tede2.pucrs.br/tede2/bitstream/tede/5113/1/427028.pdf>. Acesso em: 16 NOV 2019.

MUSSALIM, F.; BENTES, A.C. **Introdução à linguística: domínios e fronteiras** Vol. 1. 8ª ed. São Paulo: Cortez, 2008.

NAVEGA, S. Inteligência Artificial, Educação de Crianças e o Cérebro Humano. **Leopoldianum: Revista de Estudos de Comunicações da Universidade de Santos**. Ano 25, n. 72, Fev. 2000, pp 87-102. Disponível em: [http://186.231.12.240/bitstream/uniepro/259/1/Sumario\\_tendencias\\_web.pdf](http://186.231.12.240/bitstream/uniepro/259/1/Sumario_tendencias_web.pdf). Acesso em: 22 Mai. 2019.

NEVES, A. M. **iAIML: Um Mecanismo para Tratamento de Intenção em Chatterbots**. Tese de Doutorado. Universidade Federal de Pernambuco. Departamento de Ciências da Computação, 2005. Disponível em [https://repositorio.ufpe.br/bitstream/123456789/2095/1/arquivo7155\\_1.pdf](https://repositorio.ufpe.br/bitstream/123456789/2095/1/arquivo7155_1.pdf). Acesso em 14 NOV 2019.

OLIVEIRA, M. E. **Formação ao Longo da Vida na Docência: narrativas que evidenciam agenciamentos do professor rumo à consolidação de competências profissionais**. Dissertação de Mestrado. Centro Federal de Educação Tecnológica de Minas Gerais. Departamento de Linguagem e Tecnologia, 2019.

ONU, "Pandemic disruption to learning is an opportunity to reimagine, revitalize education," Acesso em 04 JAN 2021. Acesso em : <https://news.un.org/en/story/2021/01/1082792>.

OSIEK, B. A. **Reconhecimento de sentimento em texto abordado através da computação afetiva**. Tese (doutorado). UFRJ, COPPE, Programa de Engenharia de Sistemas e Computação, 2014.

PAIVA, V. L. M.; NASCIMENTO, M. (Org.). **Sistemas adaptativos complexos: lingua(gem) e aprendizagem**. Belo Horizonte: Faculdade de Letras/FAPEMIG, 2009.

PAIVA, V. L. M. Reflexões sobre ética e pesquisa. **Rev. Brasileira de Lingüística Aplicada**, 44 v. 5, n. 1, p. 43-61, 2005.

PARDO, T. A. S. **Métodos para Análise Discursiva Automática**. Tese de Doutorado. Universidade de São Paulo. Instituto de Ciências Matemáticas e de Computação, 2005. Disponível em: <http://www.nilc.icmc.usp.br/dizer2/PDF/THESIS-Pardo.pdf>. Acesso em: 15 NOV 2019.

PANTELIMON, F-V et al. The Evolution of AI-Driven Educational Systems during the COVID-19 Pandemic. **Sustainability**, v. 13, n. 23, p. 13501, 2021. Acesso em 24 JAN 2022. Disponível em <https://www.mdpi.com/2071-1050/13/23/13501>

PÉREZ, J. Q., DARADOUMIS, T., PUIG, J. M. M. Rediscovering the use of *chatbots* in education: A systematic literature review. **Computer Applications in Engineering Education**, 2020. doi:10.1002/cae.22326

PORTAL DA AVALIAÇÃO. **Matriz de Referência**. Disponível em <http://www.portalavaliacao.caedufjf.net/pagina-exemplo/matriz-de-referencia/>. Acesso em 10 NOV 2019.

RADZIWILL, N. M.; BENTON, M. C. Evaluating Quality of *Chatbots* and Intelligent Conversational Agents. **arXiv preprint arXiv:1704.04579**, 2017. Disponível em <https://arxiv.org/abs/1704.04579>. Acesso em 10 NOV 2019.

RAHMAN, M. F. B. A.; SUGUNA, V. S. **Chatbots: Friends or Fiends?** 2017. (RSIS Commentaries, No. 182). RSIS Commentaries. Singapore: Nanyang Technological University.

RANGEL, E. O. Avaliar para melhor usar – Avaliação e seleção de materiais e livros didáticos. In: **Materiais Didáticos**: escolha e uso. Boletim 14, MEC, 2005.

RANJAN, R. et al. Adopting A New Hybrid Force Model: A Survey During Covid-19 In Indian Higher Education. International. **Journal of Emerging Technologies in Learning (iJET)**, v. 16, n. 16, p. 169-185, 2021. Acesso em 26 JAN 2022. Disponível em <https://www.learntechlib.org/p/219992/>

RAUPP, F. M.; BEUREN, I. M. **Metodologia da pesquisa aplicável às ciências sociais**. In: BEUREN, I. M. (Org.). **Como elaborar trabalhos monográficos em Contabilidade: teoria e prática**. São Paulo: Atlas, 2003. p. 76-97.

REN, R. et al. Evaluation techniques for *chatbot* usability: A systematic mapping study. **International Journal of Software Engineering and Knowledge Engineering**, v. 29, n. 11n12, p. 1673-1702, 2019

REPOLÊS, M. C. P. **Coleções didáticas de língua inglesa/PNLD: a apropriação do material didático por professores e alunos de escolas públicas mineiras**. Tese (Doutorado). Centro Federal de Educação Tecnológica de Minas Gerais, Programa de Pós-Graduação em Estudo de Linguagens, Belo Horizonte, 2019.

RICHTER, M. G. O Material Didático no Ensino de Línguas. **Linguagens & Cidadania**, SANTA MARIA - RS, v. 14, n.1, p. 1-19, 2006.

RIGA, M.. Com pandemia, WhatsApp ganha status de instrumento pedagógico no Brasil. **Forbes**, 2021. Acesso em 08 JAN 2022. Disponível em: <https://forbes.com.br/forbes-tech/2021/08/com-pandemia-whatsapp-ganha-status-de-instrumento-pedagogico-no-brasil/>

SANTOS, Z. B. A linguística sistêmico-funcional: algumas considerações. **SOLETRAS**, n. 28, p. 164-181, 2014.

SCHMIDT, M. A. R.; DELAZARI, L. S. Proposta de metodologia para testes qualitativos de mapas 3D. **Bol. Ciênc. Geod., sec. Artigos**, v. 16, n. 4, p. 575-588, 2010.

SILVA, A. et. al. Descoberta de Conhecimento Através de Métodos de Aprendizagem de Máquina Supervisionados Aplicados ao Ambiente Virtual de Aprendizagem Sigaa/Ufpi. In: **23º CIAED - Congresso Internacional ABED de Educação a Distância**, 2017, Curitiba. Anais do 23º CIAED, 2017. Disponível em <http://www.abed.org.br/congresso2017/trabalhos/pdf/191.pdf>>. Acesso em: 14 Mar 2019.

SILVA, E. J. O Design SPRINT como ferramenta para engajamento da equipe: um estudo de

caso. **Human Factors in Design**, v. 7, n. 13, p. 191-202, 2018. Disponível em <<http://www.revistas.udesc.br/index.php/hfd/article/view/2316796307132018191/8179>>.

Acesso em: 09 Fev. 2020.

SILVA, L. B. **Ambigüidades da língua portuguesa: recorte Classificatório para a elaboração de um modelo ontológico**. 2006. 135 f. Dissertação (Mestrado em Biblioteconomia e Documentação) - Universidade de Brasília, Brasília, 2006.

SILVA, R. C. A pertinência da utilização do sistema de avaliatividade como referência em pesquisas sobre recepção de livros didáticos: reflexões teóricas e metodológicas. **Letras**, Santa Maria, v. 25, n. 50, p. 359-382, jan./jun. 2015.

SILVA, R. C.; PARREIRAS, V.; FERNANDES, G. Avaliação e escolha de livros didáticos de inglês a partir do PNLD: uma proposta para guiar a análise. **Linguagem & Ensino**, Pelotas, v.18, n.2, 2015, p.355-377.

SILVA, T. P. **Desenvolvimento de um sistema do tipo Chatbot para o curso de Sistemas de Informação**. Monografia. Graduação em Sistemas de Informação, Universidade Federal de Santa Catarina, 2021.

SIMON, R. M. et al. **Adaptação como mídia para o conhecimento: uma análise de ambientes virtuais de aprendizagem utilizados em disciplinas de graduação**. 2017. Tese de Doutorado. Universidade Federal de Santa Catarina.

SMUTNY, P.; SCHREIBEROVA, P. *Chatbots for learning: A review of educational chatbots for the Facebook Messenger*. **Computers & Education**, 2020 103862. doi:10.1016/j.compedu.2020.103862.

SWEIDAN, S. Z. et al. SIAAA-C: A student interactive assistant android application with chatbot during COVID-19 pandemic. **Computer Applications in Engineering Education**, v. 29, n. 6, p. 1718-1742, 2021. Acesso em 24 JAN 2022. Disponível em [https://onlinelibrary.wiley.com/doi/full/10.1002/cae.22419?casa\\_token=LHb5zLO9jQgAAA%3ALTRfTjQfkC0qSFtf7P10sr-CaBaMTmO9Y4soFxpZPf8RuGUZcc-K2lY55kSRxTRz7NqEjwh7fa2FjWfei](https://onlinelibrary.wiley.com/doi/full/10.1002/cae.22419?casa_token=LHb5zLO9jQgAAA%3ALTRfTjQfkC0qSFtf7P10sr-CaBaMTmO9Y4soFxpZPf8RuGUZcc-K2lY55kSRxTRz7NqEjwh7fa2FjWfei)

URQUIZA, M.; MARQUES, D. Análise de conteúdo em termos de Bardin aplicada à comunicação corporativa sob o signo de uma abordagem teórico-empírica. **Entretextos**, v. 16, n. 1, p. 115-144, 2016.

TOPAL, A.; EREN, C.; GEÇER, A. *Chatbot* application in a 5th grade science course. **Education and Information Technologies**, v. 26, n. 5, p. 6241-6265, 2021. Acesso em 18 JAN 2022. Disponível em <https://link.springer.com/content/pdf/10.1007/s10639-021-10627-8.pdf>

TROJAN, R. M.; RODRÍGUEZ, J. Os PCN's e os Materiais Didáticos para o Ensino da Arte: o que propõem? **LINHAS**, Florianópolis, v. 9, n. 1, p. 49-71, jan. / jun. 2008.

TOMLINSON, B. (ed.), **Materials Development in Language Teaching**. Cambridge, New York: Cambridge University Press, 1998 (9a reimpressão, 2006) – Glossary, p.viii- xiv

TOMLINSON, B.; MASUHARA, H. **Developing Language Course Materials**. Singapura: RELC Portfolio Series, 2004.

VICARI, R. M. **Tendências em inteligência artificial na educação no período de 2017 a 2030**. Brasília: SENAI, 2018. Disponível em <[http://186.231.12.240/bitstream/uniepro/259/1/Sumario\\_tendencias\\_web.pdf](http://186.231.12.240/bitstream/uniepro/259/1/Sumario_tendencias_web.pdf)>. Acesso em: 10 Mai 2019.

VIEIRA, R.; LOPES, L. Processamento de Linguagem Natural e o Tratamento Computacional de Linguagens Científicas. 2010. In: PERNA, C. L. DELGADO, H.; FINATTO, M. J (Orgs). **Linguagens especializadas em corpora: modos de dizer e interfaces de pesquisa**. Porto Alegre: EDIPUCRS, 2010.

VIEIRA, R.; LIMA, V. L. S. Linguística computacional: princípios e aplicações. In: Ana Teresa Martins; Díbio Leandro Borges. (Org.). **As Tecnologias da informação e a questão social**. Fortaleza: SBC, 2001, v. 3, p. 47-88. Disponível em: <http://www.inf.unioeste.br/~jorge/MESTRADOS/LETRAS%20->

[%20MECANISMOS%20DO%20FUNCIONAMENTO%20DA%20LINGUAGEM%20%20PROCESSAMENTO%20DA%20LINGUAGEM%20NATURAL/ARTIGOS%20INTERESSANTES/lingu%edstica%20computacional.pdf](#). Acesso em: 15 NOV 2019.

WEIZENBAUM, J. ELIZA - a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v. 9, n. 1, p. 36-45, 1966.

XAVIER, N.B; NETTO, JFM; LIMA, DPR. Relato de um Estudo de Caso: Analisando as Interações dos Alunos por meio de Sociogramas. In: **21º CIAED - Congresso Internacional ABED de Educação a Distância**, 2015, Bento Gonçalves. Anais do 21º CIAED, 2017. Disponível em [http://www.abed.org.br/congresso2015/anais/pdf/BD\\_137.pdf](http://www.abed.org.br/congresso2015/anais/pdf/BD_137.pdf). Acesso em: 19 Abr. 2019.

YIN, R. K. **Estudo de caso: planejamento e métodos**. Trad.: Daniel Grassi - 2.ed. -Porto Alegre: Bookman, 2001.

ZHU, P. et al. Lingke: A fine-grained multi-turn *chatbot* for customer service. **arXiv preprint arXiv:1808.03430**, 2018.

## APÊNDICES

### Apêndice 1

## TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Projeto CAAE: \_\_\_\_\_, aprovado pelo Sistema CEP/CONEP, em  
\_\_\_\_\_ de \_\_\_\_\_ de 20 \_\_\_\_.

Prezado(a) \_\_\_\_\_,

Você está sendo convidado(a) a participar da pesquisa intitulada: **CRITÉRIOS PARA PRODUÇÃO E AVALIAÇÃO DE CHATBOTS EDUCACIONAIS**. Este convite se deve ao fato de você ser aluno do curso *Teaching and Learning in Higher Education* e ter acesso ao *chatbot* utilizado como suporte ao curso, o que é critério de inclusão para a participação na pesquisa.

O pesquisador responsável pela pesquisa é Leonardo Drummond Vilaça Lima Camargo, RG 10231743, doutorando no programa de Estudos de Linguagens do CEFET-MG. A pesquisa refere-se a análise das contribuições de um *chatbot* para o processo de aprendizagem e o entendimento de quais critérios podem auxiliar a produção e avaliação de *chatbots* usados como material didático. O objetivo é **criar uma matriz de critérios e descritores para produção e avaliação de chatbots utilizados como materiais didáticos**. O presente projeto de pesquisa justifica-se pela necessidade de se investigar a produção, avaliação e o uso de uma tecnologia digital que tem alcançado destaque recentemente: os *chatbots*, ou agentes de conversação, que em contextos educacionais podem ser entendidos como um tipo de sistema de aprendizagem adaptativo.

Estão previstas as seguintes formas de participação por parte dos informantes/participantes na investigação que correspondem a etapas de coleta/geração de *dados da pesquisa (2020)*: a) aplicação do *Think Aloud Test*, em que o participante irá interagir com o *chatbot*, falando em voz alta (haverá gravação em áudio da voz do participante) o que está tentando encontrar no sistema, quais os resultados encontrados e qual a sua opinião sobre esses resultados (haverá gravação em vídeo da tela do computador utilizado no teste); b) análise

das interações dos participantes com o *chatbot* usado como material didático no curso *Teaching and Learning in Higher Education*, ofertado pelo Instituto Ânima e a Universidade de Tampere (Finlândia); c) Aplicação de questionário; d) Entrevista (gravada em áudio).

Todas as formas de participação apontadas são voluntárias. Além disso, você não será identificado quando o material de seu registro for utilizado, sempre com propósitos de publicação científica ou educativa. Saiba, portanto, que, em hipótese alguma, haverá identificação de qualquer informante da pesquisa na divulgação de seus resultados.

Espera-se que você contribua com este trabalho fornecendo informações sobre como você utilizou o *chatbot* durante o curso *Teaching and Learning in Higher Education*, como foram as suas interações com o *chatbot* e quais as suas opiniões sobre o sistema.

Faz-se necessário informar também que o risco que esta pesquisa apresenta ao participante diz respeito ao possível desconforto e constrangimento que pode surgir ao se levar em consideração a pessoalidade dos dados coletados, além do cansaço físico e mental em função do tempo dedicado às entrevistas. Entretanto, o pesquisador, a fim de minimizar esses riscos, assume o compromisso de ser o mais breve possível, atentando-se exclusivamente aos objetivos da pesquisa. Como benefício, o participante contribuirá, a partir dos seus relatos e da análise de suas interações com o sistema, para o debate a respeito do uso dessa tecnologia (os *chatbots*) para fins educacionais.

Como participante de uma pesquisa e de acordo com a legislação brasileira, você é portador de diversos direitos, além do anonimato, da confidencialidade, do sigilo e da privacidade, mesmo após o término ou interrupção da pesquisa. Assim, lhe é garantido:

- A observância das práticas determinadas pela legislação aplicável, incluindo as Resoluções 466 (e, em especial, seu item IV.3) e 510 do Conselho Nacional de Saúde, que disciplinam a ética em pesquisa e este Termo;
- A plena liberdade para decidir sobre sua participação sem prejuízo ou represália alguma, de qualquer natureza;
- A plena liberdade de retirar seu consentimento, em qualquer fase da pesquisa, sem prejuízo ou represália alguma, de qualquer natureza. Nesse caso, os dados colhidos

- de sua participação até o momento da retirada do consentimento serão descartados a menos que você autorize explicitamente o contrário;
- O acompanhamento e a assistência, mesmo que posteriores ao encerramento ou interrupção da pesquisa, de forma gratuita, integral e imediata, pelo tempo necessário, sempre que requerido e relacionado a sua participação na pesquisa, mediante solicitação ao pesquisador responsável;
- O acesso aos resultados da pesquisa;
- O ressarcimento de qualquer despesa relativa à participação na pesquisa (por exemplo, custo de locomoção até o local combinado para a entrevista), inclusive de eventual acompanhante, mediante solicitação ao pesquisador responsável;
- A indenização diante de eventuais danos decorrentes da pesquisa;
- O acesso a este Termo. Este documento é rubricado e assinado por você e por um pesquisador da equipe de pesquisa, em duas vias, sendo que uma via ficará em sua propriedade. Se perder a sua via, poderá ainda solicitar uma cópia do documento ao pesquisador responsável.

Qualquer dúvida ou necessidade – nesse momento, no decorrer da sua participação ou após o encerramento ou eventual interrupção da pesquisa – pode ser dirigida ao pesquisador, por e-mail: [leonardodrummondvilaca@gmail.com](mailto:leonardodrummondvilaca@gmail.com), telefone (31)988443932, pessoalmente ou via postal para Rua Santa Bárbara, 635, ap 32 – Bairro Sagrada Família, Belo Horizonte-MG. CEP 31030150.

Se preferir, ou em caso de reclamação ou denúncia de descumprimento de qualquer aspecto ético relacionado à pesquisa, você poderá recorrer ao Comitê de Ética em Pesquisa (CEP) do Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), vinculado à CONEP (Comissão Nacional de Ética em Pesquisa), comissões colegiadas, que têm a atribuição legal de defender os direitos e interesses dos participantes de pesquisa em sua integridade e dignidade, e para contribuir com o desenvolvimento das pesquisas dentro dos padrões éticos. Você poderá acessar a página do CEP, disponível em: <<http://www.cep.cefetmg.br>> ou contatá-lo pelo endereço: Av. Amazonas, n. 5855 - Campus VI; E-mail: [cep@cefetmg.br](mailto:cep@cefetmg.br); Telefone: +55 (31) 3379- 3004 ou presencialmente, no horário de atendimento ao público: às terças-feiras: 12:00 às 16:00 horas e quintas-feiras: 07:30 às 12:30 horas.

Se optar por participar da pesquisa, peço-lhe que rubrique todas as páginas deste Termo, identifique-se e assine a declaração a seguir, que também deve ser rubricada e assinada pelo pesquisador.

DECLARAÇÃO

Eu, \_\_\_\_\_,

abaixo assinado, de forma livre e esclarecida, declaro que aceito participar da pesquisa como estabelecido neste TERMO.

Assinatura do participante da pesquisa: \_\_\_\_\_

Assinatura do pesquisador: \_\_\_\_\_

Belo Horizonte, \_\_\_\_\_ de \_\_\_\_\_ de 20\_\_

Se quiser receber os resultados da pesquisa, indique seu e-mail ou, se preferir, endereço postal, \_\_\_\_\_ no \_\_\_\_\_ espaço \_\_\_\_\_ a seguir: \_\_\_\_\_

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

## Apêndice 2

### Roteiro do *Think Aloud Test*

Prezado participante, durante este teste de usabilidade, você deve utilizar o *chatbot* do curso Teaching and Learning in Higher Education, procurando as informações que deseja encontrar e fazendo perguntas para as quais deseja obter respostas que possam te ajudar a estudar. Por favor, fale em voz alta o que estiver pensando durante o teste. Conte o que pretende fazer, o que está buscando no *chatbot*, o que estiver escrevendo, onde está clicando e quais resultados encontrou e qual é a sua opinião sobre os resultados da sua interação. Em alguns momentos, você vai interagir livremente e em outros momentos pediremos que você busque informações específicas. Além disso, poderemos te fazer algumas perguntas durante o teste. Tanto a tela do computador quanto o seu áudio serão gravados para coleta de dados desta pesquisa.

#### Tarefas e perguntas:

- 1 – Leia a apresentação e as opções de menu iniciais. Qual é a sua impressão inicial sobre a apresentação do *chatbot* e as primeiras opções que aparecem na tela?
- 2 – Procure o calendário do curso.
- 3 – Procure informações sobre o projeto pedagógico que você vai realizar no curso.
- 4 – Procure um conteúdo do curso sobre o qual você gostaria de obter mais informações.
- 5 – O que você achou da imagem e personalidade do *chatbot*?
- 6 – As opções que aparecem como botões no *chatbot* são temas que te interessam?
- 7 – O *chatbot* conseguiu entender as perguntas que você fez?
- 8 – O *chatbot* conseguiu te responder adequadamente?
- 9 – Há algum tipo de informação ou conteúdo do curso que você gostaria que estivesse no menu ou que o *chatbot* soubesse te responder e que você não conseguiu encontrar?
- 10 – Como você formularia textualmente as perguntas que gostaria de fazer ao *chatbot*?
- 11- Este *chatbot* te auxiliaria nos seus estudos?
- 12 – Como ele poderia te auxiliar melhor nos seus estudos?

### Apêndice 3

#### Questionário para alunos

##### Perfil 1 – alunos que interagiram com o *chatbot* em diferentes etapas do curso

- 1 – A Katri (*chatbot*) contribuiu para os seus estudos? Se sim, como?
- 2 – Você se lembra quais os tipos de informações que você buscou com a Katri? Quais conteúdos você solicitou com mais frequência?
- 3 – A Katri conseguia entender bem o que você perguntava ou buscava? Por favor, explique como você se lembra que a Katri respondia/retornava informações.
- 4 – Como você avalia a suficiência/quantidade de informações que você encontrava?
- 5 - O uso do *chatbot* te estimulou a estudar outros conteúdos do curso? Se sim, dê um ou mais exemplos que você se recorda de ter estudado após conversar com a Katri.
- 6 – Como você avalia a linguagem utilizada pela Katri (formalidade ou informalidade, quantidade de texto etc)?
- 7- O quão fácil foi entender o que a Katri respondia?
- 8 – Você sentia que estava participando de uma experiência individualizada de aprendizagem ou que todos os alunos tinham acesso ao mesmo conteúdo? O que te fez entender isso?
- 9 – Qual é a sua opinião sobre a imagem e personalidade da Katri? 10 – O que você achou do *layout* da página em que funciona a Katri?
- 11 – Qual é a sua opinião sobre as opções de temas que aparecem no menu inicial? Eles são os que mais te interessam num primeiro acesso ou deveriam ser trocados por outros temas?
- 12 – Você preferia clicar nos botões do *chatbot* ou escrever livremente para interagir com ele? Justifique.
- 13 – Há algum tipo de informação que a Katri deveria saber como te responder e que você não conseguiu encontrar? Qual/quais?
- 14 – Quais são os principais pontos positivos do uso desse *chatbot* no curso? 15 – O que você gostaria que você diferente na Katri?

**Perfil 2 – alunos que não interagiram com o *chatbot***

- 1 - Qual o motivo de você não ter usado o *chatbot* disponível no curso?
- 2 - Você acredita que um *chatbot* pode facilitar um processo de aprendizagem? Se sim, como? Se não, por quê?
- 3 - Em que situações você usaria um *chatbot* como aluno de um curso?
- 4 - Você usaria um *chatbot* como apoio à sua prática docente? Como um *chatbot* poderia te auxiliar?
- 5 - Você acredita que os seus alunos gostariam de usar um *chatbot* no apoio aos estudos? Por que você tem essa opinião?

**Perfil 3 – Alunos que interagiram com o *chatbot* apenas no início do curso e não voltaram a interagir.**

- 1- Qual o motivo de você ter usado o *chatbot* no início do curso e não tem usado mais?
- 2- Você encontrou informações relevantes quando usou o *chatbot*? Explique como foi a experiência.
- 3- A Katri (*chatbot*) contribuiu para os seus estudos? Se sim, como?
- 4- Você se lembra quais os tipos de informações que você buscou com a Katri? Quais foram?
- 5- A Katri conseguia entender bem o que você perguntava ou buscava? Por favor, explique o que acontecia na sua busca, como você se lembra que a Katri respondia.
- 6- Como você avalia a suficiência/quantidade de informações que você encontrava?
- 7- Você sentia que estava participando de uma experiência individualizada de aprendizagem ou que todos os alunos tinham acesso ao mesmo conteúdo? O que te fez entender isso?
- 8- O uso do *chatbot* te estimulou a estudar outros conteúdos do curso? Se sim, dê um ou mais exemplos que você se recorda de ter estudado após conversar com a Katri.

- 9- Como você avalia a linguagem utilizada pela Katri (formalidade ou informalidade, quantidade de texto etc)?
- 10- O quão fácil foi entender o que a Katri respondia?
- 11- Qual é a sua opinião sobre a imagem e personalidade da Katri?
- 12- O que você achou do layout da página em que funciona a Katri?
- 13- Qual é a sua opinião sobre as opções de temas que aparecem no menu inicial? Eles são os que mais te interessam num primeiro acesso ou deveriam ser trocados por outros temas?
- 14- Você preferia clicar nos botões do *chatbot* ou escrever livremente para interagir com ele? Justifique.
- 15- Há algum tipo de informação que a Katri deveria saber como te responder e que você não conseguiu encontrar? Qual/quais?
- 16- Quais são os principais pontos positivos do uso desse *chatbot* no curso?
- 17- O que você gostaria que fosse diferente na Katri?

#### **Perfil 4 – Alunos que só interagiram com o *chatbot* na etapa final do curso**

- 1- Qual o motivo de você ter usado o *chatbot* somente no final do curso?
- 2- Você acredita que teria usado mais esse recurso após conhecer a Katri? O que te faz ter essa impressão?
- 3- Você encontrou informações relevantes quando usou o *chatbot*? Explique como foi a experiência.
- 4- A Katri contribuiu para os seus estudos? Se sim, como?
- 5- Você se lembra quais os tipos de informações que você buscou com a Katri? Quais foram?
- 6- A Katri conseguia entender bem o que você perguntava ou buscava? Por favor, explique o que acontecia na sua busca, como você se lembra que a Katri respondia.
- 7- Você sentia que estava participando de uma experiência individualizada de aprendizagem ou que todos os alunos tinham acesso ao mesmo conteúdo? O que te fez entender isso?
- 8- Como você avalia a suficiência/quantidade de informações que você encontrava?
- 9- O uso do *chatbot* te estimulou a estudar outros conteúdos do curso? Se sim, dê um ou mais exemplos que você se recorda de ter estudado após conversar com a Katri.

- 10- Como você avalia a linguagem utilizada pela Katri (formalidade ou informalidade, quantidade de texto etc)?
- 11- O quão fácil foi entender o que a Katri respondia?
- 12- Qual é a sua opinião sobre a imagem e personalidade da Katri?
- 13- O que você achou do layout da página em que funciona a Katri?
- 14- Qual é a sua opinião sobre as opções de temas que aparecem no menu inicial? Eles são os que mais te interessam num primeiro acesso ou deveriam ser trocados por outros temas?
- 15- Você preferia clicar nos botões do *chatbot* ou escrever livremente para interagir com ele? Justifique.
- 16- Há algum tipo de informação que a Katri deveria saber como te responder e que você não conseguiu encontrar? Qual/quais?
- 17- Quais são os principais pontos positivos do uso desse *chatbot* no curso?
- 18- O que você gostaria que fosse diferente na Katri?

## Apêndice 4

### **Roteiro de entrevista semiestruturada com produtores do *chatbot***

- 1 - Quais as especificidades desse *chatbot*?
- 2 - Há diferença entre produzir um *chatbot* educacional e outros tipos de *chatbots*? O que foi diferente do processo de criação desse *chatbot* em relação a outros *chatbots*?
- 3 - Quais as principais características linguísticas que o diferem de outros *chatbots*?
- 4 - Quais as vantagens e desvantagens de se trabalhar esse *chatbot* com expressões em português e inglês?
- 5 - Como vocês avaliam o número de interações e engajamento dos usuários com o *chatbot*?
- 6 - Como foi feita a análise das interações? O que era considerado para decidir alterações ou complementação de conteúdos?
- 7 - Algo na linguagem do *chatbot* mudou ao longo do processo? Se sim, o quê?
- 8 - Que tipos de interações eram esperadas e foram verificadas no uso do *chatbot*? Quais tipos de interações surpreenderam a equipe?
- 9 - O nível de confiabilidade das respostas apresentadas nos relatórios automáticos do *chatbot* correspondem ao nível de confiabilidade que vocês podem aferir analisando as perguntas e respostas?
- 10 - Após as interações dos usuários, a equipe se mantém satisfeita com a escolha da persona e a maneira como ela interage e redige as respostas?
- 11 - O que vocês aprimorariam ou mudariam no *chatbot* após analisar as interações já realizadas?