

Universidade Federal de Campina Grande - UFCG
Centro de Engenharia Elétrica e Informática - CEEI
Coordenação de Pós-Graduação em Ciência da Computação

Arcabouço Semiautomático para o Apoio à Avaliação da Participação em Fóruns de EaD

Danielle Chaves de Medeiros

Joseana Macêdo Fachine Régis de Araújo, DSc.

José Eustáquio Rangel de Queiroz, DSc.

(Orientadores)

Campina Grande, Paraíba, Brasil
Dezembro – 2014

Universidade Federal de Campina Grande - UFCG
Centro de Engenharia Elétrica e Informática - CEEI
Coordenação de Pós-Graduação em Ciência da Computação

**Arcabouço Semiautomático para o Apoio à
Avaliação da Participação em Fóruns de EaD**

Danielle Chaves de Medeiros

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande - Campus I como parte dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Joseana Macêdo Fechine Régis de Araújo, DSc.

José Eustáquio Rangel de Queiroz, DSc.

(Orientadores)

Área de Concentração: Ciência da Computação
Linha de Pesquisa: Modelos Computacionais e Cognitivos

Campina Grande, Paraíba, Brasil
Dezembro – 2014

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

M488a Medeiros, Danielle Chaves de.
 Arcabouço semiautomático para o apoio à avaliação da participação em fóruns de EAD / Danielle Chaves de Medeiros. – Campina Grande, 2014.
 131 f. : il. color.

 Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática, 2014.

 "Orientação: Prof. Dr. Joseana Macêdo Fechine Régis de Araújo, Prof. Dr. José Eustáquio Rangel de Queiroz".

 Referências.

 1. Educação à Distância – Ciência da Computação. 2. Avaliação da Participação. 3. Fóruns de Discussão. 4. Modelagem do Conhecimento. 5. Análise de Similaridade entre textos. I. Araújo, Joseana Macêdo Fechine Régis de. II. Queiroz, José Eustáquio Rangel de. III. Título.

CDU 004:37.018.43(043)

"Pessoas científicas são sempre curiosas e eu serei um científico."
(Collin Craven, em *O Jardim Secreto*, por Frances Hodgson Burnett)

*Dedico esta pesquisa a minha família, pelo
amor, compreensão e incentivo dados
ao longo de toda minha vida.*

Agradecimentos

A Deus, por ser minha fonte de força inabalável ao longo de toda minha caminhada e motivo de conforto e luz em todos os momentos.

A meus pais, Gilberto e Maria de Fátima, pela certeza que sempre tiveram do meu sucesso e por todo o esforço que dedicaram para nos proporcionar uma boa educação. A minha irmã, Gisele, por ser meu exemplo de dedicação e pelas conversas sempre cheias de encorajamento. A meu irmão, Antonio Eduardo, por ser um companheiro de estudos (quase) sempre paciente e por me oferecer palavras de apoio sempre quando mais preciso delas.

Aos professores Joseana Fachine e, Eustáquio Rangel, por me oferecerem sua amizade e orientação ao longo destes anos e por terem me dado a oportunidade e a confiança necessárias para realizar esta pesquisa.

Aos colegas do Laboratório de Inteligência Artificial (LIA), pelas palavras de incentivo e conhecimento compartilhado. Em especial, a Alana Marques, por dividir comigo as responsabilidades e atividades de pesquisa e por estar sempre disposta a oferecer ajuda e discutir ideias.

A meus amigos Jéssica Fontes, Elizabel Aluska, Tony Herbert, Mayara Nóbrega, Naja Queiroz, Adrielle Rosa, Stéfani Brito, assim como a todos aqueles que não foram aqui mencionados, pelo constante apoio e por me proporcionarem sorrisos sempre quando necessários.

Ao Instituto Federal da Paraíba (IFPB), especialmente aos professores Francisco Alves, Marcos Vinícius e Adriana Costeira, pela confiança e boa vontade em me ajudar a obter os dados para minha pesquisa.

Aos docentes do IFPB Virtual, pela colaboração, em especial, ao amigo e tutor Alexandre Sales, por toda paciência e apoio.

A todos os funcionários da COPIN, especialmente a Rebeka e a Vera, pela atenção e boa vontade com as quais sempre me atenderam.

À CAPES, pelo apoio financeiro.

A todos os professores de EaD que, voluntariamente, responderam o questionário de pesquisa administrado.

Por fim, a todos aqueles que, direta ou indiretamente, tornaram possível a conclusão desta pesquisa.

Resumo

A seleção de critérios para a análise de informações durante o processo de avaliação da participação em fóruns de discussão de cursos de Educação a Distância (EaD) é um grande desafio. São muitas as variáveis que devem ser consideradas neste processo, além da subjetividade inerente à análise realizada pelo docente, passível de erro humano. Os docentes geralmente não possuem a seu dispor todos os recursos necessários, se tornando necessário o uso de uma metodologia ou ferramenta que o auxilie no processo de avaliação. Diante desta demanda e, a partir de um estudo dos principais indicadores qualitativos/quantitativos utilizados pelos professores de EaD, foi desenvolvido um arcabouço para a análise da participação dos alunos em fóruns. O objetivo deste arcabouço é servir de apoio à tomada de decisão do professor, fornecendo um mecanismo mais efetivo para a mensuração da quantidade e da qualidade das interações, passível de adaptação à metodologia tradicional adotada por cada docente. A validação deste arcabouço deu-se a partir da administração de questionários para a sondagem da opinião de docentes atuantes na área de ensino a distância, assim como pela realização de estudos de caso envolvendo a avaliação da acurácia de instâncias do arcabouço para o cálculo da nota de participação de alunos. Foi desenvolvido um Sistema Especialista (SE) para o processamento dos dados, com o uso de funções de similaridade para realizar, de forma semiautomática, a avaliação do conteúdo das mensagens dos alunos. Assim, as notas de participação calculadas foram confrontadas com as notas atribuídas pelo docente utilizando a abordagem tradicional. Os resultados obtidos demonstraram que, em três das cinco turmas observadas, não foi possível verificar a existência de diferenças estatísticas significativas entre o desempenho das abordagens estudadas. Um estudo da acurácia e correlação revela que, em todos os casos analisados, há uma forte relação entre os dados e o erro médio encontrado foi inferior a 3%, demonstrando a aplicabilidade do arcabouço ao contexto da avaliação da participação em fóruns.

Palavras-chaves: educação a distância, avaliação da participação, fóruns de discussão, modelagem do conhecimento, análise de similaridade entre textos

Abstract

The selection of criteria for the information analysis during the participation evaluation process, in discussion forums of distance courses, is a major challenge. There are many variables to consider in this process, in addition to the subjectivity inherent in the analysis carried out by the instructor, which is subject to human error. Instructors, generally, do not have at their disposal all the resources necessary, thus, the use of methodologies or tools that can help them with this process are necessary. Facing this demand, and after performing a study of the major qualitative/quantitative indicators used by distance education teachers, we developed a framework for the analysis of the student participation in the forums. The aim of this framework is to support the decision-making process, by providing a more effective mechanism to measure the quantity and quality of interactions, capable of adjusting itself to the traditional methodology adopted by each teacher. The validation of this framework was performed by the administration of questionnaires that surveyed the opinion of active teachers in the distance learning area, and by the execution of case studies involving the assessment of the accuracy of instances of this framework for calculating the participation grade of students. The study involved the development of an Expert System, for the treatment and processing of the data, using similarity functions to perform, semi-automatically, the assessment of the content of the students' messages. Thus, it was possible to confront the calculated participation grades with the grades assigned by the teacher using the traditional approach. The results showed that, in three out of the five classes observed, it was not possible to verify the existence of statistically significant differences between the performance of both the approaches studied. A study of the accuracy and correlation shows that, in all the cases analyzed, there is a strong relationship between the data, and the average error was less than 3%, demonstrating the applicability of the proposed framework to the assessment of student participation in forums.

Keywords: *distance learning, participation assessment, discussion forums, knowledge modelling, text similarity analysis*

Lista de Abreviações

AVA	Ambientes Virtuais de Aprendizagem
ARS	Análise de Redes Sociais
CoI	<i>Community of Inquiry</i>
EaD	Educação a Distância
ELVT	<i>Educationally Less Valuable Talk</i>
EQM	Erro Quadrático Médio
EVT	<i>Educationally Valuable Talk</i>
IFPB	Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
IFPE	Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco
IFTM	Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro
LMS	<i>Learning Management System</i>
MEC	Ministério da Educação e Cultura
PaI	<i>Participation Indicator</i>
PeI	<i>Performance Indicator</i>
PLN	Processamento de Linguagem Natural
RB	Redes Bayesianas
RBC	Raciocínio Baseado em Casos
RBR	Raciocínio Baseado em Regras
SCAFFOLD	<i>Scale For Forums/Online Discussion Assessment</i>
SE	Sistemas Especialistas
UAB	Universidade Aberta do Brasil
UEPB	Universidade Estadual da Paraíba
UERJ	Universidade Estadual do Rio de Janeiro
UFC	Universidade Federal do Ceará
UFF	Universidade Federal Fluminense

UFMA	Universidade Federal do Maranhão
UFMG	Universidade Federal de Minas Gerais
UFPB	Universidade Federal da Paraíba
UFPE	Universidade Federal de Pernambuco
UFRPE	Universidade Federal Rural de Pernambuco
UFSC	Universidade Federal de Santa Catarina
UMESP	Universidade Metodista de São Paulo
UNIFEI	Universidade Federal de Itajubá
UNIP	Universidade Paulista
UNOPAR	Universidade Norte do Paraná

Lista de Figuras

Figura 2.1 – Arquitetura básica de um Sistema Especialista	31
Figura 3.1 – Perfil relativo à área de atuação e tempo (em anos) de experiência dos docentes consultados em diversas instituições brasileiras	39
Figura 3.2 – Recursos utilizados pelos professores consultados para a avaliação da participação do aluno em fóruns de discussão	39
Figura 3.3 – Opinião dos docentes a respeito da utilidade do uso de fóruns de discussão para a avaliação da participação	40
Figura 3.4 – Resultados da sondagem de opinião de docentes de EaD a respeito de alguns fatores de influência na avaliação da participação de alunos em fóruns	42
Figura 3.5 – Resultados da sondagem de opinião de docentes de EaD a respeito de alguns indicadores de participação para a avaliação de alunos em fóruns	42
Figura 3.6 – Resultados da sondagem de opinião de docentes de EaD sobre algumas formas de mensuração dos indicadores de participação de quantidade e qualidade das mensagens	43
Figura 3.7 – Modelo conceitual para avaliação da participação de alunos em fóruns de discussão	45
Figura 3.8 – Arcabouço para a avaliação da participação do aluno em fóruns de discussão em cursos de EaD	49
Figura 4.1 – Importância média (em %) atribuída pelos respondentes para cada medida proposta para a implementação dos indicadores de participação incorporados no arcabouço	65
Figura 4.2 – Técnicas de Avaliação das mensagens enviadas ao fórum utilizadas pelos docentes consultados	65

Figura 4.3 – Gráfico de dispersão entre o conjunto de dados obtidos pelo uso da abordagem proposta e tradicional, para o Estudo de Caso 1	76
Figura 4.4 – Gráfico de dispersão entre o conjunto de dados obtidos pelo uso da abordagem proposta e tradicional, para o Estudo de Caso 2	76
Figura 4.5 – Gráfico de dispersão entre o conjunto de dados obtidos pelo uso da abordagem proposta e tradicional, para o Estudo de Caso 3	77

Lista de Quadros

Quadro 2.1 – Níveis conceituais da participação de alunos em cursos online e suas respectivas unidades de análise (Hrastinski, 2008)	13
Quadro 2.2 – Sumarização dos indicadores de participação utilizados pelas pesquisas estudadas, de acordo com a divisão proposta por Hrastinski (2008)	15
Quadro 2.3 – Exemplos de categorias da ferramenta SCAFFOLD	19
Quadro 3.1 – Sumarização da abordagem utilizada para a validação do arcabouço proposto	54
Quadro 4.1 – Fatores do modelo conceitual considerados pelos docentes para a avaliação da participação dos alunos nos fóruns estudados	70

Lista de Tabelas

Tabela 2.1 – Exemplos do uso de frequência de palavras-chaves, utilizados por Wu e Chen (2005)	22
Tabela 3.1 – Sumarização da base de dados disponibilizada pelo IFPB Virtual, após filtragem dos dados	53
Tabela 4.1 – Sumarização dos dados de delineamento do perfil dos docentes consultados	59
Tabela 4.2 – Utilização dos fatores de avaliação empregados pelos docentes consultados	61
Tabela 4.3 – Importância média (em %) atribuída pelos respondentes para cada medida proposta para a implementação dos indicadores de participação incorporados no arcabouço	64
Tabela 4.4 – Uso de técnicas, pelos docentes consultados, para a avaliação das mensagens em fóruns de discussão	65
Tabela 4.5 – Sumarização de dados referentes a cada estudo de caso utilizado para a avaliação da acurácia do arcabouço proposto	68
Tabela 4.6 – Resultados dos testes não-paramétrico Wilcoxon-Mann-Whitney para os três Estudos de Caso (nível de significância $\alpha = 0,05$)	78
Tabela 4.7 – Resultados para o teste de correlação entre as notas simuladas e as atribuídas pelo docente, para cada estudo de caso analisado	79
Tabela 4.8 – Resultados para o cálculo do EQM entre as notas simuladas e atribuídas pelo docente, para cada estudo de caso analisado	81

Sumário

Capítulo 1 – Considerações Iniciais	1
1.1 Contextualização do Problema	1
1.2 Objetivos.....	5
1.2.1 Geral	5
1.2.2 Específicos.....	5
1.3 Relevância	6
1.4 Estrutura da Dissertação.....	7
Capítulo 2 – Fundamentação Teórica	9
2.1 Participação do Aluno em EaD	9
2.2 Avaliação da Participação em Fóruns de Discussão	11
2.2.1 Indicadores de Participação.....	14
I Acesso ao Ambiente	16
II Leitura das Mensagens.....	16
III Percepção do Aluno	17
IV Qualidade das Mensagens	17
V Quantidade das Mensagens	22
VI Tamanho das Mensagens	23
VII Tempo Gasto	23
VIII Outros Indicadores.....	24
2.2.2 Abordagens para Mensuração da Participação Encontradas na Literatura	26
2.3 Sistemas Especialistas (SE).....	29
2.3.1 Arquitetura de um Sistema Especialista	30
2.3.2 Classificação de Sistemas Especialistas	31
2.4 Análise da Semelhança entre Dados Textuais.....	32

2.4.1	Pré-processamento de Dados Textuais.....	33
2.4.2	Similaridade Entre Textos	34
2.5	Considerações sobre o Conteúdo do Capítulo	34
Capítulo 3 – Arcabouço para a Modelagem da Avaliação da Participação em Fóruns de Discussão		36
3.1	Descrição da Abordagem Proposta	36
3.1.1	Investigação da Opinião da Comunidade Acadêmica	37
3.1.2	Modelo Conceitual	43
	I Presença Social	44
	II Qualidade das Interações.....	47
3.1.3	Arcabouço Proposto	48
3.2	Abordagem Proposta para a Validação do Arcabouço.....	51
3.3	Considerações sobre o Conteúdo do Capítulo	55
Capítulo 4 – Validação do Arcabouço		56
4.1	Investigação sobre o Perfil e Opinião do Docente.....	56
4.1.1	Questões de Pesquisa	57
4.1.2	Metodologia	57
4.1.3	Participantes.....	58
4.1.4	Execução.....	58
4.1.5	Resultados e Discussão	58
	I Delineamento do Perfil do Docente	58
	II Sondagem da Opinião do Docente.....	60
4.2	Avaliação da Acurácia do Arcabouço Proposto	67
4.2.1	Questões de Pesquisa	67
4.2.2	Metodologia	67

I	Seleção dos Estudos de Caso e Realização de Pesquisas de Campo	63
II	Elaboração de Instâncias do Arcabouço Proposto	69
III	Processamento e Análise dos Dados	73
4.2.3	Participantes	74
4.2.4	Execução	75
4.2.5	Resultados e Discussão	75
I.	Análise Estatística dos Dados	75
II.	Estudo do Erro e Correlação entre os Dados	79
4.3	Limitações da Abordagem Proposta	81
4.4	Considerações sobre o Conteúdo do Capítulo	84
Capítulo 5 – Considerações Finais e Sugestões para Pesquisas Futuras		85
5.1	Considerações Finais	85
5.2	Contribuições da Pesquisa	87
5.3	Sugestões para Pesquisas Futuras	88
Referências Bibliográficas		91
Apêndice A – Questionário sobre Técnicas para Avaliação da Participação de Alunos em EaD		102
Apêndice B – Documentação da Parceria entre LIA-UFCG e IFPB Virtual		110
Apêndice C – Questionário sobre Avaliação da Participação de Alunos em Fóruns de EaD no IFPB Virtual		113
Apêndice D – Formulário de Entrevista com Docentes		120
Apêndice E – Análise de Funções de Similaridade para Verificação do Conteúdo de Mensagens em Fóruns de Discussão		122

Capítulo 1

Considerações Iniciais

Neste capítulo, é apresentada uma visão geral sobre a abordagem proposta nesta dissertação. Na Seção 1.1 deste capítulo, é contextualizado o problema abordado nesta pesquisa e descrita brevemente a solução proposta para sua resolução. Em seguida, na Seção 1.2, são apresentados os objetivos desta pesquisa. Na Seção 1.3, descreve-se a relevância da pesquisa. Por fim, na Seção 1.4, apresenta-se a estrutura dos demais capítulos da dissertação.

1.1 Contextualização do Problema

De acordo com o Censo Brasileiro de Educação Superior (CENSUP, 2013), no período 2012-2013, as matrículas em cursos de Educação a Distância (EaD) possuíram uma participação superior a 18% nas matrículas em cursos de graduação, o que totalizou mais de um milhão, dentre os mais de sete milhões de estudantes do ensino superior. Em comparação com o número de alunos matriculados em cursos a distância em 2001, que não passava de 5.359 alunos (CENSUP, 2001), é notável o crescimento desta modalidade de educação.

O termo *e-learning* é utilizado para designar qualquer modelo de ensino não presencial que faz uso de tecnologia, utilizando Ambientes Virtuais de Aprendizagem (AVA), tais como o Moodle (DOUGIAMAS, 2001) e o Blackboard (BLACKBOARD INC., 1997). Tais ambientes auxiliam na comunicação e interação entre os participantes do curso, oferecendo ferramentas que tornem a aprendizagem a distância o mais próximo possível do ensino presencial.

Com a popularização da EaD e o uso crescente destes ambientes, viu-se também crescer o número de pesquisas dedicadas ao estudo de diferentes

aspectos pedagógicos nesta modalidade de ensino, e.g., a análise do processo de ensino do conteúdo e de avaliação do aluno (GHOSH; KLEINBERG, 2013). As pesquisas cujo foco está na investigação de fatores relacionados à participação dos alunos ao longo do curso, buscam estudar e propor métodos e mecanismos destinados a auxiliar o professor na tarefa de análise e avaliação das interações realizadas pelos alunos. Para isto, podem ser utilizados dados provenientes de algumas das funcionalidades oferecidas pela maioria dos AVA, como as ferramentas de discussão síncronas e assíncronas. Estas ferramentas possibilitam o intercâmbio de informações entre os estudantes e docentes, a partir do compartilhamento de ideias e do conhecimento individual.

Ferramentas síncronas, tais como *chats*, oferecem meios para a elucidação de dúvidas, de forma rápida, a partir do intercâmbio de mensagens instantâneas entre professor e aluno. As ferramentas assíncronas, por sua vez, permitem que os alunos desenvolvam habilidades cognitivas e sociais (BIRCH; VOLKOV, 2005; GREENLAND, 2011), promovendo, durante um período de tempo, o intercâmbio de informações e a construção de conhecimento por meio da discussão entre todos os alunos, intermediada pelo docente. A extração de dados gerados a partir do uso destas ferramentas é usualmente utilizada em pesquisas na área de EaD. Dados estatísticos obtidos a partir das interações entre alunos em AVA são utilizados, por exemplo, para a inferência de correlação entre esses dados e o resultado da avaliação final dos estudantes (BEER *et al.*, 2009; MACFADYEN; DAWSON 2010).

Dentre os meios de comunicação assíncronos, um dos mais populares é o *fórum de discussão* (RAGHAVAN *et al.*, 2010). A utilização de fóruns é uma forma eficiente de promover o compartilhamento de ideias, pois os participantes podem postar mensagens a respeito de um determinado tema proposto (por eles próprios ou, na maioria dos casos, pelo docente), ler as mensagens enviadas pelos colegas e/ou deixar comentários sobre o tema em discussão, além de obter comentários sobre seu progresso na discussão (RAGHAVAN *et al.*, 2010). Assim sendo, esta funcionalidade vem se tornando um elemento-chave para o

compartilhamento de experiências e para a construção do conhecimento, fatores muito importantes em cursos não presenciais (CALVANI *et al.*, 2009). A partir da análise de interações dessa natureza, o professor pode, além de ter uma boa noção da evolução do aprendizado de sua turma, verificar também a colaboração individual de cada aluno e avaliar sua participação ao longo do curso.

No entanto, embora seja uma forma prática e efetiva para a verificação da participação do aluno, o processo de avaliação pode envolver diferentes formas e fatores de análise. Alguns exemplos são: (i) a análise de indicadores numéricos; (ii) a análise do conteúdo, visando verificar a qualidade das mensagens; (iii) o relacionamento entre a leitura do material da disciplina e a redação das mensagens; e (iv) a percepção do próprio aluno acerca da discussão (AGUILAR *et al.*, 2013).

Um dos fatores mais importantes ao considerar a avaliação da participação do aluno em fóruns é a possibilidade de que um número maior de alunos ingressantes¹, em um determinado curso a distância, pode gerar um volume considerável de mensagens a serem avaliadas pelo professor (ANDRESEN, 2009). Essa atividade pode demandar grande dispêndio de tempo e, até mesmo, se tornar uma tarefa impraticável, a destacar também a componente de erro humano passível de introdução no processo de avaliação.

Os AVA comumente utilizados ainda não oferecem informação suficiente para auxiliar efetivamente o professor nesta tarefa. A forma mais comum de verificar a participação dos alunos nesta modalidade de ensino é a utilização de medidas quantitativas baseadas em estatísticas (MAZZOLINI; MADDISON, 2007). Medidas tais como a *quantidade de mensagens enviadas* pelo aluno, o *número total de tópicos iniciados* por ele e a *quantidade de acessos* ao ambiente de aprendizagem podem ser extraídas utilizando-se recursos do AVA. Mesmo assim, a utilização de dados dessa natureza não é, muitas vezes, suficiente para avaliar a participação dos alunos, já que os dados estatísticos levam em consideração apenas a quantidade e não a qualidade das interações (ERLIN *et al.*, 2008).

¹ Em EaD, um mesmo curso pode possuir centenas ou até milhares de alunos matriculados ao mesmo tempo.

Um dos grandes obstáculos para o auxílio à automatização da verificação do que está sendo discutido nos fóruns é a subjetividade existente na avaliação feita pelo professor. O docente é capaz de verificar, a partir da leitura de cada mensagem, se o conteúdo é relevante ao tema, se as informações trazidas apresentam alguma novidade em relação ao que já foi discutido, se seu conteúdo foi extraído integralmente da *Internet* ou se se trata, de fato, de opiniões formadas a partir do conhecimento do aluno. Utilizando esta abordagem de análise tradicional, o docente se depara com a dificuldade de interpretação e avaliação do conhecimento e da qualidade da participação dos alunos (ROMERO *et al.*, 2013).

No intuito de automatizar tal processo de investigação realizado pelo professor, levando-se em consideração a subjetividade presente em todo o procedimento de avaliação, a aplicação de técnicas de Inteligência Artificial pode fornecer meios alternativos e que considerem aspectos mais relevantes. Alguns exemplos de aspectos que podem ser considerados são: (i) o perfil do aluno, (ii) a verificação automática do conteúdo das mensagens ou (iii) uma avaliação mais aprofundada do histórico de mensagens intercambiadas pelos alunos.

Nesta pesquisa, pretendeu-se, pois, identificar e adaptar indicadores de participação que refletissem uma forma efetiva de avaliar a participação de alunos em fóruns de EaD, com a finalidade de construir um arcabouço para tal fim. De forma a verificar a importância dos indicadores que compõem o arcabouço, foram realizadas pesquisas de campo junto a professores de EaD.

A acurácia do arcabouço foi, ainda, validada a partir de estudos de caso, realizados por meio do emprego de dados provenientes de cursos a distância de uma instituição de ensino técnico. Para isso, os docentes responsáveis pela avaliação da participação dos alunos das turmas selecionadas foram entrevistados, a fim de determinar qual o processo de avaliação empregado por eles ao atribuir a nota de participação dos alunos. Assim, de posse deste conhecimento, foram criadas instâncias do arcabouço para cada caso investigado. Com o auxílio de um Sistema Especialista (SE) destinado à análise e à avaliação

das interações, representadas por *logs* de acesso e mensagens trocadas nos fóruns, foi possível investigar quão eficaz era a solução proposta, ao confrontar os resultados obtidos pelo uso do arcabouço e a abordagem tradicional utilizada pelos docentes.

A partir da sondagem de opinião dos docentes de EaD, foi possível corroborar os resultados obtidos por pesquisadores da área (HRASTINSKI, 2009; HRASTINSKI, 2008), o que atesta a relevância dos indicadores utilizados pelo modelo conceitual e a estrutura que o compõe. O segundo estudo envolveu a análise da acurácia do arcabouço proposto. Nas cinco turmas analisadas, os resultados gerados mostraram que, em três delas, não foram registradas diferenças estatisticamente significativas entre a avaliação realizada por meio do uso do arcabouço e a avaliação tradicional feita pelo docente. Para todos os casos estudados, o erro médio calculado foi menor que 3%. Tal resultado é evidenciado pela forte correlação existente, para todas as turmas analisadas, entre as notas calculadas a partir de instâncias do arcabouço e as notas atribuídas pelo docente.

1.2 Objetivos

Nesta seção, serão enunciados os objetivos geral e específicos da pesquisa ora descrita.

1.2.1 Geral

O objetivo geral desta pesquisa foi utilizar indicadores de participação adotados por professores durante a avaliação da participação dos alunos, em fóruns de discussão, para a elaboração de um arcabouço para modelagem da participação de cada aluno, buscando diminuir a subjetividade proveniente do processo de avaliação convencional.

1.2.2 Específicos

A fim de atingir o objetivo geral supraexplicitado, os seguintes objetivos

específicos fizeram-se necessários:

1. Verificar os indicadores utilizados no processo de avaliação da participação de alunos em fóruns, a partir de uma revisão da literatura e da realização de um estudo de sondagem de opinião, junto a docentes de cursos de EaD;
2. A partir da literatura revisada da área e do resultado da sondagem dos indicadores adotados por docentes, definir um conjunto de indicadores para compor o arcabouço de avaliação da participação;
3. Avaliar a acurácia do arcabouço definido, a partir do uso de um SE, utilizando técnicas para análise semântica das mensagens e de dados extraídos de cursos de EaD; e
4. Confrontar os resultados obtidos por meio do uso do arcabouço e da avaliação realizada da forma convencional pelos docentes de cada disciplina analisada nos experimentos.

1.3 Relevância

Várias abordagens metodológicas podem ser encontradas na literatura, visando à análise da participação de alunos em cursos *online*. Uma parte dessas pesquisas tem como objetivo prover ao docente uma forma de verificar, de forma mais eficiente, o andamento da discussão, apontando-lhe os principais tópicos em discussão e lhe fornecendo uma visão geral de todas as interações realizadas entre os alunos (ERLIN *et al.*, 2009; RABBANY *et al.*, 2011; JYOTHI *et al.*, 2012).

Há também as pesquisas cujo foco está na utilização e/ou análise de indicadores de participação, de forma a fornecer ao professor meios que lhe dêem suporte no processo de avaliação da participação dos alunos. Os autores dessas pesquisas se empenham, de forma direta ou indireta, em identificar indicadores que representem a participação do aluno de forma adequada (BLISS e LAWRENCE, 2009, ELIA; CHAMOVITZ, 2009, LIN *et al.*, 2009, ZORRILLA *et al.*, 2010, ANAYA; BOTICÁRIO, 2011, CABALLÉ *et al.*, 2011, WEE; ABRIZAH, 2011, GOMES *et al.*, 2012, OLIVEIRA; ESMIN, 2012, ROMERO *et al.*, 2013).

No entanto, ainda é notória a dificuldade de se encontrar, na literatura, técnicas para avaliar especificamente a qualidade das interações entre estudantes em fóruns de discussão. Ainda há uma carência de metodologias que auxiliem o docente no processo de avaliação dessas interações. Assim, este tema de pesquisa é relevante por objetivar uma forma de modelar o processo de avaliação realizado pelos professores, resultando em um arcabouço que seja facilmente adaptável às necessidades do docente, que o auxilie na tomada de decisões relativas à atribuição de notas de participação aos alunos e que lhe permita explorar de forma mais efetiva os recursos fornecidos pelo AVA utilizado.

Embora já existam pesquisas com propostas similares, conforme mencionado anteriormente, grande parte utiliza técnicas de análise social na visualização das interações ou indicadores unicamente quantitativos na definição do valor atribuído à participação. Outras proposições focalizam apenas o estudo do relacionamento dos indicadores com a participação dos alunos em fóruns. Na literatura revisada, não foram identificadas abordagens que visassem à padronização deste processo de avaliação.

As abordagens estudadas que fornecem uma metodologia para o cálculo de notas de participação dos alunos consideram apenas a extração de fatores quantitativos que mensuram o total de contribuições dos alunos. A verificação das mensagens, nessas pesquisas, foram realizadas de forma manual pelo docente (ROMERO *et al.*, 2013, WEE; ABRIZAH, 2011) ou de forma totalmente automática, sem nenhuma orientação ou envolvimento do docente neste processo (WU; CHEN, 2005).

1.4 Estrutura da Dissertação

Esta dissertação está organizada em 5 (cinco) capítulos. O primeiro capítulo, já apresentado, contém uma contextualização do problema abordado nesta pesquisa, uma discussão da relevância de sua resolução e os objetivos geral e específicos da pesquisa.

No Capítulo 2, são descritas diferentes abordagens para a investigação da participação de alunos em fóruns de discussão. Neste capítulo são definidos indicadores de participação nelas adotados e as diferentes formas de obtenção dessas medidas. Adicionalmente, são apresentados conceitos relevantes de IA utilizados para a avaliação da abordagem proposta nesta pesquisa, tais como Sistemas Especialistas (SE) e análise de similaridade entre dados textuais.

No Capítulo 3, é descrita a solução proposta, apresentando-se detalhes de seu processo de concepção. No Capítulo 4, por sua vez, apresentam-se e discutem-se a metodologia adotada no processo de validação do arcabouço proposto e os resultados obtidos nos experimentos conduzidos para tal validação.

Por fim, no Capítulo 5, formulam-se algumas considerações finais, assim como se descrevem as principais contribuições desta pesquisa e se propõem algumas sugestões para pesquisas futuras.

Capítulo 2

Fundamentação Teórica

Neste capítulo, apresentam-se, para o melhor entendimento do tema abordado, alguns conceitos básicos relacionados à avaliação da participação do aluno no contexto da Educação a Distância (Seção 2.1), complementados por uma discussão das diferentes abordagens revisadas, destinadas à avaliação da participação em fóruns de cursos *online* (Seção 2.2). Na Seção 2.3, discutem-se, para melhor entendimento da solução descrita nesta dissertação (ver Capítulo 3), conceitos relacionados à aplicação e ao desenvolvimento de Sistemas Especialistas (SE). Na Seção 2.4, são apresentados conceitos básicos relativos à verificação de similaridade entre textos, importantes para o pré-processamento dos dados durante o processo de validação do arcabouço proposto na presente pesquisa. Por fim, na Seção 2.5, são apresentadas algumas considerações finais sobre o capítulo.

2.1 Participação do Aluno em EaD

A expressão "Educação a Distância" é geralmente associada à transmissão de conhecimento a partir de sistemas de informação que facilitam o desenvolvimento do processo de ensino-aprendizagem. EaD, segundo Moore (1990), é uma relação existente entre o diálogo, a estrutura e a autonomia, que requer meios para mediatizar esta comunicação. Peters (1983) afirmou que a EaD fornece meios de acesso aos estudos universitários a um grande número de estudantes independentemente de seu lugar de residência e de ocupação. Estes dois pesquisadores fazem parte de um conjunto de pensadores que tentam definir a EaD por meio de uma perspectiva pedagógica, econômica e tecnológica.

Nos últimos 20 anos, a questão relativa às características básicas desta

modalidade de ensino têm sido estudadas por vários pesquisadores (HARRY *et al.*, 2013). Dentre estas características, uma das mais estudadas é a *interação*. Em cursos *online*, conforme afirmou Peters (1983), é desejável que exista um alto índice de interações, de modo a aumentar a produtividade e troca de conhecimento, devido a distância física entre professores e alunos. Moore (1989) e Hrastinski (2009) identificaram, com uma década de espaçamento em suas pesquisas, 3 (três) tipos de interações que definem a EaD: (i) a interação entre o aluno e o conteúdo, relacionada à habilidade que os estudantes possuem para acessar, manipular, sintetizar e comunicar informação, com base no tópico em discussão; (ii) a interação entre o aluno e o professor, geralmente envolvendo o fornecimento de *feedback*; e (iii) interações entre os alunos, que envolvem o intercâmbio de informações e experiências entre eles. Ao final de sua pesquisa, Hrastinski (2009) concluiu que, na prática, nenhum dos 3 (três) modos de interação funciona de forma independente. Na verdade, a participação dos alunos em cursos de EaD envolve todos os tipos de interações supracitados.

Hrastinski (2009) adotou a definição de *participação* em comunidades dada por Wenger (1998), afirmando que a participação é um processo complexo que inclui fatores como a conversação, a ação, o raciocínio e o emocional do aluno. Ele teorizou, a partir de uma análise da literatura e, posteriormente, da identificação de fatores comportamentais (HRASTINSKI; JALDEMARK, 2012), que a participação é um fator fundamental em cursos *online*, promovendo a aprendizagem. Reiterando as conclusões de Hrastinski, Jaldemark *et al.* (2006) afirmaram que, apesar da existência de diferentes perspectivas acerca da participação na EaD, todas as perspectivas apontam para uma conclusão consonante: a participação é uma condição para a aprendizagem.

Em Educação, todo processo de avaliação requer o uso de metodologias e técnicas de validação que forneçam informações suficientes para auxiliar o docente na tomada de decisão. Em EaD, a relação entre a participação individual dos alunos, as interações realizadas por eles e a avaliação elaborada pelo professor são bastante complexas (YUKSELTURK, 2010). As interações sucedidas

ao longo do curso geram um conjunto de dados a serem analisados, sintetizados e transformados em resultados. Se for realizada de forma eficiente, a avaliação dessas interações pode auxiliar o docente a compreender melhor as necessidades da turma, tornando o foco do ensino centralizado no aluno (BEEBE *et al.*, 2010).

No Brasil, a documentação utilizada como referência, que deve oferecer orientação para elaboração de cursos de nível superior em EaD, é denominada "Referências de Qualidade para Educação Superior a Distância", divulgada pela Secretaria da Educação a Distância (MEC, 2007). O objetivo desse documento é guiar as instituições de ensino com recomendações sobre o planejamento de cursos de EaD, de acordo com o padrão esperado pelo Ministério da Educação (LACHI *et al.*, 2006). No entanto, não consta no referido documento nenhuma informação sobre como a verificação da participação dos alunos deve ser conduzida.

Assim, cada instituição é responsável por definir o padrão de avaliação a ser seguido. Tal atividade fica, geralmente, a critério do coordenador de cada curso ou, na maioria dos casos, do professor de cada disciplina. Dentre o conjunto de fatores que possibilitam a observação da troca de informações e interações entre os discentes, segundo Alstete e Beutell (2004), o fator de maior impacto na avaliação da participação do aluno em cursos a distância é o uso de *fóruns de discussão*. Desta forma, os fóruns são a ferramenta mais utilizadas pelos docentes para a verificação da participação do aluno nesta modalidade de ensino (RAGHAVAN *et al.*, 2010).

2.2 Avaliação da Participação em Fóruns de Discussão

O fórum de discussão é uma ferramenta que permite o desenvolvimento de habilidades cognitivas, a partir da composição de mensagens que expressam as opiniões e os pensamentos dos participantes sobre o tema da discussão. Em sua pesquisa, Zorrilla *et al.* (2010) constataram que os alunos consultados consideravam o fórum uma ferramenta útil e os dados extraídos de seu estudo

confirmam que o fórum é a ferramenta de comunicação mais utilizada por eles.

Ao fazer uso do fórum de discussão, os alunos podem compartilhar informações que reflitam seu conhecimento, lendo, respondendo, compartilhando ideias, de forma a aprofundar seu próprio entendimento do assunto (SHAW, 2012). A utilização de fóruns para investigar a capacidade de percepção e reflexão dos alunos permite explorar a participação individual efetiva na discussão (AZEVEDO *et al.*, 2012).

Ao avaliar a participação do aluno utilizando dados extraídos desses ambientes, é necessário dispor de um método claro e que represente bem as interações realizadas. Não obstante, a avaliação é geralmente conduzida a partir de medidas quantitativas (MAZZOLINI; MADDISON, 2007). Todavia, avaliar a participação de um aluno em fóruns transcende a verificação de estatísticas (VONDERWELL; ZACARIAH, 2005).

Wee e Abrizah (2011) descreveram 2 (duas) abordagens para realizar a avaliação da participação de alunos em discussões *online*. A primeira fundamenta-se na visão holística do processo de avaliação, segundo a qual as mensagens enviadas pelos alunos são avaliadas pelo docente de uma forma geral, sendo atribuído um único valor que represente sua participação, a partir de determinados critérios de avaliação pré-estabelecidos. Alguns exemplos são o aprofundamento do conteúdo, o envio de comentários a outros alunos, a regularidade nas postagens e a liderança. A visão analítica, por sua vez, utiliza uma abordagem multidimensional, por meio de métodos de alocação de pontos, em que o aluno é avaliado segundo diferentes critérios, que totalizados, refletem o valor final de sua participação. A vantagem desta abordagem é o fato de serem levadas em consideração diferentes informações a respeito do progresso do aluno ao longo da discussão. Do ponto de vista analítico, o aluno pode ser avaliado de acordo com a adoção de medidas tanto qualitativas como quantitativas, tais como o número, o tamanho e conteúdo das mensagens enviadas e o tempo gasto na discussão, dentre outros.

Alguns desses indicadores podem ser facilmente observados e quantificados,

a partir do uso de *logs* de acesso à plataforma e de envio de mensagens aos fóruns de discussão. Hrastinski (2008) apontou evidências que corroboram esta afirmação. O autor descreveu como "conceitos de baixo nível" definições de participação *online* que se referem à avaliação como algo simples, podendo facilmente ser mensurada a partir de fatores quantitativos, tais como o número de vezes em que o aluno acessou o AVA ou o número de mensagens que o estudante leu ou escreveu durante a discussão. As definições de "alto nível", para o autor, envolvem, além dos fatores associados aos "conceitos de baixo nível", definições que enfatizam a complexidade envolvida no fenômeno da participação (HRASTINSKI, 2009).

O autor propôs uma divisão destes conceitos em seis níveis, de acordo com o que ele denominou de *unidades de análise*, utilizadas pela literatura estudada, apresentadas no Quadro 2.1. É importante ressaltar que esta denominação pode variar, de acordo com os grupos de pesquisadores, sendo chamadas, por exemplo: *critérios de avaliação* (e.g., ELIA; CHAMOVITZ, 2009, GOMES *et al.*, 2012, NANDI *et al.*, 2012) ou *indicadores de participação* (e.g., NISTOR; NEUBAUER, 2010, WEE; ABRIZAH, 2011, AZEVEDO *et al.*, 2012, CACCIAMANI *et al.*, 2012, WISE *et al.*, 2012, ROMERO *et al.*, 2013, TOBARRA *et al.* 2014).

Quadro 2.1 Níveis conceituais da participação de alunos em cursos *online* e suas respectivas unidades de análise (Hrastinski, 2008).

Unidade de Análise	Níveis					
	1	2	3	4	5	6
Acesso ao Ambiente	•	•	•	•	•	•
Leitura das Mensagens				•		•
Percepção do Aluno			•		•	•
Qualidade das Mensagens			•	•	•	•
Quantidade de Mensagens		•	•	•	•	•
Tamanho das Mensagens		•	•	•	•	•
Tempo Gasto				•	•	•

No entanto, a utilização de uma abordagem que utilize diversos indicadores, ao tratar simultaneamente de diferentes aspectos da participação do aluno, pode demandar muito tempo do docente e acabar por omitir a visão geral das ideias intercambiadas na discussão. Assim, ferramentas de análise podem ser utilizadas como forma de verificar estes fatores de forma automática, a fim de reduzir o tempo e o esforço exigidos na interpretação (SCHRIRE, 2003).

Vários pesquisadores se empenharam em investigar como esses indicadores afetam a participação do aluno, utilizando diferentes abordagens. O ponto em comum entre estas pesquisas é o fato de que todas utilizam, de forma direta ou indireta, métricas de avaliação para estudar e/ou calcular as interações realizadas entre os alunos em fóruns de discussão.

2.2.1 Indicadores de Participação

De acordo com a literatura revisada, a participação pode ser detectada pelo uso de indicadores numéricos, que descrevem a iniciativa, a atividade, a assiduidade e ações resultantes das decisões tomadas pelo aluno (ANAYA; BOTICARIO, 2010). Contudo, também deve ser levada em consideração a análise qualitativa das mensagens, realizada pelo docente, e a percepção que os estudantes possuem da discussão (CACCIAMANI *et al.*, 2012).

O *indicador de participação*, apesar de se tratar de uma medida de natureza matemática, engloba variáveis relativas às características qualitativas e quantitativas das interações geradas durante a discussão. Tipicamente, os dados extraídos do AVA, podem ser processados, analisados e sumarizados para prover essas medidas (COLLAZOS *et al.*, 2007). Os indicadores de participação utilizados mais frequentemente nas pesquisas analisadas podem ser categorizados utilizando-se a estrutura de níveis descrita por Hrastinski (2008) (Quadro 2.1).

No Quadro 2.2, é apresentada uma classificação das pesquisas analisadas durante a revisão de literatura e os indicadores de participação utilizados na abordagem descrita por cada uma delas.

Quadro 2.2 Sumarização dos indicadores de participação utilizados pelas pesquisas estudadas, de acordo com a divisão proposta por Hrastinski (2008).

Autores	Indicadores de Participação							Técnica(s) Utilizada(s)
	Acesso ao Ambiente	Leitura das Mensagens	Percepção do Aluno	Qualidade das Mensagens	Quantidade das Mensagens	Tamanho das Mensagens	Tempo Gasto	
Anaya e Boticario (2011)					•			Clustering e Árvores de Decisão
Azevedo et al. (2011, 2012)				•	•			Sistema Especialista e Análise de Dados Textuais
Bliss e Lawrence (2009)				•	•			Contagem e Análise Estatística
Caballé et al. (2011)		•	•	•	•			Contagem e Análise Estatística
Cacciamani et al. (2012)				•	•			Contagem e Análise Estatística
Caspi et al. (2008)	•				•			Contagem e Análise Estatística
Chan e Chan (2011)		•			•			Contagem e Análise Estatística
Davies e Graff (2005)	•							Contagem e Análise Estatística
Dringus e Ellis (2004)				•				Sistema Especialista
Dringus e Ellis (2010)							•	Contagem e Análise Estatística
Ellis (2003)	•		•		•			Contagem e Análise Estatística
Gomes et al. (2012)	•				•		•	Contagem e Análise Estatística
Kay (2006)				•	•	•	•	Contagem e Análise Estatística
Lin et al. (2009)				•				Análise de Dados Textuais
Nistor e Neubauer (2010)					•	•		Contagem e Análise Estatística
Romero et al. (2013)		•		•	•	•	•	ARS e Clustering
Shaw et al. (2012)					•			Contagem e Análise Estatística
Tobarra et al. (2014)					•			Contagem e Análise Estatística
Wee e Abrizah (2011)				•	•	•		Contagem e Análise Estatística
Wise et al. (2012)							•	Contagem e Análise Estatística
Wu e Chen (2005)				•	•	•		Análise de Dados Textuais

Nas próximas subseções, será dada a definição de cada um dos indicadores listados, bem como serão listadas algumas formas de manipulação dessas medidas encontradas na literatura.

I. Acesso ao Ambiente

A *frequência de acesso* ao fórum é um indicador quantitativo, que pode ser verificado a partir de registros obtidos por meio de *logs* de acesso, que são disponibilizados pela maioria dos AVA adotados atualmente em cursos de EaD. A partir da análise da frequência de acesso à discussão, é possível ter a primeira impressão sobre a participação de um aluno no fórum.

Em algumas pesquisas (e.g., DAVIES; GRAFF, 2005, CASPI *et al.*, 2008), utiliza-se a média de *logins* realizadas pelos alunos como forma de investigar a relação entre participação de alunos por gênero, por exemplo. Davies e Graff (2005) utilizaram o acesso ao ambiente e às áreas de interação para representar a participação dos alunos.

Caspi *et al.* (2008) mensuraram a participação do aluno pela frequência com que ele acessava as discussões. Na ferramenta proposta por Gomes *et al.* (2012), um dos indicadores que também pode ser utilizado pelo professor ao configurar os requisitos para compor a participação dos alunos é o total de *logins*.

II. Leitura das Mensagens

A verificação da leitura das mensagens é uma medida importante no processo de mensuração da participação do aluno. Zhang *et al.* (2009) afirmaram que não é possível que um aluno participe efetivamente de uma discussão se ele não realizou a leitura do que foi discutido até então. Este indicador está usualmente associado à contagem de mensagens lidas pelo aluno.

A abordagem adotada por Caballé *et al.* (2011) considera o *feedback* dos alunos, de forma que, ao acessar o fórum, eles indiquem, se foi realizada a leitura da postagem. Outra forma de analisar a leitura da discussão é por meio da análise

do tempo gasto pelo aluno no fórum, que será mais bem detalhado na Seção VII. Romero *et al.* (2013) e Chan e Chan (2011) também utilizaram este indicador em suas pesquisas.

III. Percepção do Aluno

A percepção do aluno está relacionada a qualquer forma de avaliação realizada pelos alunos, de forma direta (e.g., atribuição de notas às mensagens dos demais alunos) ou indireta (e.g., por meio de questionários ou relatórios de experiências administrados pelo docente após o término da discussão).

A análise deste indicador pode ser feita de forma quantitativa ou qualitativa. O total de mensagens lidas, por exemplo, pode ser adotada como uma forma de mensurar a percepção dos alunos.

Uma forma mais direta, adotada por Caballé *et al.* (2011), consiste em solicitar dos alunos um valor para a utilidade das mensagens dos demais em relação ao tema proposto, tornando o aluno um participante ativo do processo de avaliação. Além disto, os alunos também podem ser consultados por meio de questionários ou relatórios sobre sua experiência no fórum de discussão (ELLIS, 2003).

IV. Qualidade das Mensagens

A análise das interações entre alunos deve envolver não apenas a verificação do "quanto" o aluno interagiu, mas também do "como" ele realizou tais interações. Essa investigação pode ser feita a partir da verificação do conteúdo das mensagens sob 2 (dois) pontos de vista: sintaticamente e/ou semanticamente. A análise semântica corresponde à verificação da adequação do conteúdo da mensagem ao tema da discussão. A sintática, por sua vez, é a observação de como o conteúdo foi redigido, e.g., frequência das palavras e sua ordenação, o vocabulário utilizado, as características gramaticais e estilísticas empregadas, entre outras (BAUER; GASKELL, 2002).

Este indicador é um dos mais complexos e envolve diversas metodologias e modelos cognitivos em sua concepção. Quanto mais clara for a definição do indicador de participação para avaliar a qualidade da interação, mais efetivo será o uso de uma ferramenta educacional que auxilie o docente a avaliar os alunos na discussão (NANDI *et al.*, 2009).

Duas das abordagens encontradas na literatura para tratar deste indicador estão relacionadas a atribuição de: (i) categorias às mensagens; e (ii) notas às mensagens.

A) Atribuição de Categorias às Mensagens

A classificação das mensagens em categorias pode auxiliar o professor a identificar melhor as necessidades dos alunos e ter uma visão geral das informações trocadas entre eles. Além disso, a categoria de uma mensagem também pode indicar o conteúdo presente nela, ajudando o docente também no processo de avaliação da qualidade das contribuições.

Em fóruns categorizados, as mensagens são classificadas de acordo com o assunto principal representado pelo seu conteúdo ou a intenção que se quer transmitir.

Na literatura revisada, este indicador foi analisado, por exemplo, (i) por meio da categorização realizada pelo próprio aluno ao compor a mensagem (CABALLÉ *et al.*, 2011); ou (ii) a partir da análise do conteúdo da mensagem, realizada pelo docente, a fim de que ele determinasse a qual categoria cada mensagem deveria pertencer (DRINGUS; ELLIS, 2004; WEE e ABRIZAH, 2011).

Dringus e Ellis (2004) apresentaram uma ferramenta para a categorização de mensagens, que descreve o objetivo geral das contribuições enviadas pelos alunos em fóruns, denominada *Scale for Forums/Online Discussion Assessment* (SCAFFOLD). A abordagem dos autores é fundamentada no modelo Analítico de Henri (HENRI, 1992), no modelo de Bloom (BLOOM *et al.*, 1956) e no modelo proposto por Garrison e Anderson (2003), perfazendo o total de 19 possíveis

categorias para as mensagens. No Quadro 2.3, são apresentados exemplos das categorias utilizadas pelo SCAFFOLD. Ao utilizar esta ferramenta, o docente deve categorizar ou descrever as mensagens, em uma escala entre 1 a 3, sendo 1 equivalente à "Não Evidente", 2 equivalente à "Mais ou Menos Evidente" e 3, "Fortemente Evidente". A ferramenta, então, realiza o processamento para representação numérica de atributos das mensagens de acordo com sua categorização.

Quadro 2.3 Exemplos de categorias da ferramenta SCAFFOLD.

Categoria	Descrição
Avaliativa	Avaliou as ideias compartilhadas em relação a sua validade e significância.
Esclarecimento	Forneceu ou buscou esclarecimento solicitado nas respostas.
Problema	Identificou um problema importante dentro do tema.
Resumo	Resumiu de forma geral o tema.
Social	Enviou mensagens de caráter conversacional ou de natureza social.
Questionamento	Levantou questionamentos relevantes ao tema.

Wee e Abrizah (2011) propuseram um modelo cuja análise da qualidade das mensagens foi realizada a partir de sua categoria e/ou descrição. Os professores consultados faziam uso do SCAFFOLD para classificar as mensagens. Segundo as autoras, tal esquema é bastante útil para descrever e categorizar mensagens em fóruns *online*, sendo de fácil uso e compreensão.

No entanto, as autoras argumentaram que, embora a utilização desta ferramenta possa prover informações úteis sobre os tipos de informação que são trocadas durante a discussão, um dos problemas no seu uso é a complexidade de implementação. Por ser um processo de verificação manual, o docente deve avaliar cada mensagem e seu julgamento pode conter inconsistências.

A abordagem utilizada por Caballé *et al.* (2011), por sua vez, resultou na integração de diversos modelos e métodos (CLARK; SCHAEFER, 1989, MARTIN,

1992, SELF, 1994). Os autores dividiram a classificação de mensagens em 3 (três) tipos: fornecimento de informação (*give-information exchange*), solicitação de informação (*elicit-information exchange*) e levantamento de problema (*raise-na-issue exchange*). A atribuição de uma destas 3 (três) categorias era realizada pelo docente durante a composição das mensagens. Oliveira *et al.* (2012), em um estudo inicial, utilizaram técnicas de mineração de dados para identificar um escopo menos abrangente, classificando as mensagens em 2 (duas) categorias: Dúvida e Esclarecimento.

Cacciamani *et al.* (2012) também estudaram o emprego da classificação de mensagens, baseando-se em um esquema proposto por Cesareni e Martini (2005). As categorias utilizadas foram: Propor Informação (e.g., contribuições que incluíam exemplos ou dados extraídos de fontes confiáveis), Elaborar Informação (e.g., síntese de ideias já propostas durante a discussão), Explorar (e.g., levantamento de questões ou problemas) e Avaliar (e.g., comentários ou reflexões metacognitivas).

Lin *et al.* (2009) apresentaram um modo automático para a categorização de mensagens, na qual estas eram classificadas, por meio de técnicas de mineração de dados e árvores de decisão, em: Aviso, Pergunta, Esclarecimento e Interpretação, Conflito, Afirmação e Outras.

B. Atribuição de Notas às Mensagens

Uma forma de mensurar o indicador de qualidade das mensagens é a partir da atribuição de notas às mensagens enviadas pelo aluno.

A metodologia utilizada por Azevedo *et al.* (2012) para a verificação de qualidade das postagens incluiu a análise de 3 (três) fatores principais: (i) a relevância temática da contribuição textual; (ii) a relevância de citações da mensagem; e (iii) a similaridade entre as mensagens postadas no fórum.

Romero *et al.* (2013) e Kay (2006) também utilizaram esta abordagem para a análise qualitativa das mensagens. Entretanto, estes autores fizeram uso de

uma metodologia mais objetiva, ao calcular a média das avaliações das mensagens atribuídas de forma tradicional pelo próprio docente.

C. Outras Abordagens para a Análise da Qualidade das Mensagens

Outra forma de verificação da qualidade das mensagens foi proposta por Bliss e Lawrence (2009), por meio do uso de 2 (dois) conceitos: EVT (*Educationally Valuable Talk*) e ELVT (*Educationally Less Valuable Talk*). Esses padrões foram inspirados na pesquisa de Uzuner (2007) e foram adotados para determinar, a partir da análise tradicional das mensagens realizada pelo docente, se a postagem possuía valor educacional ou não.

Wu e Chen (2005), por sua vez, partiram da premissa de que a qualidade das interações pode ser mensurada a partir do uso de conceitos-chaves (*key concepts*), representados por palavras-chaves. O número de palavras-chaves encontradas em uma mensagem reflete o quão abrangente é o conhecimento de seu autor. Para isso, Wu e Chen (2005) usaram um conjunto de palavras, descritas como "frases substantivas simples e não-recursivas", que podem conter outros modificadores, tais como adjetivos.

A seleção das palavras-chaves foi feita de forma automática, não-supervisionada, de acordo com a frequência de aparição na mensagem. A cada palavra-chave identificada, foi atribuído um peso, variável conforme a quantidade de vezes que a palavra-chave aparecia ao longo da discussão.

Na Tabela 2.1, são apresentados exemplos de resultados obtidos durante a etapa de validação do indicador de qualidade utilizado pelos autores. A palavra-chave *company*, por ter sido utilizada várias vezes e por todos os alunos do curso, possuiu seu peso nulo. É possível observar que, quanto menor a frequência de alunos utilizando uma determinada palavra-chave, maior o seu peso. Isso se dá porque, segundo os autores, mensagens contendo uma grande quantidade de termos gerais tendem a ter um conteúdo mais superficial e carecem de uma análise mais aprofundada e de argumentos fortes.

Tabela 2.1 Exemplos do uso de frequência de palavras-chaves, utilizadas por Wu e Chen (2005).

Palavra-chave	Peso Atribuído	Frequência	Número de Autores
<i>company</i>	0,00	212	31
<i>business decision</i>	5,81	1	1
<i>business structure change</i>	7,21	1	1
<i>business complexity</i>	11,86	3	3
<i>content management</i>	17,44	3	1

V. Quantidade das Mensagens

O total de postagens enviadas é o modo de mensuração de quantidade de mensagens mais utilizada para avaliar a participação dos alunos em fóruns (RYABOV, 2012). Ela pode indicar o nível de envolvimento do aluno na discussão, dado que, teoricamente, ela irá refletir o número de vezes que ele acessou ao fórum e contribuiu, de alguma forma, para a discussão.

De fato, a maioria das pesquisas analisadas fez uso do total de mensagens enviadas na descrição de suas abordagens (e.g., ROMERO *et al.*, 2013, GOMES *et al.*, 2012, SHAW *et al.*, 2012, ANAYA; BOTICARIO, 2011, BLISS; LAWRENCE 2009). É possível encontrar variações desta medida em pesquisas cujo escopo envolveu a construção de uma métrica para calcular a participação. Nestes estudos, os pesquisadores utilizaram um valor normalizado, dividindo o total de mensagens enviadas pelo discente pelo valor máximo de mensagens enviadas ao fórum (e.g., CACCIAMANI *et al.*, 2012, WEE; ABRIZAH 2011, WU; CHEN, 2005).

Além deste método de mensuração, foram encontradas na literatura revisada outras medidas quantitativas relacionadas às mensagens, tais como o total de respostas recebidas e enviadas pelo aluno e o total de tópicos (*threads*) iniciados por ele. Segundo Anaya e Boticario (2011), estas medidas refletem a iniciativa e o impacto das mensagens do aluno na discussão. Para Tobarra *et al.* (2014), um

aluno pode ser considerado proativo quando ele responde às mensagens dos demais. Chan e Chan (2011) utilizaram, em sua investigação sobre os fatores que ajudam a prever a participação dos alunos, o termo *build-on notes* associado ao número de respostas enviadas a uma *thread*.

Uma variante deste indicador foi implementada por Bliss e Lawrence (2009), que definiram, em sua abordagem para a mensuração da quantidade de mensagens, 3 (três) categorias de mensagens: (i) postagens apenas reconhecidas; (ii) postagens que obtiveram resposta; ou (iii) postagens que incentivaram a troca de informações e alimentaram a discussão.

VI. Tamanho das Mensagens

O tamanho de uma mensagem pode ser mensurado de várias maneiras, seja pelo total de palavras, sentenças ou caracteres presentes na mensagem. O tamanho das mensagens pode refletir o esforço investido pelo aluno na construção de uma mensagem. Segundo Wee e Abrizah (2011), é necessário um número mínimo de palavras para construir sentenças que possuam significado.

A contagem de palavras foi utilizada por Kay (2006) e Wu e Chen (2005) para estudar este indicador. Romero *et al.* (2013), além do total de palavras, também analisaram a contagem total de sentenças de uma mensagem. Outros pesquisadores que fizeram uso deste indicador em sua pesquisa foram Nistor e Neubauer (2010), a partir do uso da contagem de caracteres.

VII. Tempo Gasto

Os indicadores temporais são medidas quantitativas relacionadas ao tempo gasto pelos discentes, geralmente utilizados na análise de perfis e desempenho dos alunos ao longo da discussão.

O estudo de medidas temporais pode ser útil para identificar padrões de comportamentos e até indicar o esforço e comprometimento do aluno. O tempo utilizado para leitura e composição de mensagens pode sugerir que houve maior

reflexão sobre o tema abordado.

Dringus e Ellis (2010) estudaram indicadores temporais, tais como a duração do fluxo de mensagens no fórum e seu relacionamento com a qualidade da discussão. Wise *et al.* (2012) utilizaram a análise de *logs* e registros de datas e horários de acesso (*date-time stamps*) para investigar o tempo gasto pelo aluno na leitura, escrita e edição de mensagens.

Romero *et al.* (2013) e Gomes *et al.* (2012) também fizeram uso deste indicador, por meio da análise do registro do tempo de dedicação do aluno na discussão, em segundos.

A abordagem descrita por Kay (2006) utilizou a medida de tempo de resposta do aluno, representada pelo total de minutos que o aluno leva para responder a uma mensagem postada no fórum. O autor também verificou o horário de envio, indicando se a mensagem foi enviada em horário de aula ou não.

VIII. Outros Indicadores

Além dos indicadores descritos anteriormente, alguns pesquisadores utilizaram outras medidas para investigar a participação de alunos em fóruns. Tratam-se, em grande parte, de indicadores que não podem ser mensurados quantitativamente, mas que podem influenciar na participação do aluno de um modo geral.

Dois desses indicadores, citados por Kay (2006), são o nível de dificuldade do tema abordado na discussão e o nível de conhecimento dos alunos. Estes indicadores estão mais fundamentados na teoria educacional e são difíceis de serem quantificados, devido à subjetividade envolvida em sua atribuição, embora possam ser observados, por exemplo, na análise do perfil dos alunos ou na qualidade das interações.

Chan e Chan (2011) e Wise *et al.* (2012) estudaram, dentre outros indicadores, o número de edições feitas pelos alunos em mensagens já enviadas

ao fórum. Segundo Chan e Chan (2011), as tentativas de revisar o conteúdo escrito demonstram, visto em uma perspectiva de construção de conhecimento, um esforço em trabalhar melhor as ideias.

Há, ainda, outros fatores que podem influenciar na participação dos alunos. Mazzoline e Maddison (2007) compararam a variação da participação no fórum e a natureza do conteúdo postado pelos alunos quando o docente participava ativamente da discussão, mantendo o diálogo dentro do tema proposto ou quando ele apenas observava a discussão, intervindo unicamente quando julgava necessário.

Bliss e Lawrence (2009) discorreram sobre algumas medidas para quantificar a presença do docente, como, por exemplo, o uso de níveis que representem o seu envolvimento na discussão (alto, médio ou baixo). As autoras também utilizaram um indicador binário em sua métrica de avaliação para identificar o fornecimento ou não de expectativas e diretrizes. Esse indicador corresponde à comunicação clara do que é esperado do aluno e, mais importante ainda, devido à modalidade de ensino ser a distância, ao fornecimento de diretrizes bem definidas para guiar a discussão.

Ainda há outras pesquisas que fazem uso da análise de redes sociais para investigar as interações entre os alunos ou construir ferramentas que forneçam uma forma de auxílio visual ao professor, ajudando-o na tomada de decisão em relação à avaliação da participação dos alunos (e.g., CALVANI *et al.*, 2009, ERLIN *et al.*, 2009, RABBANY *et al.*, 2011, JYOTHI *et al.*, 2012).

Nos estudos analisados durante a revisão de literatura, todos os pesquisadores utilizaram abordagens que envolveram o uso de indicadores de participação. Entretanto, foi possível observar que as formas para a obtenção destes indicadores podem variar.

Na prática, a metodologia de avaliação utilizada pelo docente pode envolver diferentes indicadores, implementados por meio de diversas medidas (e.g., a obtenção do indicador de quantidade de mensagens pode ser dada por meio do

total de mensagens enviadas pelo aluno, em conjunto com o total de tópicos iniciados por ele).

Assim, durante a elaboração de um modelo para a avaliação da participação de alunos em fóruns deve-se considerar que cada docente possui sua própria metodologia de avaliação e, desta forma, o modelo deve ser capaz de prover indicadores que reflitam esta metodologia.

2.2.2 Abordagens para Mensuração da Participação Encontradas na Literatura

Wu e Chen (2005) e Wee e Abrizah (2011) são responsáveis pelas 2 (duas) pesquisas que apresentam métricas para inferir matematicamente a participação, dentre as abordagens investigadas durante a revisão de literatura.

I. Abordagem proposta por Wu e Chen (2005)

Wu e Chen (2005) utilizaram 3 (três) indicadores específicos em sua abordagem: o *tamanho da mensagem*, o *total de mensagens enviadas* e a *contagem de palavras-chaves*. Esta última faz uso de técnicas de mineração de dados, conforme descrito na Seção 2.2.1.

Os autores popuseram uma fórmula matemática, denominada *Indicador de Participação* (PaI – *Participation Indicator*), para o cálculo da participação de um aluno i no fórum, definida pela equação 2.1.

$$PaI_i = \alpha KC_i + \beta ML_i + \gamma MC_i, \quad (2.1)$$

na qual os coeficientes α , β e γ representam os pesos de cada um dos 3 (três) indicadores e são ajustáveis de acordo com a opinião do docente. KC_i representa a média de palavras-chaves (*Keyword Count*), ML_i é a média de tamanho das mensagens enviadas pelo aluno i (*Message Length*) e MC_i , a média de mensagens enviadas pelo aluno (*Message Count*).

Utilizando dados de 5 (cinco) turmas, que possuíam entre 15 a 31 alunos, os autores realizaram a validação do modelo proposto confrontando os resultados obtidos por meio do cálculo do PaI e a avaliação tradicional realizada pelos docentes. Ao final, foram encontrados valores de correlação entre os dados variando entre 62% e 92%.

Em suas conclusões, Wu e Chen (2005) afirmaram que uma forma de melhorar o modelo proposto é considerar as interações entre os alunos. O indicador do total de mensagens enviadas não diferencia respostas de *threads* iniciadas, por exemplo. Para isso, os autores sugeriram o emprego de técnicas de Análise de Redes Sociais (ARS).

Outra sugestão proposta por Wu e Chen (2005) foi a atribuição de pesos pelos docentes, uma vez que o professor pode considerar algumas palavras mais relevantes que outras.

II. Abordagem proposta por Wee e Abrizah (2011)

De forma análoga, Wee e Abrizah (2011), preocupadas em reduzir o trabalho manual dos professores em cursos *online* durante o processo de avaliação da participação, também propuseram um modelo baseado em uma medida matemática para a geração de um cálculo aproximado da participação dos alunos em fóruns de discussão.

As autoras realizaram uma análise dos critérios de avaliação educacionais encontrados na literatura, além de consultar educadores especializados em EaD, a partir de um levantamento dos requisitos destinado à atribuição do peso da participação nessa modalidade de ensino. Os professores consultados faziam uso do SCAFFOLD, uma ferramenta fundamentada no esquema de categorização de mensagens proposto por Dringus e Ellis (2004) (uma descrição mais detalhada sobre a ferramenta pode ser encontrada na Seção 2.2.1).

As autoras partiram do entendimento de que a qualidade da aprendizagem em fóruns *online* está diretamente relacionada ao tipo de mensagens nele

postadas, ou seja, a qualidade da aprendizagem manifesta-se a partir da categorização ou descrição das mensagens trocadas nos fóruns.

Assim, foi utilizada uma combinação de indicadores de participação para definir o que elas denominaram de *Indicador de Desempenho* (PeI – *Performance Indicator*), representado por uma expressão matemática baseada nos 4 (quatro) indicadores de participação que julgaram ser mais importantes. O modelo é representado pela equação 2.2.

$$\begin{aligned}
 PeI = & a * Tot_{Mess} + b * Tot_{Length} + c * Count_{Acknowledging} + d * Count_{Analysis} + \\
 & e * Count_{broadened} + f * Count_{Evaluative} + g * Count_{Clarification} + h * \\
 & Count_{Questioning} + j * Count_{Resolution} + k * Count_{Resources} + m * Count_{Social} + \\
 & n * Count_{Summarizing} + p * Count_{Synthesis},
 \end{aligned} \tag{2.2}$$

na qual os coeficientes $a, b, c, d, e, f, g, h, j, k, m, n$ e p representam pesos atribuídos pelo docente a cada indicador. As demais variáveis representam os indicadores de participação utilizados pelas autoras: média do total de mensagens postadas (Tot_{Mess}); tamanho médio das mensagens do aluno (Tot_{Length}); e a média de mensagens enviadas em cada categoria SCAFFOLD.

A avaliação do modelo de Wee e Abrizah (2011) deu-se a partir de dados extraídos de 2 (duas) disciplinas do segundo semestre acadêmico do curso de Ciência da Computação e Tecnologia da Informação de uma universidade de pesquisa intensiva da Malásia, com um total de 64 alunos matriculados.

Os resultados gerados permitiram às autoras constatar que o modelo proposto apresenta resultados similares quando comparados à nota atribuída pelo professor das disciplinas estudadas e pode ser implementado em fóruns de discussão *online*, com o fim de servir como referência a ser adotada pelo professor no processo de atribuição de notas de participação.

No entanto, as autoras concluíram que, apesar do modelo apresentar resultados satisfatórios, é necessário analisar outros fatores e realizar a validação com a utilização de dados de outros cursos.

O escopo da avaliação realizada pode apresentar uma limitação, devido ao tamanho da amostra utilizada por elas. Outro fator importante a ser citado é que os valores para grande parte dos indicadores foram atribuídos manualmente pelos docentes.

Em resumo, foi possível concluir que nenhuma das pesquisas ora apresentadas descreve uma metodologia que considere uma análise mais aprofundada da literatura e apresente um modelo conceitual da abordagem proposta que seja facilmente replicável ou adaptável à metodologia utilizada por docentes de EaD de forma geral.

Além disto, nenhuma das pesquisas estudadas considera a subjetividade no contexto da atribuição mecânica de notas de participação, apesar da alusão feita constantemente à subjetividade do processo de avaliação. Em Wu e Chen (2005), por exemplo, a avaliação das mensagens é baseada na busca automática de termos, em que o docente não tem nenhum controle sobre a real relevância destes para discussão. Em Wee e Abrizah (2011), ocorre o oposto: a análise do conteúdo das mensagens, feita por meio da definição de categorias, é realizada totalmente de forma manual pelo docente.

2.3 Sistemas Especialistas (SE)

Um dos campos de aplicabilidade de técnicas da Inteligência Artificial está relacionado às ciências cognitivas, com o objetivo de estudar o raciocínio humano e formas de modelagem do conhecimento. Os sistemas que fazem uso de tais técnicas necessitam conhecer o meio em que são aplicadas, a fim de tomar decisões. Estes sistemas são denominados Sistemas Especialistas (KHANNA *et al.*, 2010).

O conhecimento idealizado em sistemas especialistas se dá a partir da extração de informações de um especialista humano a respeito de um domínio. São estas informações que irão compor a base de conhecimento do sistema, a fim de simular o raciocínio do especialista humano em questão. Assim, a

modelagem do conhecimento torna-se parte fundamental no desenvolvimento desses sistemas.

O desenvolvimento de um SE envolve a reprodução do comportamento humano na resolução de problemas do mundo real. No entanto, é importante salientar que o domínio desses problemas é bastante específico (BARR e FEIGENBAUM, 1981).

Os SE possuem campos de aplicação em uma grande diversidade de problemas, que vão desde a manipulação do conhecimento e resolução de heurísticas, até processos de inferência. Há também, atualmente, uma grande variedade de formalismos para modelar o conhecimento em SE, e.g., regras de produção, raciocínio baseado em casos (RBC), redes neurais, redes probabilísticas e a lógica nebulosa.

A modelagem do conhecimento é necessária para representar a base de conhecimento do sistema, formando um dos componentes de SE. Além disto, o SE possui um motor de inferência, componente ativo de sua arquitetura, que funciona como um interpretador.

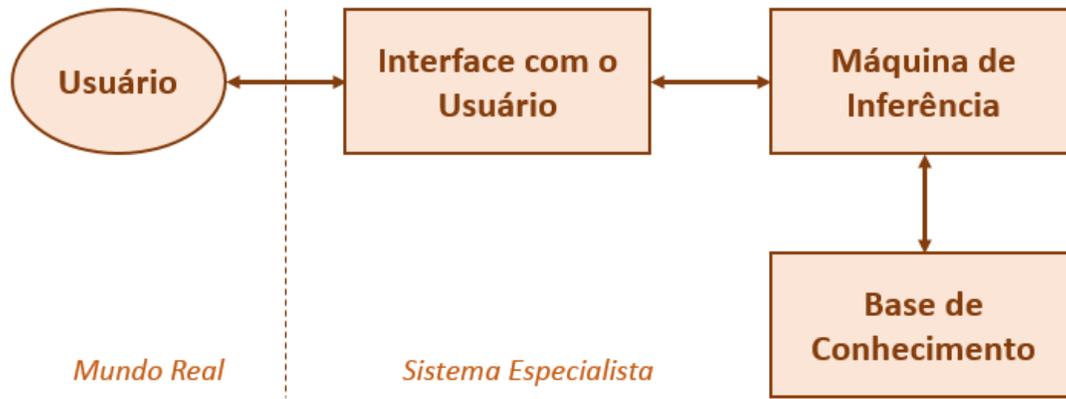
2.3.1 Arquitetura de um Sistema Especialista

A representação de um SE possui 2 (dois) componentes principais: (i) a *base de conhecimento* ou *base de regras*, que corresponde a um conjunto de declarações específicas ao domínio do problema; e (ii) o *motor de inferência* ou *máquina de inferência*, o núcleo do SE.

Na Figura 2.1, é representada uma arquitetura básica de um sistema especialista: a base de conhecimento e os mecanismos para a aquisição e inferência de conhecimento. A base do conhecimento ou base de regras é um acervo de todo o conhecimento do especialista. A modelagem desse conhecimento pode ser representada de diversas maneiras. Ainda não há um formalismo de representação mais adequado ou que seja aplicado de forma a satisfazer todo tipo de problema. No entanto, a base de conhecimento é

comumente representada por meio de regras de produção, fabricadas a partir de heurísticas.

Figura 2.1 Arquitetura básica de um Sistema Especialista.



A extração do conhecimento de um especialista é, geralmente, uma atividade desempenhada por um engenheiro do conhecimento. A composição de uma base de dados sólida não é uma atividade trivial. Geralmente, envolve a consulta a mais de um especialista e possui incertezas inerentes ao processo do raciocínio humano.

O *motor de inferência* ou *máquina de inferência* é o responsável pelo processo de resolução do problema. Os fatos, as regras e as heurísticas que formam a base de conhecimento são examinadas por este mecanismo de inferência, de maneira a tomar decisões em relação à consulta realizada pelo usuário.

2.3.2 Classificação de Sistemas Especialistas

O objetivo de um SE geralmente possui como foco a tomada de decisão, fornecendo ao especialista opções e soluções para resolver o problema em questão. Alguns dos SE descritos por Hayes-Roth (1985) são:

- *SE de Interpretação*: realizam inferências a partir de um dado conjunto de dados, i.e., levando em consideração fatos observados, o sistema realiza inferências sobre a relação existente entre eles e sua significância;

- *SE de Predição*: realizam inferências sobre os dados atuais, tomando como base dados previamente armazenados em sua base de conhecimento;
- *SE de Diagnóstico*: extraem conclusões de fatos observados, contidos em sua base de dados. As conclusões são geradas a partir de inferências sobre falhas provenientes da interpretação dos dados;
- *SE de Instrução*: realizam o diagnóstico de comportamento de aprendizado de estudantes;
- *SE de Controle*: realizam inferências sobre o comportamento de outros sistemas, por meio do monitoramento e ajuste de dados oriundos do sistema controlado.

É possível perceber, portanto, que SE são empregados em diferentes contextos, os quais abrangem desde a interpretação de regras e o diagnóstico (e.g., médico, financeiro) até a previsão e construção de novos conhecimentos.

2.4 Análise da Semelhança entre Dados Textuais

O Processamento de Linguagem Natural (PLN) é a área de conhecimento dedicada ao estudo da interpretação de particularidades da comunicação humana, por meio da análise de textos e sons. Ela se relaciona diretamente com o estudo da linguística, tratando de fatores morfológicos, sintáticos, semânticos e pragmáticos (BALDAN e MENEZES, 2012). Recentemente, o emprego de técnicas de PLN vem se tornando populares em aplicações, sendo seu principal objetivo a redução da natureza subjetiva proveniente da análise de textos (TIERNEY, 2012).

Uma das formas para realizar a comparação entre dados textuais requer a análise de seu conteúdo e a definição do grau de similaridade entre eles. Para isso, os termos contidos em cada conjunto de dados devem ser comparados, utilizando alguma medida para representar o quão semelhantes são os conteúdos dos 2 (dois) documentos.

Geralmente, para a realização da análise de similaridade entre textos, é necessária a fabricação de uma representação que torne possível a recuperação,

a visualização e a manipulação do texto. Assim, a representação dos dados textuais deve ser realizada de forma a obter o máximo de representatividade nos resultados de sua manipulação e perder o mínimo de informação possível neste processo (SALAZAR, 2012).

A seguir, são apresentadas algumas abordagens usadas para efetuar o pré-processamento destes dados e algumas medidas empregadas para o cálculo da similaridade entre textos.

2.4.1 Pré-processamento de Dados Textuais

A depender do volume do conteúdo presente em um dado textual, o custo para seu processamento pode ser elevado. Desta forma, algumas estratégias para contornar este problema envolvem a realização de um pré-processamento, mantendo o conteúdo relevante presente no texto, porém, reduzindo a sua dimensão. Algumas destas estratégias são, por exemplo, a remoção de termos irrelevantes (*stop words*) e a redução de termos a seus radicais (*stemming*), descritas a seguir.

- *Remoção de termos irrelevantes ao conteúdo do texto:* são removidos termos complementares (*stop words*), de acordo com língua utilizada no dado textual, tais como artigos, pronomes, conjunções e preposições (BAEZA-YATES; RIBEIRO-NETO, 1999). Um exemplo é a redução da frase: "*Este é o sistema operacional que me auxilia em minhas atividades diárias*" para o conjunto formado pelos termos "*sistema operacional auxilia atividades diárias*";
- *Redução de termos a seus radicais:* técnica conhecida como radicalização ou *stemming*. Parte do entendimento de que as palavras possuem variantes e que estas são semanticamente equivalentes. Assim, para realização do processamento, podem ser tratadas como um único termo. Por exemplo, os termos "*conhecimento, conhecimentos, conhecer, conhecido*" compartilham o radical comum "*conhec-*". Desta forma, poderiam ser reduzidos a esse

radical. Um dos algoritmos mais conhecidos para a implementação da radicalização é o algoritmo de Porter (1980), para a língua inglesa. Existem algumas implementações para a língua portuguesa, que são adaptações do algoritmo de Porter.

2.4.2 Similaridade Entre Textos

O grau de similaridade entre os termos de 2 (dois) documentos pode ser calculado utilizando diferentes técnicas e abordagens. Algumas das funções de similaridade mais conhecidas são descritas a seguir (SHIU; PAL, 2004).

- **Similaridade dos Cossenos:** é geralmente utilizada na comparação entre 2 (dois) documentos, por meio da medição do valor do cosseno do ângulo existente entre os 2 (dois) vetores que descrevem o texto. O valor do cosseno do ângulo entre os vetores determina o nível de similaridade dos termos existente entre eles;
- **Distância Euclidiana:** representa o valor da distância entre dos conjuntos de dados, utilizando-se, para isto, o cálculo matemático da distância euclidiana entre 2 (dois) pontos;
- **Índice de Jaccard:** também conhecido como Índice de Similaridade de Jaccard, Coeficiente de Similaridade de Jaccard ou Coeficiente de Jaccard, permite comparar a diversidade e similaridade existentes entre 2 (dois) documentos. Trata-se da razão entre o tamanho da intersecção entre os vetores que representam os textos e o tamanho da união dos 2 (dois) vetores;
- **Coeficiente de Dice:** similar ao Coeficiente de Jaccard. No entanto, a função realiza o cálculo da razão entre o dobro do valor da intersecção entre os vetores e o tamanho da união entre eles.

2.5 Considerações sobre o Conteúdo do Capítulo

O propósito deste capítulo foi descrever conceitos e definições da

fundamentação teórica necessários para a realização desta pesquisa. Adicionalmente, foram apresentadas algumas pesquisas que fizeram uso dessas informações. Este estudo da literatura forneceu o embasamento teórico necessário para o desenvolvimento da abordagem proposta, descrita no Capítulo 3. Foram, ainda, descritos alguns conceitos relacionados às técnicas de IA, tais como Sistemas Especialistas e análise de similaridade textual, necessários para melhor entendimento do processo de validação da solução proposta nesta dissertação, descrito no Capítulo 4.

Capítulo 3

Arcabouço para a Avaliação da Participação em Fóruns de Discussão

Esse capítulo destina-se à apresentação da solução proposta nesta pesquisa, fundamentada nos conceitos apresentados no Capítulo 2. Na Seção 3.1, é apresentada a solução proposta nesta dissertação e descrito o processo de finalização do arcabouço do modelo proposto, assim como os conceitos a ele associados. Na Seção 3.2, é descrita a abordagem utilizada para a validação do referido arcabouço. Por fim, na seção 3.3, são feitas algumas considerações sobre este capítulo.

3.1 Descrição da Abordagem Proposta

A avaliação do aluno, independentemente do nível ou modalidade de ensino, é essencial para a verificação de seu progresso ao longo de uma disciplina. A verificação da aprendizagem da turma deve ser realizada de forma a servir de apoio e incentivo ao processo de construção do conhecimento (VENÂNCIO; LOPES, 2013).

A atribuição de valores à participação dos alunos, em particular, desempenha a função de promover uma discussão mais substancial do tema em questão (SEETHAMRAJU, 2014). Em cursos a distância, o comprometimento dos alunos em participar ativamente em fóruns de discussão não ocorre de forma automática, o que torna a avaliação das discussões uma prática recomendada e muito adotada em diversas instituições (NANDI *et al.*, 2012). Esta avaliação é de grande importância, uma vez que a participação dos alunos é o fator que possui maior impacto em seu desempenho final no curso (FALAKMASIR; JAFAR, 2010).

Na revisão da literatura realizada, foi possível constatar, a partir da análise

de diferentes fatores que podem influenciar a participação de um aluno, que as discussões realizadas por meio do uso de fóruns em cursos de EaD servem como fator determinante para aumentar a qualidade da aprendizagem. Vale salientar que, embora existam diversas pesquisas comprometidas em investigar, validar e comparar os fatores quantitativos e qualitativos que influenciam neste processo de avaliação da participação, poucas são aquelas que fornecem uma forma concreta de mensurá-la. Ainda não existe um modelo, padrão ou sistema de simples entendimento, configuração e adaptação para este fim.

A concepção de uma metodologia de avaliação envolve um estudo aprofundado de diversos conceitos relativos à área de aplicação e dos diferentes esforços para resolução do problema já empregados pelos pesquisadores da área. A solução proposta nesta dissertação consiste na determinação de um modelo conceitual (ou abstrato), cujo objetivo é fornecer o embasamento teórico necessário à aplicação da metodologia, bem como um arcabouço, cuja finalidade é prover um recurso de fácil adaptação do modelo conceitual, que se adéque conforme as necessidades do docente, do curso e/ou da instituição que o utilizar.

Para estabelecer tal arcabouço, foi necessário realizar um estudo inicial, consultando a comunidade acadêmica usuária, representada por docentes de instituições brasileiras de EaD. Tal estudo teve como objetivo comprovar ou não a relevância dos indicadores descritos no capítulo anterior, de forma a fortalecer os conceitos envolvidos na composição da solução proposta.

3.1.1 Investigação da Opinião da Comunidade Acadêmica

Após a realização da análise do estado da arte, foi possível constatar que há uma diversidade de formas utilizadas para investigar a participação dos alunos em fóruns. No entanto, devido à natureza das pesquisas estudadas, nada se pode afirmar a respeito dos indicadores que possuem maior importância, do ponto de vista do docente, durante o processo de avaliação.

A verificação desta importância pode ser feita por meio da administração de

um questionário de sondagem (AZEVEDO *et al.*, 2012). Assim, na pesquisa ora descrita, foi administrado um questionário de sondagem, de forma a coletar informações que corroborassem a importância dos indicadores encontrados na literatura, bem como de sondar a opinião de docentes atuantes em cursos de EaD.

O questionário foi estruturado em 2 (duas) seções: uma seção inicial, para traçar um breve perfil de docentes de EaD no Brasil; e uma seção de sondagem de opinião, em que foram propostas diferentes afirmações, com o objetivo de verificar o grau de importância de alguns indicadores de participação e fatores de influência para a avaliação da participação em fóruns. O questionário exposto no Apêndice A, foi administrado durante um período de 3 (três) meses, sendo coletadas 110 opiniões no total.

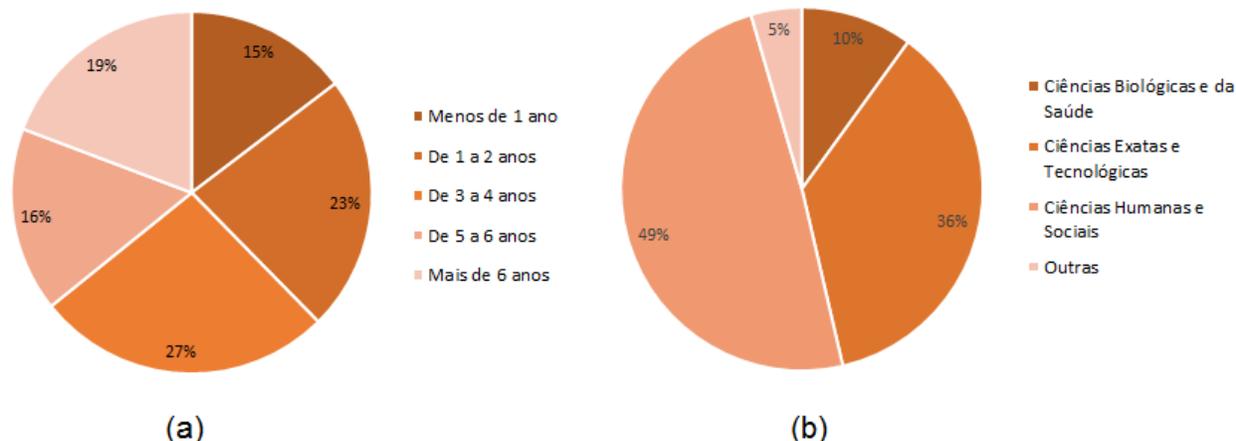
Diversas instituições brasileiras participaram da pesquisa, dentre elas: instituições públicas integrantes da Universidade Aberta do Brasil (UAB), como a Universidade Federal Rural de Pernambuco (UFPRE), Universidade Federal da Paraíba (UFPB), Universidade Federal do Ceará (UFC), Universidade Federal de Pernambuco (UFPE), Universidade Estadual da Paraíba (UEPB), Universidade Federal Fluminense (UFF), Universidade Estadual do Rio de Janeiro (UERJ), Universidade Federal de Santa Catarina (UFSC), Universidade Federal de Minas Gerais (UFMG), Universidade Federal do Maranhão (UFMA), Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB), Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco (IFPB), Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro (IFTM), Universidade Federal de Itajubá (UNIFEI); bem como instituições privadas, como a Universidade Paulista (UNIP), Universidade Norte do Paraná (UNOPAR), Universidade Metodista de São Paulo (UMESP).

I. Delineamento do Perfil do Docente

Conforme pode ser observado na Figura 3.1, o perfil dos respondentes incluiu, em sua maioria, docentes atuantes na área das Ciências Humanas e Sociais. A

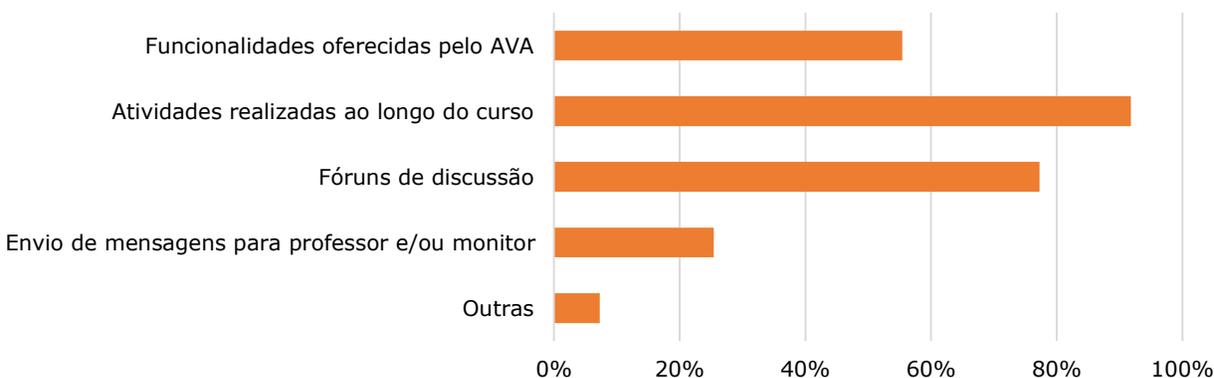
experiência dos docentes com EaD foi bem homogênea: há desde docentes ingressantes nesta modalidade de ensino até aqueles que possuem experiência superior a 6 (seis) anos de atuação em cursos de EaD.

Figura 3.1 Perfil relativo à área de atuação e tempo (em anos) de experiência dos docentes consultados em diversas instituições brasileiras.



É possível observar, a partir da análise da Figura 3.2, que o fórum de discussão foi o recurso mais utilizado para a avaliação da participação, após a análise das atividades realizadas ao longo do curso. Estas atividades consistem na realização de provas, no desenvolvimento de projetos e na criação de relatórios envolvendo o conteúdo da disciplina que, quando realizadas pelo aluno irão indiscutivelmente refletir-se em sua participação ao longo da disciplina.

Figura 3.2 Recursos utilizados pelos professores consultados para a avaliação da participação do aluno em fóruns de discussão.



No entanto, é importante lembrar que a utilização do fórum de discussão, conforme mencionado anteriormente, tem uma função importante no estabelecimento do processo de aprendizagem, o que é justificado pelos resultados apresentados.

II. Sondagem da Opinião do Docente

Na segunda etapa do questionário, o objetivo foi a coleta de opiniões a respeito de alguns fatores que podem influenciar a participação dos alunos em fóruns, bem como representar detalhes importantes levados em consideração pelo docente no momento da avaliação. Adicionalmente, alguns indicadores de participação analisados até o momento da administração do questionário foram também estudados, buscando-se delinear a opinião direta de professores a respeito de sua aplicabilidade e importância no processo de avaliação.

Quando indagados sobre a utilidade do fórum de discussão para avaliação da participação em EaD, grande parte dos docentes concordou que os fóruns de discussão assíncrona são ferramentas que exercem um papel importante neste processo de avaliação, conforme pode ser observado a partir da Figura 3.3. Estas informações também ratificam as afirmações feitas em diversas pesquisas encontradas na literatura.

Figura 3.3 Opinião dos docentes a respeito da utilidade do uso de fóruns de discussão para a avaliação da participação.



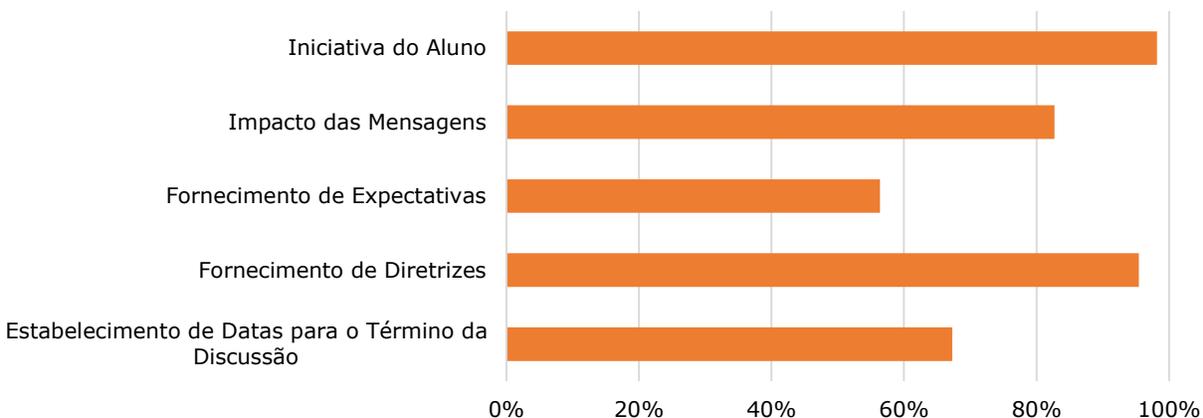
Os fatores de influência analisados são detalhados a seguir.

- a) *Estabelecimento de Datas para Término da Discussão*: corresponde à determinação de prazo final para o encerramento da discussão, a partir de sua data de início. O estabelecimento de datas finais para o encerramento das interações nos fóruns estimulam a discussão, uma vez que os alunos devem realizar o envio de mensagens dentro de um prazo previsto (WEE; ABRIZAH, 2011);
- b) *Iniciativa do Aluno*: a análise da tomada de iniciativa pode prover maior compreensão sobre sua disposição, determinação e esforço em participar da discussão. A partir deste fator, inferem-se informações sobre o comportamento do aluno e verifica-se se tal comportamento se mantém ao longo da discussão;
- c) *Impacto das Mensagens*: este fator está relacionado às atitudes de liderança do aluno em relação aos demais, ao longo da discussão. Os alunos que possuem maior tendência de liderar a discussão podem estimular os demais a participar, uma vez que suas mensagens podem trazer conteúdo de grande importância para o debate (ANAYA; BOTICARIO, 2011);
- d) *Estabelecimento de Diretrizes e Fornecimento de Expectativas*: a descrição do que é esperado dos alunos e o fornecimento de diretrizes claras, para guiar a discussão, são muito importantes em cursos *online* devido à ausência de reações imediatas entre alunos e professores (BLISS; LAWRENCE, 2009).

Como é possível observar na Figura 3.4, embora todos os fatores estudados sejam considerados importantes, visto que todos obtiveram mais de 50% de opiniões favoráveis, é possível destacar alguns fatores principais.

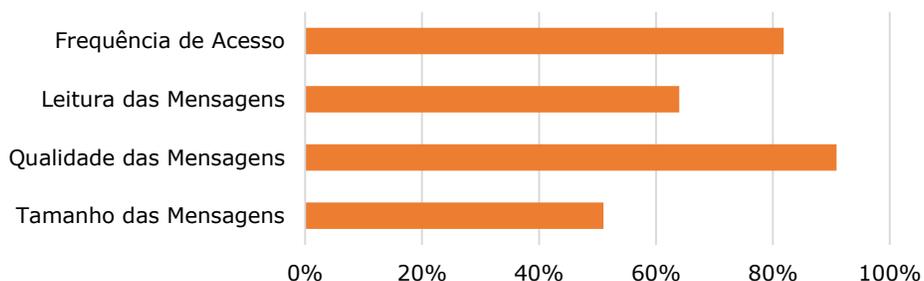
Segundo os respondentes, os fatores de maior relevância para a promoção e a verificação da participação de alunos em fóruns são: a *iniciativa do aluno*, o *prévio estabelecimento de diretrizes para a discussão* e o *impacto das mensagens do aluno na discussão*.

Figura 3.4 Resultados da sondagem de opinião de docentes de EaD a respeito de alguns fatores de influência na avaliação da participação de alunos em fóruns.



Foram listados, no questionário, alguns indicadores quantitativos e qualitativos, estudados até o momento da administração do questionário, tais como: a *qualidade das mensagens*, o *tamanho das mensagens*, a *frequência de acesso ao ambiente* e *leitura das mensagens*. Conforme pode ser observado na Figura 3.5, apesar do fato de que todos os indicadores obtiveram opiniões favoráveis acima de 50%, os indicadores de qualidade das mensagens e frequência de acesso se destacaram, com 91% e 82% das opiniões favoráveis, respectivamente.

Figura 3.5 Resultados da sondagem de opinião de docentes de EaD a respeito de alguns indicadores de participação para a avaliação de alunos em fóruns.

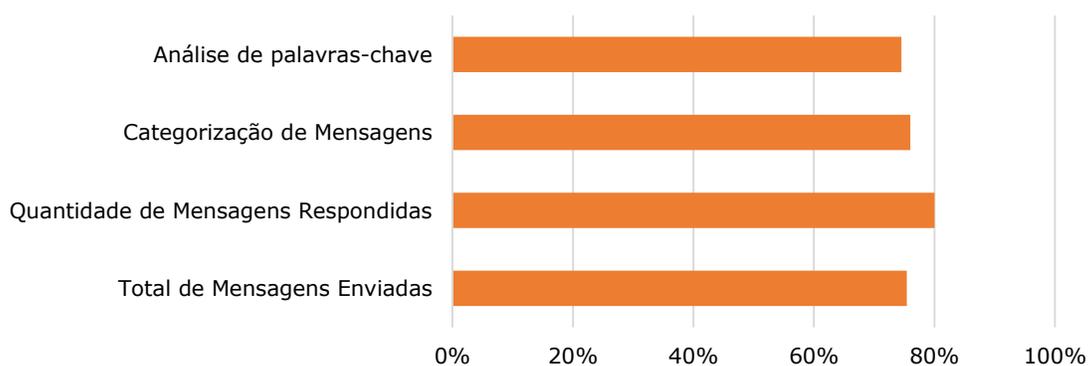


Os respondentes também opinaram a respeito de algumas medidas utilizadas para a mensuração de indicadores de participação. A *análise de palavras-chaves* e a *categorização de mensagens* são alguns dos indicadores analisados que podem representar a qualidade da discussão. Em relação ao uso

de medidas estatísticas, foi verificada a opinião dos respondentes dos indicadores: o *total de mensagens enviadas* pelo aluno e a *quantidade de mensagens respondidas* pelo aluno.

Na Figura 3.6, são apresentados os resultados percentuais obtidos de opiniões favoráveis a cada forma de mensuração supracitada. É possível verificar que todos os métodos foram considerados válidos para a implementação dos indicadores de participação de quantidade e qualidade das mensagens.

Figura 3.6 Resultados da sondagem de opinião de docentes de EaD sobre algumas formas de mensuração dos indicadores de participação de quantidade e qualidade das mensagens.



Embora não tenham sido suficientemente representativos para refletir a opinião de forma geral da comunidade docente atuante em EaD no Brasil, os resultados obtidos por meio da administração deste questionário serviram de fundamentação para as etapas posteriores da presente pesquisa, além de terem fornecido uma perspectiva geral da opinião dos docentes e a forma como eles compreendem o processo de avaliação da participação em fóruns.

3.1.2 Modelo Conceitual

Com base no estado da arte e de posse das observações realizadas por meio do levantamento de opiniões da comunidade acadêmica interessada, foi possível constatar a importância do uso de indicadores de participação na avaliação da participação.

Assim, foi possível dispor destes indicadores de forma a sistematizá-los em um modelo conceitual que determinasse, na visão do docente, fatores de influência que reflitam os aspectos quantitativos e qualitativos considerados por eles durante o processo de avaliação em fóruns de discussão. O objetivo entrevisto para tal modelo foi prover uma forma visual de apresentação dos conceitos utilizados para a avaliação da participação do aluno, que fosse de fácil compreensão e adaptação e fornecesse uma forma tangível para apoiar a decisão do professor no processo de avaliação.

Na Figura 3.7, é apresentada uma visão geral do modelo conceitual proposto. Nela, são apresentados 2 (dois) módulos principais: *Qualidade das Interações* e *Presença Social* – fatores de influência atribuídos à participação do aluno. Tais fatores representam, respectivamente, indicadores qualitativos e quantitativos.

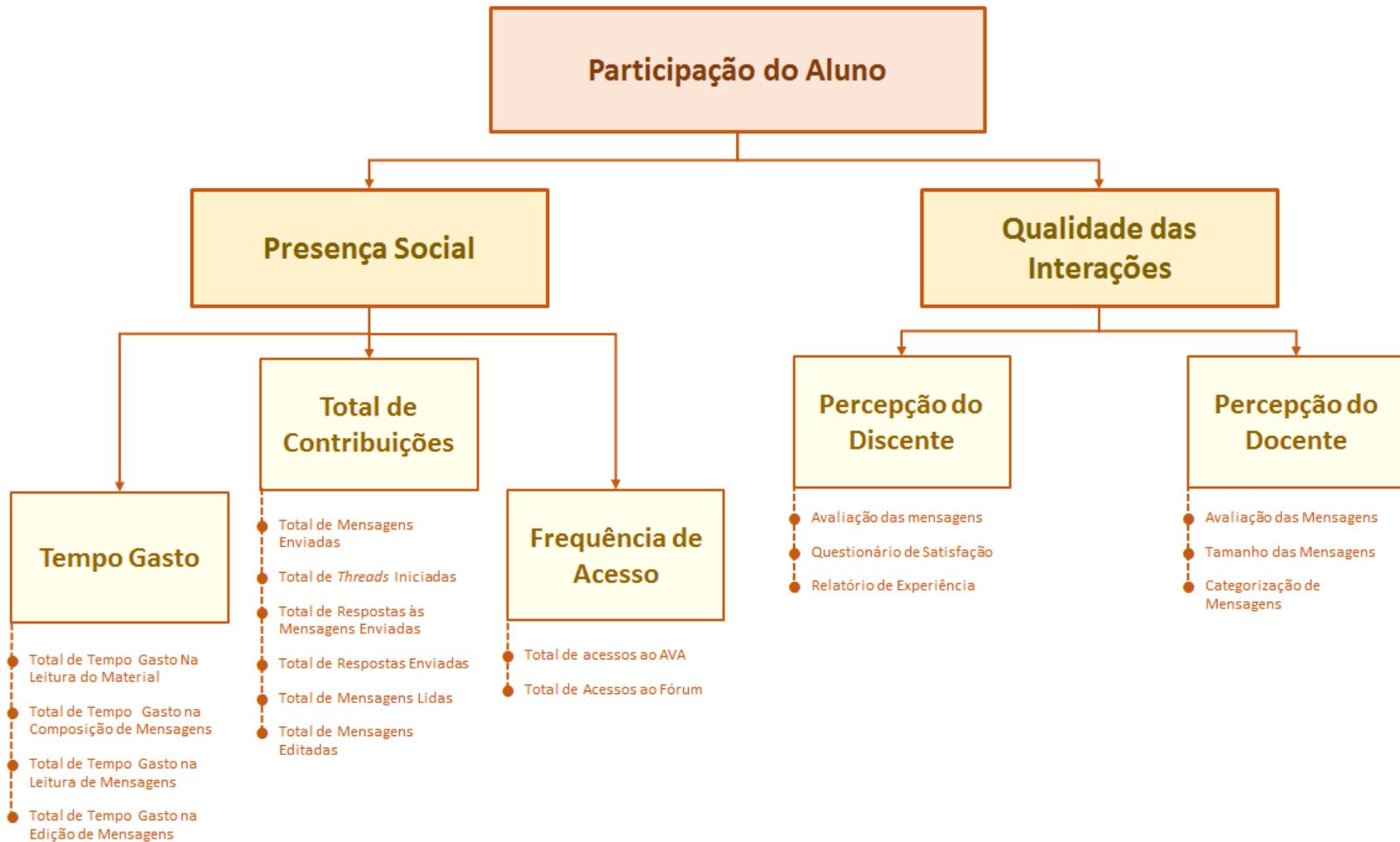
Cada módulo possui um conjunto de indicadores de participação atribuídos a ele. A qualidade das interações, por exemplo, pode ser determinada por meio da visão do próprio discente em relação à discussão ou por meio da verificação do conteúdo de cada mensagem enviada ao fórum. A presença social do aluno, por sua vez, é determinada a partir da análise da frequência de acesso, do tempo aplicado pelo aluno à leitura e à participação na discussão, e pelo total de contribuições resultantes do período em que ele interagiu com os demais.

I. Presença Social

A análise de fatores envolvendo a presença física do aluno é a forma mais comum de avaliar o aluno em uma sala de aula presencial. Em geral, o docente mantém um histórico de aulas, com observações sobre a presença do aluno na sala. Além disto, muitos professores também procuram gerar discussão ao longo da aula, de forma a pontuar ou, de forma mais simples, observar o comportamento e verificar quais os alunos que possuem maior destaque.

Uma vez que cursos de EaD tentam se aproximar do modelo de aula presencial, um modo de verificar a presença do aluno na "aula" é por meio de

Figura 3.7 Modelo conceitual para avaliação da participação de alunos em fóruns de discussão.



estatísticas. Um aluno que não interagiu em nenhum momento da discussão ou sequer esteve presente para sua leitura, não esteve presente, de fato, na discussão. Esta situação poderia ser comparada, de forma similar, ao aluno não comparecer presencialmente às aulas, resultando no registro de sua ausência nas discussões e atividades realizadas na aula.

Garrison *et al.* (2000) propuseram um arcabouço conceitual denominado de Comunidade da Investigação (CoI – *Community of Inquiry*). Neste arcabouço, os autores descreveram 3 (três) elementos que caracterizam a aprendizagem *online*: *presença social*, a qualidade das interações realizadas pelo aluno (*presença cognitiva*) e a *presença do docente*.

A presença social está relacionada ao envolvimento social e emocional de um aluno em relação aos demais participantes de uma discussão (GARRISSON; ARBAUGH, 2007). A presença cognitiva está relacionada à qualidade da discussão, uma vez que a quantidade de interações não gera, necessariamente, desenvolvimento cognitivo ou facilita a aprendizagem e a compreensão de forma significativa (HE, 2013). A presença docente está relacionada às diretrizes fornecidas ao aluno, que irão guiá-lo durante os processos sociais e cognitivos, de forma a ajudá-los, ao longo da discussão, em sua evolução pessoal e intelectual (SWAN *et al.*, 2008).

Considerando os fatores determinados pelas interações geradas entre os alunos, a presença social e cognitiva, que se apoiam nas atitudes dos alunos em relação aos demais, ao docente e à própria discussão, são fatores de grande influência para a avaliação de sua participação. McKerlich *et al.* (2011) afirmaram que não há, na literatura, um consenso entre os pesquisados sobre a definição do termo "presença". Assim, o autor o define como o "conceito de estar presente, apesar da distância física".

A presença do aluno no fórum está diretamente relacionada a sua presença social no fórum. Embora existam diversos esforços em mensurar esta presença em ambientes de aprendizagem *online* e em discutir sobre como avaliá-la, os

pesquisadores concordam que se trata de um fator crucial em qualquer tipo de educação mediada por computador (LIN *et al.*, 2004).

Na literatura estudada, propõem-se diferentes formas para mensurar a presença social de um aluno em fóruns, a partir da análise de medidas quantitativas. Baseando-se na ideia de divisão de indicadores proposta por Hrastinski (2008), o módulo *Presença Social* do arcabouço conceitual ora apresentado subdivide-se em 3 (três) indicadores de participação: o *total de contribuições* do aluno (*e.g.*, total de envios, de *threads* iniciadas, de respostas recebidas e enviadas, de edições realizadas); a *frequência de acesso* à plataforma e ao fórum; e o *tempo gasto* pelo aluno na discussão. Estes indicadores, quando utilizados em conjunto, podem significar que o aluno esteve presente na discussão e representar seu esforço em participar do debate.

No entanto, deve-se ressaltar que a análise dos relacionamentos sociais do aluno não fornece garantia alguma de que os estudantes estejam educacionalmente engajados na discussão de forma significativa (HE, 2013). Assim, é importante utilizar indicadores que representem o desenvolvimento cognitivo do aluno na discussão, representando sua presença cognitiva.

II. Qualidade das Interações

Dentre os estudos realizados no contexto da avaliação da participação em fóruns, embora sejam utilizadas diferentes abordagens para resolver ou investigar este problema, grande parte das pesquisas utiliza, de forma direta ou indireta, a verificação do conteúdo das mensagens e de *logs* de interação e a observação do comportamento dos alunos (NISTOR, 2013; HAKKARAINEN *et al.*, 2011).

Em fóruns de discussão, mesmo que um aluno tenha interagido com os demais, afirmando sua presença social no fórum, suas contribuições podem apresentar pouco ou nenhum valor semântico e em nada contribuir para a discussão (*e.g.*, *spam*, fuga do tema proposto, mensagens sem sentido ou mal elaboradas). Tais circunstâncias reafirmam a necessidade do acompanhamento

do docente, requerendo sua atenção e análise, de forma a manter o padrão de qualidade esperado da discussão.

No modelo conceitual ilustrado na Figura 3.7, o módulo relacionado à avaliação de aspectos cognitivos da discussão é denominado *Qualidade das Interações*, sendo representado por 2 (dois) indicadores de participação: a análise da qualidade determinada pelo próprio aluno, dada por medidas que reflitam a *percepção do discente*; e medidas que reflitam a forma de análise das mensagens realizada pelo professor, representadas pelo indicador *percepção do docente*. No modelo proposto por Garrison *et al.* (2000), este módulo compreende a definição de presença cognitiva, que trata da investigação do conteúdo trocado entre os alunos.

A relevância da análise da qualidade das interações e as metodologias existentes para tal análise já foram discutidas anteriormente, no Capítulo 2. No entanto, apesar de sua importância para o processo de avaliação, a verificação do conteúdo das mensagens pode representar um processo complexo, que sofre forte influência da inconsistência inerente ao processamento da linguagem natural (DRINGUS; ELLIS, 2005).

Vale ressaltar, desta forma, que é necessário levar em consideração esta subjetividade, ao se utilizar indicadores de participação para determinar a qualidade das mensagens de forma automática ou semiautomática¹.

Para este fim, pode-se utilizar técnicas de IA, como redes neurais, RBC, medidas de similaridade entre as mensagens enviadas pelo aluno e o conhecimento do docente ao avaliar o seu conteúdo, que pode ser representado por palavras ou expressões importantes para o tema proposto.

3.1.3 Arcabouço Proposto

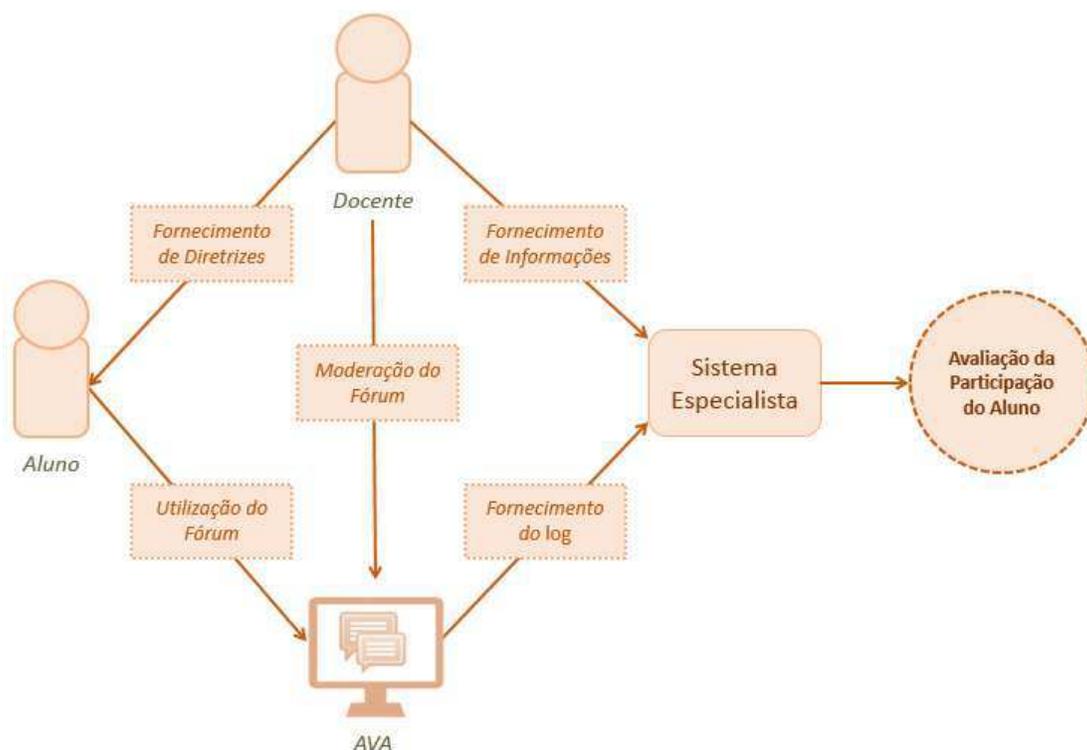
Dada a descrição do modelo conceitual, se faz necessário apontar uma forma genérica de concretizar a adaptação deste modelo, uma vez que se compreende

¹ A modelagem da análise das mensagens, visando o desenvolvimento de uma ferramenta para tal fim, pode não ser totalmente automatizada. Esse processo deve ser guiado pelo docente, uma vez que cada docente pode possuir uma metodologia diferente para realizar esta análise. 48

que os indicadores de participação selecionados para uma dada instância não serão extraídos e analisados de forma manual pelo docente. De acordo com as especificações de uma determinada instituição, curso ou do próprio docente, deve ser determinada uma instância do modelo conceitual de forma a acomodar a forma de avaliação da participação dos alunos na disciplina.

Para isto, o arcabouço proposto nesta dissertação descreve uma visão genérica dos elementos que devem compor a estrutura para o desenvolvimento de uma ferramenta computacional, que implemente a adaptação do modelo previamente definida pelo docente. Tal ferramenta deve extrair dados do AVA, provenientes de interações geradas pelos alunos no fórum e fazer uso do conhecimento do docente para realizar o cálculo da participação dos alunos. Assim, o arcabouço efetua a integração das ações realizadas pelos alunos e pelo docente e das informações coletadas automaticamente pelo AVA. Na Figura 3.8, é ilustrado o relacionamento existente entre estes atores com a ferramenta computacional a ser desenvolvida, representada por um Sistema Especialista.

Figura 3.8 Arcabouço para a avaliação da participação do aluno em fóruns de discussão em cursos de EaD.



Idealmente, o docente deve interagir com o aluno, fornecendo-lhe diretrizes sobre o que é esperado durante a discussão. Adicionalmente, o docente também irá interagir com o AVA, uma vez que sua presença é necessária para manter a discussão dentro do tema e prover incentivos à participação dos alunos. Por fim, o docente também interagirá com o SE, ao definir o modelo que será utilizado para verificação da participação em sua disciplina, o professor definirá termos e fornecerá informações sobre a discussão e os alunos, de acordo com os indicadores de participação que julgou relevantes. Esta última dependência reflete o aspecto semiautomático do arcabouço proposto: apesar de o SE realizar o processamento e avaliação da participação, é inevitável que o docente interfira neste processo, configurando, supervisionando e alimentando o sistema com informações relevantes ao processo de avaliação.

Ainda observando-se a Figura 3.8, é possível verificar que o aluno irá interagir com o AVA, a partir da utilização do fórum, enviando mensagens ou apenas acessando o ambiente para a leitura da discussão. Do AVA, provêm as informações que compõem os indicadores quantitativos utilizados, a partir da análise do *log* de interações e informações para a realização da análise qualitativa, por meio do fornecimento do *log* de mensagens enviadas ao fórum.

Desta forma, o SE representa a implementação de uma instância do arcabouço, que utiliza o modelo conceitual definido previamente, seja pelo próprio professor, pelo departamento ou instituição de ensino. Assim, ao final, espera-se que o SE desenvolvido realize a avaliação da participação do aluno. Desta forma, torna-se indispensável a especificação de um método para o cálculo da avaliação da participação. Neste contexto, pode-se utilizar, por exemplo, diferentes técnicas da Inteligência Artificial: mineração de dados e classificação automática das informações; o uso de redes sociais, em conjunto com modelos matemáticos e aprendizado semisupervisionado, dentre outras. A extração de informações do *log* de interações entre os alunos, provido pelo AVA, deve ser realizada de acordo com a estrutura fornecida pelo próprio ambiente.

Deve-se destacar, no entanto, que a saída do sistema não deve ficar limitada

apenas ao indicador de participação final (e.g., nota de participação do aluno no fórum). A avaliação da participação pode ser representada por uma série de gráficos, estatísticas ou relatórios que detalhem a quantidade e/ou a qualidade das interações realizadas pelos alunos ao longo da discussão. Com base nestas informações, bem como na sugestão de notas fornecidas pelo SE, o docente poderá, então, tomar uma decisão em relação à participação final de cada aluno.

É válido ressaltar a importância da escolha de uma técnica eficaz para a interpretação das mensagens, caso ela seja realizada pelo SE. A literatura aponta o desafio encontrado na automatização da avaliação do conteúdo das mensagens em fóruns (KERR *et al.*, 2013). Levar em consideração a acurácia de técnicas empregadas para este propósito é de grande importância, uma vez que a avaliação final, provida pelo sistema, pode afetar de forma significativa a atribuição de notas para o desempenho geral do aluno no curso. Para tanto, podem ser empregadas técnicas de PLN, RBC e Redes Neurais, dentre outras.

3.2 Abordagem Proposta para a Validação do Arcabouço

O planejamento da estratégia de validação do arcabouço levou em consideração que, para atingir os melhores resultados, a avaliação da abordagem proposta deveria utilizar dados de cursos a distância reais, de diferentes áreas, a fim verificar a sua eficácia em diferentes contextos, quando comparada com a abordagem tradicional de avaliação da participação realizada pelos docentes de EaD.

É necessário reconhecer que o planejamento, o desenvolvimento e a aplicação de pesquisas na área de EaD são atividades complexas (NEJAD; HAJIHEYDARI, 2012). Os AVA utilizados como suporte na realização de cursos nesta modalidade de ensino acumulam uma vasta riqueza de informações, extraídas por meio de arquivos de *logs*, base de dados ou ferramentas desenvolvidas especificamente para este fim. No entanto, a obtenção de uma

base de dados que contenha as configurações adequadas para a realização do processo de validação de uma abordagem pode não ser uma tarefa trivial.

Assim, durante a realização da etapa inicial de busca de uma base de dados, foi concretizada uma parceria com o Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) (Apêndice B). De posse dos dados, foi realizado um levantamento inicial do tipo de informações que poderiam ser extraídas do ambiente, de forma a obter-se uma visão geral da base.

Foram verificados fóruns temáticos de disciplinas ofertadas em 2 (dois) cursos de nível técnico, durante os períodos que abrangem o segundo semestre de 2013 e o primeiro semestre de 2014.

Os fóruns investigados possuíam grande variação em relação a sua finalidade, a depender da área da disciplina ao qual pertence. Além disso, o uso desta ferramenta nas disciplinas estava condicionado também à definição de fórum compreendida pelo próprio docente. Grande parte dos fóruns disponíveis, embora de natureza avaliativa, servia apenas como repositório de atividades da disciplina, não possuindo, assim, características de um fórum de discussão.

Outro aspecto que pôde ser observado nos dados está relacionado à atividade nos fóruns: existia uma grande variabilidade na participação dos alunos. Em alguns fóruns, havia um grande fluxo de mensagens enviadas e de interações entre os alunos. A maioria, porém, era caracterizada por uma baixa participação, em que cada aluno apenas enviava sua opinião sobre o assunto abordado, havendo pouca ou nenhuma interação entre os alunos.

Assim, foi necessária realizar uma filtragem com objetivo de investigar os temas em discussão, o objetivo do fórum e o próprio comportamento dos alunos.

Na Tabela 3.1, é apresentada uma visão geral dos dados após a filtragem. Observa-se que o espaço amostral possível de ser utilizado para a validação foi composto por 17 fóruns temáticos. Após a análise inicial dos dados, pôde-se, então, dar continuidade ao processo de avaliação.

Tabela 3.1 Sumarização da base de dados disponibilizada pelo IFPB Virtual, após filtragem dos dados.

Curso	Total de Disciplinas Ofertadas	Total de Fóruns Temáticos
Profuncionário - Secretaria Escolar	5	6
Segurança no Trabalho	9	11

Em uma situação ideal, a validação do arcabouço deveria ser realizada por meio da adaptação do modelo conceitual definido por docentes de diferentes disciplinas, utilizando diferentes combinações de indicadores de participação. No entanto, por questões de tempo e incompatibilidade de períodos nos quais novas turmas para estas disciplinas seriam ofertadas, um experimento com estas características não pôde ser realizado.

Assim, por tais razões, foi realizado um estudo comparativo utilizando dados de fóruns temáticos de cursos já concluídos, confrontando a metodologia utilizada pelos docentes responsáveis pela avaliação da participação com o uso de instâncias do arcabouço que refletiam esta metodologia.

Após a análise da base de dados, foi definido que o processo de validação seria composto por 2 (duas) etapas: uma etapa inicial, com o objetivo de coletar informações sobre o processo de avaliação da participação praticado na instituição na qual o estudo foi realizado, bem como a opinião dos docentes a respeito dos indicadores de participação comumente utilizados; e uma segunda etapa, para verificação da acurácia do modelo, confrontando as 2 (duas) abordagens: (i) a avaliação semiautomática, proposta nesta pesquisa e (ii) a avaliação tradicional, realizada pelo docente.

A segunda etapa, então, trataria da extração, tratamento, processamento e análise dos dados coletados nos fóruns temáticos disponibilizados pela instituição. Dentre os docentes consultados na primeira etapa da validação, seriam selecionados alguns voluntários para participar de estudos de caso. Para cada estudo de caso, seria concebida uma adaptação do modelo conceitual, de acordo

com os objetivos do fórum e as necessidades do docente.

Os dados provenientes do uso do arcabouço, *i.e.*, a nota de participação de cada aluno da turma ao qual o fórum pertence, seriam, então, confrontados com os dados reais, *i.e.*, a nota de participação atribuída a cada aluno feita pelo docente, após a análise manual das interações do fórum.

O Quadro 3.1 apresenta uma síntese da estrutura de validação proposta para cada uma etapas descritas anteriormente.

Quadro 3.1 Sumarização da abordagem utilizada para a validação do arcabouço proposto.

Etapa de Validação		
	Investigação sobre o Perfil e Opinião do Docente	Avaliação da Acurácia do Arcabouço Proposto
Descrição	Coleta de informações sobre o perfil e estratégias de avaliação da participação utilizadas pelos docentes do IFPB Virtual.	Verificação da aplicabilidade da solução proposta, a partir da comparação entre os resultados obtidos por meio do emprego de instâncias do arcabouço e os resultados provenientes do uso da abordagem tradicional realizada pelos docentes.
Estrutura	<ul style="list-style-type: none"> • Elaboração e administração do questionário para o delineamento do perfil dos docentes e sondagem da opinião a respeito do processo de avaliação da participação em fóruns; • Tabulação e análise dos dados coletados; • Confronto dos resultados obtidos com os resultados do estudo realizado para finalização do arcabouço, descrito na Seção 3.1.1 e de pesquisas analisadas durante a revisão de literatura. 	<ul style="list-style-type: none"> • Seleção dos estudos de caso e realização de pesquisas de campo, na forma de entrevistas semi-estruturadas; • Elaboração de adaptações do modelo conceitual, que reflitam a realidade dos estudos de caso analisados; • Realização do processamento e análise dos dados, a partir da aplicação das instâncias do arcabouço; • Análise estatística dos resultados obtidos, de forma a determinar a viabilidade do uso da solução proposta na prática; • Avaliação da acurácia da abordagem proposta, por meio da análise do erro da correlação entre os dados.

3.3 Considerações sobre o Conteúdo do Capítulo

Neste capítulo, foram apresentados os resultados de um estudo inicial, realizado por meio da administração de um questionário junto a docentes de diversas instituições brasileiras de EaD. Os resultados deste estudo tiveram como objetivo o fornecimento de informações úteis que confirmem os resultados das pesquisas analisadas no Capítulo 2.

Com base no conhecimento extraído dessa análise, foi proposto um arcabouço semiautomático para avaliação da participação do aluno em fóruns de discussão. A partir da utilização do modelo conceitual, que provê uma visão geral, de fácil entendimento, das definições utilizadas por docentes ao avaliar a participação dos alunos, o arcabouço proposto define as características desejadas para a implementação da adaptação do modelo conceitual elaborada previamente.

Assim, a partir do arcabouço apresentado, propõe-se uma forma de verificação da participação, considerando a subjetividade inerente ao processo de avaliação, que seja de fácil adaptação à realidade das instituições, curso ou disciplina que o utilize durante o processo de avaliação. Ele deverá prover ao docente informações úteis sobre a participação dos alunos de forma individual, fornecendo apoio ao docente na tomada de decisão.

Por fim, foi apresentada uma visão geral da abordagem proposta para a validação do arcabouço, que será descrita mais detalhadamente no Capítulo 4.

Capítulo 4

Validação do Arcabouço

A fim de realizar a avaliação da abordagem proposta no Capítulo 3, foi definida uma estratégia envolvendo 2 (duas) etapas. Inicialmente, na Seção 4.1, apresenta-se um estudo exploratório do perfil, das práticas de avaliação e da relevância de indicadores de participação para este processo de avaliação, realizado junto a docentes da instituição na qual a validação da abordagem proposta tornou-se possível. Na Seção 4.2, descreve-se o processo de verificação da acurácia do arcabouço de avaliação da participação, a partir da análise de estudos de caso envolvendo sua aplicação em diferentes contextos, para os quais o arcabouço foi adaptado às necessidades do docente de cada caso estudado. Na seção 4.3, apresentam-se algumas limitações identificadas no processo de avaliação do arcabouço. Por fim, na Seção 4.4, são feitas algumas considerações sobre o conteúdo deste capítulo.

4.1 Investigação sobre o Perfil e Opinião do Docente

Durante o processo de finalização do modelo conceitual, foi feito um estudo junto à comunidade acadêmica brasileira de EaD, visando a coleta de informações que apoiassem as informações obtidas durante a investigação do Estado da Arte. A descrição deste estudo pode ser observada na Seção 3.1.1 do Capítulo 3 desta dissertação. Analogamente, para a validação do arcabouço proposto, foi conduzida uma análise do perfil e da opinião da comunidade docente, embora envolvendo uma amostra mais reduzida de participantes. Esta análise possui o intuito de traçar o perfil e coletar opiniões sobre o processo de avaliação realizado por docentes da instituição de ensino parceira, o IFPB Virtual. O principal objetivo desta etapa foi confrontar os resultados desta segunda sondagem com aqueles

obtidos a partir da sondagem inicial, realizada com docentes de instituições de EaD do Brasil, assim como com os resultados obtidos nas pesquisas revisadas na literatura, de forma a comprovar a importância do uso de indicadores de participação e, conseqüentemente, de elaboração de métodos para a sua aplicação.

4.1.1 Questões de Pesquisa

Este estudo foi realizado visando a responder as seguintes questões de pesquisa:

1. *Qual o perfil do docente responsável pela avaliação da participação dos alunos em fóruns na instituição estudada?*
2. *Há um padrão, entre os docentes da instituição, na forma de realizar a verificação da participação em fóruns de EaD?*
3. *Como os docentes realizam a avaliação das mensagens enviadas pelos alunos nos fóruns de discussão?*
4. *Há diferença de opiniões a respeito do emprego sobre determinados indicadores?*
 - a. *Em caso afirmativo, qual(is) o(s) indicador(es) de maior importância para a avaliação da participação, na opinião dos docentes consultados?*
 - b. *Em caso negativo, a literatura explica o motivo da atribuição de baixa relevância aos fatores destacados?*

4.1.2 Metodologia

A sondagem sobre o perfil e a opinião do docente do IFPB foi realizada de forma presencial, por meio da administração de um questionário de pesquisa (Apêndice C). O questionário foi dividido em 2 (duas) etapas. A primeira continha questionamentos sobre o perfil do respondente, de forma a melhor compreender as características particulares dos docentes da instituição. A segunda etapa do questionário teve 2 (dois) propósitos: (i) verificar se havia uma padronização na forma de avaliar a participação do aluno em fóruns, entre os docentes do IFPB Virtual; e (ii) investigar a opinião dos docentes sobre os indicadores propostos no

modelo conceitual descrito nesta pesquisa.

4.1.3 Participantes

A amostra coletada totalizou 31 instâncias, tendo sido consultados docentes de diferentes cursos e disciplinas dos campus do IFPB das cidades de Campina Grande, Guarabira, João Pessoa, Patos e Sousa.

4.1.4 Execução

O questionário foi administrado no período que compreendeu os meses de agosto e setembro de 2014. Foram realizadas sucessivas visitas aos pólos de EaD do IFPB Virtual, nos quais o questionário foi administrado, após uma breve apresentação dos objetivos da pesquisa em questão.

4.1.5 Resultados e Discussão

A seguir, são apresentados e discutidos os resultados obtidos após a administração do questionário. A discussão foi decomposta em 2 (duas) subseções para averiguar melhor os resultados e inferir conclusões a respeito de cada uma das 2 (duas) etapas previamente mencionadas.

I. Delineamento do Perfil do Docente

Após a tabulação dos dados coletados, foi possível traçar um perfil geral dos respondentes, apresentado na Tabela 4.1. É possível observar que os responsáveis pela avaliação da participação dos alunos, representado pelo corpo docente do IFPB Virtual, são, em geral, indivíduos de 20 a 30 anos, destros, sem predominância de sexo ou da utilização de corretivos visuais. O nível de familiaridade com sistemas computacionais é, em geral, intermediário e possuem, em sua maioria, especialização em sua área de conhecimento. Em relação às plataformas computacionais utilizadas, grande parte dos consultados faz uso de PC e *notebooks* e utiliza o Sistema Operacional Windows.

Tabela 4.1 Sumarização dos dados de delineamento do perfil dos docentes consultados.

Sexo	Feminino		Masculino		
	15 (48%)		16 (52%)		
Faixa Etária	Entre 20 e 30 anos	Entre 31 e 40 anos	Entre 41 e 50 anos	Mais de 50 anos	
	17 (55%)	9 (29%)	5 (16%)	0 (0%)	
Destreza	Destro		Canhoto		Ambidestro
	26 (84%)		5 (16%)		0 (0%)
Uso de Corretivos Visuais	Sim			Não	
	14 (45%)			17 (55%)	
Nível de Conhecimento de Sistemas Computacionais	Iniciante		Intermediário		Experiente
	0 (0%)		22 (71%)		9 (29%)
Grau de Escolaridade	Graduação	Especialização	Mestrado	Doutorado	
	7 (23%)	15 (48%)	9 (29%)	0 (0%)	
Plataformas Computacionais Utilizadas	PC	Notebook	Smartphone	Tablet	
	29 (93%)	29 (93%)	24 (71%)	25 (81%)	
Sistemas Operacionais Utilizados	Windows		Windows e GNU/Linux		Windows e MAC OS
	12 (39%)		12 (39%)		7 (22%)
AVA Utilizados	Moodle			TelEduc	
	31 (100%)			1 (3%)	
Experiência do Docente	Menos de 1 ano	De 1 a 2 anos	De 3 a 4 anos	De 5 a 6 anos	Mais de 6 anos
	22 (71%)	8 (26%)	1 (3%)	0 (0%)	0 (0%)
Fóruns são ferramentas importantes para a avaliação da participação	Concordo Totalmente	Concordo Parcialmente	Nem concordo, nem discordo	Discordo Parcialmente	Discordo Totalmente
	22 (71%)	9 (29%)	0 (0%)	0 (0%)	0 (0%)

Analisando-se os fatores relacionados à EaD, é possível perceber que todos os respondentes estão familiarizados com o AVA Moodle, o que é justificável, por este ser o ambiente utilizado pelo IFPB Virtual. O único respondente a possuir conhecimento em mais de um ambiente (TelEduc, 2002) foi um dos poucos docentes a possuir mais de 3 (três) anos de experiência em EaD.

Ao considerar a sumarização do perfil dos docentes na tabela, é facilmente perceptível que grande parte deles possui menos de um ano de experiência em cursos a distância.

Em relação à opinião dos consultados a respeito da utilidade de fóruns de discussão para a análise da participação dos discentes, foi possível verificar que há um consenso de que fóruns são, de fato, ferramentas importantes para a realização deste processo. Esta informação corrobora a opinião dos docentes consultados no estudo descrito na Seção 3.1.1 e as referências a este fato verificadas na literatura consultada.

II. Sondagem da Opinião do Docente

A segunda etapa do questionário visava verificar as metodologias empregadas pelos docentes para a avaliação da participação dos alunos, mais especificamente questionando-os sobre a aplicabilidade de cada fator utilizado pelo modelo conceitual. Os docentes foram encorajados a responder de acordo com a metodologia atualmente empregada por eles, para realizar a avaliação em fóruns com os recursos disponibilizados pelo AVA da instituição.

Na Tabela 4.2, percebe-se que o fator mais utilizado para verificar as interações entre os alunos é a avaliação da qualidade destas interações. De fato, a literatura estudada cita, em diversos momentos, a importância deste fator, contextualizando os desafios de se realizar tal avaliação por meio do uso de sistemas computacionais. Este fator também foi considerado de grande relevância, obtendo cerca de 90% de opiniões favoráveis, no primeiro estudo realizado (Seção 3.1.1, Capítulo 3).

Tabela 4.2 Utilização dos fatores de avaliação empregados pelos docentes consultados.

	Fator Utilizado	Fator Não Utilizado
Frequência de Acesso	16 (52%)	15 (48%)
Tempo Gasto	6 (19%)	25 (81%)
Total de Contribuições	22 (71%)	9 (29%)
Percepção do Discente	1 (3%)	30 (97%)
Qualidade das Interações	31 (100%)	0 (0%)

Outro fator que demonstra ser continuamente utilizado pelos docentes é a quantidade de contribuições, que trata de indicadores quantitativos, e.g., a quantidade de mensagens enviadas, editadas e/ou respondidas.

Na realidade da instituição na qual o estudo foi realizado, é provável que os docentes que afirmaram fazer uso de tais indicadores necessitem de uma ferramenta que os auxilie na extração desses critérios ou realizem esse processo de forma tradicional, porque o AVA da instituição não possui nenhum outro recurso, além dos *logs* de acesso, para a obtenção dessas informações.

Além disso, este fator muitas vezes é apresentado como diretriz ao início do fórum. Em geral, os fóruns temáticos da instituição possuem baixa atividade, com uma quantidade muito reduzida de contribuições sendo geradas. Assim, os alunos estão constantemente sendo incentivados a participar (e.g., o estabelecimento de regras básicas de participação no fórum).

Um exemplo é a determinação de uma quantidade mínima de mensagens que deve ser enviada ao fórum: o aluno deve expressar sua opinião e, adicionalmente, comentar ao menos uma vez a opinião de algum colega. Assim, em alguns casos, a análise da quantidade de contribuições se torna necessária, dado que a avaliação da participação deve refletir o cumprimento ou não das diretrizes estabelecidas ao início da discussão. O docente realizará esta verificação, possivelmente, apenas a partir da observação do número de

mensagens que cada aluno irá postar no fórum.

No entanto, vale ressaltar que, na ausência de uma ferramenta de apoio, o total de contribuições, embora seja um critério importante, é uma medida de difícil obtenção. Um dos docentes que participou dos estudos de caso para a validação da arcabouço afirmou que "*(...) de certa forma, (o professor) já possui informações na plataforma que podem ajudar, mas de uma forma mais superficial. Como os dados são apresentados, não dá muito para aproveitar. (O professor) acaba perdendo muito tempo e tendo que realizar uma tabulação das informações que aparecem para poder utilizar essas informações. O que pode ser até impraticável.*"

Ao realizar uma comparação destes resultados com os da pesquisa inicial realizada (Seção 3.1.1) e a literatura estudada, pode-se facilmente entender a necessidade deste indicador para a avaliação da participação, uma vez que a maior parte dos docentes e pesquisas na área utiliza ao menos uma medida que expressa o indicador de quantidade de contribuições.

Ainda na Tabela 4.2, é perceptível que a frequência de acesso é um fator que divide opiniões. Por meio da observação das interações registradas no ambiente e durante a administração do questionário, notou-se que os docentes possuem diferentes visões sobre o fator em questão.

O AVA utilizado pela instituição disponibiliza *logs* descritivos, em que podem ser extraídas informações como data, horário e tipo de ação realizada pelos alunos de uma dada disciplina. No entanto, esses *logs* não possuem uma estrutura de fácil visualização e, a depender da quantidade de alunos matriculados e do tipo de atividade que se quer observar, podem possuir uma quantidade de informação muito grande, o que torna impraticável sua utilização na avaliação da participação dos alunos sem o auxílio de uma ferramenta computacional.

Para um dos temas de fóruns utilizados durante a realização dos estudos de caso descritos na próxima seção, por exemplo, o *log* de interações possui cerca

de 11.000 interações. Assim, sem o uso de uma ferramenta para extrair tais dados dos *logs*, o docente deveria tabular manualmente as informações de interesse, a fim de utilizá-las posteriormente, de forma justa e efetiva, na avaliação dos alunos.

Os dados contidos na Tabela 4.2 também sugerem que o fator relativo ao tempo gasto pelo discente na discussão ou na leitura do material não é levado em consideração por grande parte dos consultados.

Assim como o fator de frequência de acesso, o tempo gasto não pode ser facilmente observado no ambiente adotado pela instituição estudada. Em virtude da presente estrutura dos *logs* disponibilizados na plataforma, nada se pode inferir sobre o tempo que os alunos gastaram lendo as mensagens do fórum ou o material disponibilizado pelo professor ou compondo as mensagens.

A opinião do discente também não é levada em consideração durante a avaliação da participação realizada por grande parte dos respondentes. De acordo com a literatura estudada, o uso deste indicador não é muito comum e, a depender da forma de implementação, pode resultar em resultados enviesados, caso não seja constantemente supervisionado.

Os respondentes foram, então, indagados acerca da importância de cada medida incorporada pelo arcabouço para a implementação dos indicadores de participação.

Como pôde ser observado no AVA adotado pela instituição, além dos *logs* de interações, não há qualquer outra informação que possa servir de apoio ao docente para realizar a avaliação da participação. Assim sendo e se levando em consideração que grande parte dos docentes não realiza a extração manual desses dados, solicitou-se do respondente sua opinião sobre cada medida, independentemente de este utilizá-la ou não em sua avaliação.

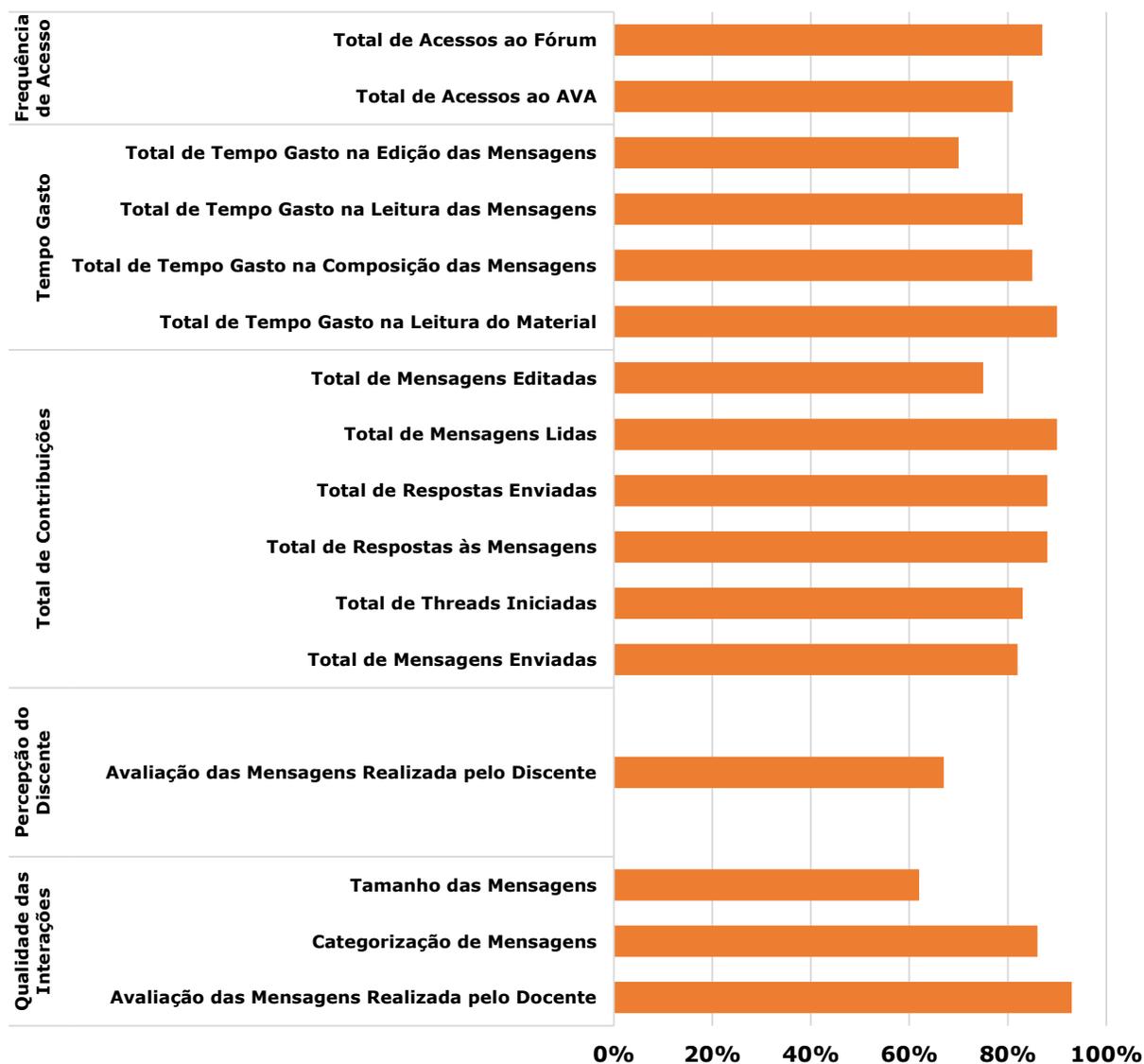
Os resultados obtidos sobre a importância média atribuída a cada fator podem ser visualizados na Tabela 4.3 e na Figura 4.1. Foi possível concluir que todos indicadores propostos foram considerados relevantes (Importância Média

acima de 80%), com exceção dos indicadores *Total de Mensagens Editadas*, *Total de Tempo Gasto na Edição das Mensagens*, *Avaliação das Mensagens Realizada pelo Discente* e *Tamanho das Mensagens*.

Tabela 4.3 Importância média atribuída pelos respondentes para cada medida proposta para a implementação dos indicadores de participação incorporados no arcabouço.

Fator	Indicadores	Importância Média
Frequência de Acesso	<i>Total de Acessos ao AVA</i>	8,1 (81%)
	<i>Total de Acessos ao Fórum</i>	8,7 (87%)
Tempo Gasto	<i>Total de Tempo Gasto na Leitura do Material</i>	9,0 (90%)
	<i>Total de Tempo Gasto na Composição das Mensagens</i>	8,5 (85%)
	<i>Total de Tempo Gasto na Leitura das Mensagens</i>	8,3 (83%)
	<i>Total de Tempo Gasto na Edição das Mensagens</i>	7,0 (70%)
Total de Contribuições	<i>Total de Mensagens Enviadas</i>	8,2 (82%)
	<i>Total de Threads Iniciadas</i>	8,3 (83%)
	<i>Total de Respostas às Mensagens</i>	8,8 (88%)
	<i>Total de Respostas Enviadas</i>	8,8 (88%)
	<i>Total de Mensagens Lidas</i>	9,0 (90%)
	<i>Total de Mensagens Editadas</i>	7,5 (75%)
Percepção do Discente	<i>Avaliação das Mensagens pelo Discente</i>	6,7 (67%)
Qualidade das Interações	<i>Avaliação das Mensagens pelo Docente</i>	9,3 (93%)
	<i>Categorização de Mensagens</i>	8,6 (86%)
	<i>Tamanho das Mensagens</i>	6,2 (62%)

Figura 4.1 Importância média (em %) atribuída pelos respondentes para cada medida proposta para a implementação dos indicadores de participação incorporados no arcabouço.



A medida que considera o tamanho das mensagens foi motivo de discussão, não apenas na literatura estudada, mas entre os docentes consultados. Um dos docentes participantes nos estudos de caso descritos na próxima seção, reflete acerca deste fator, afirmando:

"(...) não acho que o tamanho das mensagens seja relevante, principalmente em uma avaliação qualitativa. Muitas vezes o aluno posta duas ou três linhas, que

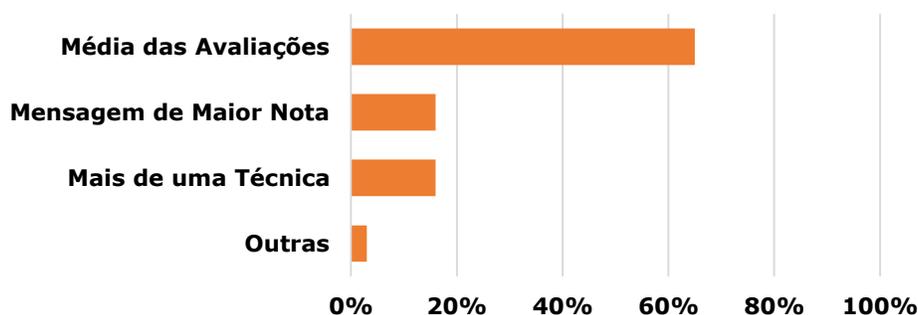
tem muito mais informação relevante do que um que preencheu diversos parágrafos. No primeiro caso, o aluno pode inserir sua subjetividade e até fazer uma crítica em relação ao conteúdo que ele está estudando. No segundo caso, ele pode ter feito apenas um 'copiou, colou' e, então, em que este aluno contribuiu para discussão?"

Em se tratando da avaliação das mensagens enviadas ao fórum, os docentes também forneceram informações a respeito das metodologias empregadas para avaliar o conteúdo das postagens. Os resultados podem ser observados na Tabela 4.4 e na Figura 4.2.

Tabela 4.4 Uso de técnicas, pelos docentes consultados, para a avaliação das mensagens em fóruns de discussão.

Média das Avaliações	Mensagem de Maior Nota	Mais de uma Técnica	Outras
20 (65%)	5 (16%)	5 (16%)	1 (3%)

Figura 4.2 Técnicas de Avaliação das mensagens enviadas ao fórum utilizadas pelos docentes consultados.



Grande parte dos docentes consultados atribui notas a todas as mensagens enviadas pelo aluno, calculando a média aritmética entre as notas concedidas. No entanto, é possível perceber que alguns docentes utilizam mais de uma técnica para realizar esta avaliação, o que é explicável pelo fato de um mesmo docente poder realizar atividades simultâneas em mais de uma disciplina.

4.2 Avaliação da Acurácia do Arcabouço Proposto

A avaliação da acurácia das instâncias do arcabouço proposto envolveu a análise dos dados provenientes das adaptações do modelo conceitual, processadas por meio do uso de um sistema computacional. O processo para a concepção, extração e a análise destes dados é descrito nas seções a seguir.

4.2.1 Questões de Pesquisa

O estudo da acurácia do arcabouço envolveu a análise das seguintes questões de pesquisa:

- 1. Os resultados obtidos por meio do uso do arcabouço apresentam desempenho similar aos da avaliação tradicional da participação de alunos em fóruns, realizada pelo docente?*
- 2. O quão próximas estão as notas de participação simuladas das notas de participação atribuídas pelo docente?*

A partir da primeira questão de pesquisa, a fim de ser realizada uma análise estatística com os dados do estudo, foi determinada a seguinte hipótese nula:

H₀: Não há diferença estatística significativa entre os resultados obtidos por meio da utilização do arcabouço e a avaliação tradicional da participação de alunos em fóruns, realizada pelo docente.

O segundo questionamento foi respondido por meio da análise de correlação entre as notas e o estudo do erro, descritos a seguir.

4.2.2 Metodologia

A metodologia de avaliação da participação empregada pelos docentes de cada estudo de caso pode ser dividida em 3 (três) subetapas, a saber: (i) a seleção de fóruns temáticos e a realização de entrevistas com docentes responsáveis pela avaliação nestes fóruns; (ii) a adaptação do modelo conceitual para cada fórum temático estudado (para fins descritivos, cada fórum temático será denominado

um estudo de caso diferente) e a criação de uma instância do arcabouço, de forma a refletir o processo de avaliação realizado pelos docentes; e, por fim, (iii) o processamento e análise dos dados para cada estudo de caso.

I. Seleção dos Estudos de Caso e Realização de Pesquisas de Campo

O contato inicial para o recrutamento de docentes para participar do processo de validação deu-se de forma presencial, após visitas ao pólo de EaD da instituição para a administração do questionário de sondagem da opinião (descrito na seção anterior). Posteriormente, o contato foi realizado por mensagens eletrônicas (*e-mail*) e por telefonemas.

Ao final da pré-seleção de docentes, foram definidos 3 (três) estudos de caso, compostos por fóruns temáticos das áreas de Ciências Humanas, Saúde e Ciências Exatas. A Tabela 4.5 contém o curso, a área e o total de fóruns ao qual pertence cada tema de discussão analisado, com o propósito de fornecer uma visão geral dos dados coletados para cada estudo de caso, *i.e.*, total de alunos matriculados (total de instâncias utilizadas para o estudo de caso) e total de mensagens enviadas aos fóruns.

Tabela 4.5 Sumarização de dados referentes a cada estudo de caso utilizado para a avaliação da acurácia do arcabouço proposto.

Estudo de Caso	Tema do Fórum	Total de Fóruns	Período de Atividade	Total de Instâncias	Total de Mensagens
1	Discutindo sobre Sistemas Operacionais	8	2014.1	201	159
2	A Educação por meio da História	6	2013.2	355	385
		8	2014.1	227	222
3	Conhecimentos Iniciais sobre GSST	8	2013.2	71	61
		8	2014.1	148	143

Uma pesquisa de campo foi realizada junto aos docentes responsáveis pela criação, manutenção e avaliação destes fóruns, na forma de uma entrevista semiestruturada (Apêndice D).

II. Adaptação do Modelo Conceitual

Conforme previamente descrito, o arcabouço proposto deve ser de fácil adequação às condições impostas pelo docente e/ou pela instituição para a avaliação. Assim, determinou-se que a análise de sua aplicabilidade poderia ser atribuída à adaptação do modelo conceitual, que devem ser implementadas por meio de uma ferramenta computacional. No entanto, devido à natureza do estudo previsto, o domínio do modelo conceitual necessitou ser reduzido, pois as interações e coleta dos dados não puderam ser realizadas durante o período em que as disciplinas foram ofertadas.

A seguir, descreve-se a metodologia adotada para a adaptação do modelo conceitual para cada um dos 3 (três) estudos de caso, bem como da implementação da instância do arcabouço desenvolvida para a execução da validação proposta.

A) Adaptação do Modelo Conceitual para os Estudos de Caso

Após a realização das entrevistas com os docentes responsáveis pelos fóruns de cada estudo de caso, foi possível ter uma visão geral dos fatores considerados pelos mesmos durante o processo de avaliação da participação.

Conforme pode ser verificado no Quadro 4.1, ambos os docentes dos estudos de caso 1 e 2 fizeram uso de apenas um indicador – a opinião do docente, tendo considerado apenas uma medida de participação – a avaliação das mensagens. Para o estudo de caso 3, o docente considerou não somente a avaliação das mensagens, mas também uma das métricas que descrevem o indicador do total de contribuições: o total de mensagens enviadas pelos alunos. O cálculo da participação para os estudos de caso 1 e 2, conforme descrito pelos docentes, é realizado apenas por meio do cálculo da média aritméticas das avaliações das mensagens do aluno.

Quadro 4.1 Indicadores de participação considerados pelos docentes para a avaliação da participação dos alunos nos fóruns estudados.

Estudo de Caso	Presença no Ambiente			Qualidade das Interações	
	Tempo Gasto	Total de Contribuições	Frequência de Acesso	Percepção do Discente	Percepção do Docente
1					•
2					•
3		•			•

Devido à utilização de ambos os módulos para o terceiro estudo de caso, se fez necessária a elaboração de um modelo matemático para o cálculo da participação. Na entrevista conduzida, o docente foi perguntado sobre a influência de cada fator durante a avaliação da participação. Essa influência foi representada por meio de pesos. Assim, para o estudo de caso 3, foi possível definir o cálculo da participação dos alunos a partir da equação 4.1.

$$Part = \left(\frac{NTME_{Aluno}}{NMME_{Forum}} \times P_{Presença} \right) + (MAM_{Aluno} \times P_{Qualidade}), \quad (4.1)$$

na qual:

$NTME_{Aluno}$: Número Total de Mensagens Enviadas pelo Aluno;

$NMME_{Forum}$: Número Máximo de Mensagens Enviadas ao Fórum;

MAM_{Aluno} : Média das Avaliações das Mensagens do Aluno;

$P_{Presença}$: Peso do fator Presença no Ambiente;

$P_{Qualidade}$: Peso do fator Qualidade das Interações.

B) Implementação do Arcabouço

Após a adaptação do modelo conceitual para cada estudo de caso, foi possível definir um escopo para o desenvolvimento do SE, utilizando uma estrutura

reduzida do arcabouço proposto. Assim, inicialmente, foi realizada a coleta dos *logs* de interação e de mensagens dos fóruns para cada caso estudado, com o objetivo de verificar qual a melhor estratégia para o tratamento e manipulação dos dados extraídos dos fóruns. No entanto, ao averiguar as interações dos alunos nos fóruns de discussão selecionados, constatou-se que a plataforma IFPB Virtual não possui um método para extração da discussão. A extração dos *logs* de mensagens, portanto, foi realizada de forma manual.

Desta forma, foi desenvolvido um sistema simplificado, responsável pelo pré-processamento das mensagens e dos *logs* de interação e de mensagens, para posterior obtenção das notas de participação provenientes do emprego das instâncias do arcabouço abstrato. O sistema foi desenvolvido na linguagem Java, a partir do IDE Eclipse.

Os dados de entrada do sistema foram compostos por informações necessárias ao processamento da avaliação da participação: (i) os pesos de cada fator considerado na adaptação do modelo conceitual; (ii) os conjuntos de palavras-chaves, utilizados no cálculo da nota das mensagens do fórum; e (iii) os arquivos de *log* de interação do fórum de discussão, do *log* de mensagens e da listagem de alunos matriculados na turma. A partir das entradas descritas, o sistema realizou o processamento em 4 (quatro) etapas: (i) pré-processamento dos dados; (ii) análise dos *logs* para extração de dados estatísticos; (iii) cálculo da avaliação das mensagens; e (iv) cálculo da avaliação final da participação de cada aluno.

Etapa 1: Pré-processamento dos dados

Esta etapa inicial consistiu na padronização dos dados de entrada, realizada da seguinte forma:

1. Remoção de sinais de pontuação e acentuação de palavras-chaves fornecidas e dos arquivos de *log*;
2. Conversão de caracteres maiúsculos e minúsculos, de forma a assegurar que não haveria diferenciação entre caracteres;

3. Remoção de *stopwords*, por meio da redução dos dados textuais a um vetor de termos de maior relevância, visando reduzir o tempo de processamento. A base de termos comuns da língua portuguesa foi criada baseando-se nos *stopwords* para a língua portuguesa do Snowball (Snowball 2012); e
4. Redução de termos aos seus radicais, com a utilização da adaptação do algoritmo de Porter para a língua portuguesa, utilizando-se a biblioteca PTStemmer (PTSTEMMER, 2010). A utilização da técnica de *stemming* auxiliou a otimizar a técnica da utilização de conjuntos de palavras-chaves para a avaliação das mensagens, uma vez que muitos dos termos fornecidos pelos docentes oscilavam entre verbos e substantivos, podendo assim, ser generalizados.

Etapa 2: Análise dos Logs de Interação

Dentre os 3 (três) estudos de caso realizados, apenas um dos docentes utilizou uma medida estatística no processo de avaliação da participação dos alunos – o estudo de caso 3. Por esta razão, foi realizada uma etapa para a verificação do *log* de interações, a fim de realizar a contagem do total de mensagens enviadas por cada aluno ao longo da discussão.

Neste processo, obteve-se, também, o valor máximo de mensagens enviadas ao fórum, utilizado para o cálculo da nota de participação, conforme a Equação 4.1.

Etapa 3: Avaliação das Mensagens

Diante dos resultados favoráveis provenientes de estudos anteriores (Apêndice E), além de outras contribuições encontradas na literatura (e.g., WU; CHEN, 2005), foi proposta a utilização de palavras-chaves no processo de análise qualitativa das mensagens do fórum, por meio do uso de funções de similaridade. Desta maneira, esta etapa consistiu na análise da similaridade entre cada mensagem enviada ao fórum e os conjuntos de palavras-chaves previamente definidos pelo docente, que refletiam o tema principal do fórum.

As palavras-chaves foram divididas em grupos que representassem um conceito ou ideia geral que deve ser discutido ou mencionado pelos alunos. Assim, cada conjunto de palavras-chaves foi, inicialmente, comparado a cada mensagem do fórum, para o cálculo posterior da nota atribuída à mensagem, de acordo com a quantidade de conceitos/ideias identificados na mensagem. A medida de similaridade utilizada fez uso da função de Similaridade dos Cossenos, disponibilizada pela biblioteca SimMetrics (CHAPMAN, 2004). Ao final, conforme definido pelos 3 (três) docentes consultados, foi realizado o cálculo da média aritmética entre as mensagens enviadas por cada aluno.

Etapa 4: Avaliação Final da Participação

A última etapa de processamento do sistema consistiu no cálculo da nota final de participação de cada aluno matriculado na disciplina, de acordo com os indicadores utilizados pelo docente.

A saída esperada para o sistema consistia em um arquivo contendo a nota de participação dos alunos para cada fórum atribuída pelo docente e da nota de participação para o aluno calculada por meio da implementação do modelo conceitual definida pelo docente.

III. Processamento e Análise dos Dados

Após o processamento dos dados, foi realizado um estudo estatístico para determinar se havia diferença na utilização do arcabouço proposto e o uso da avaliação tradicional realizada pelos docentes. Tal estudo envolveu o uso de funções para a análise da normalidade das amostras, seguido pelo uso de um teste estatístico de hipótese.

Outro fator analisado foi a acurácia do arcabouço proposto, determinada pelo estudo de correlação e estudo do erro entre as notas dadas pelo docente e as notas calculadas por meio do uso do arcabouço. Assim, esse estudo possibilitou a verificação do quão próximo o valor simulado pelo arcabouço estava do valor

real atribuído pelo docente. Para a realização da comparação entre as metodologias de avaliação, foi necessário estabelecer que:

- Os alunos de cada curso foram submetidos ao mesmo conteúdo e participaram de fóruns com o mesmo objetivo, independente do período em que ele esteve aberto para discussão;
- Em uma mesma disciplina, a avaliação da participação foi realizada pelos docentes, a partir da mesma metodologia;
- Todos os dados receberam o mesmo tratamento durante a etapa de pré-processamento. O sistema de processamento e o cálculo da nota de participação foi configurado de acordo com as informações fornecidas por cada docente sobre seu processo de avaliação; e
- As notas de participação derivadas do emprego do arcabouço e da avaliação tradicional realizada pelo docente seriam representadas por uma variável quantitativa contínua, contida no intervalo [0,00; 1,00].

Em relação ao estudo da acurácia da metodologia proposta neste trabalho foi realizado um estudo do erro médio e da correlação entre as notas calculadas com o uso da abordagem proposta pelo arcabouço e da abordagem tradicional, realizada pelo docente.

4.2.3 Participantes

Foram convidados a participar do estudo docentes de cursos do IFPB Virtual, tendo em vista serem os responsáveis pela manutenção dos fóruns temáticos e pela avaliação da participação dos alunos nos cursos a distância ofertados pela instituição. A realização das entrevistas com os docentes foi feita de forma presencial, nos campus do IFPB das cidades de Campina Grande (para os Estudos de Caso 1 e 2) e Patos (Estudo de Caso 3), durante os meses de agosto, setembro e outubro de 2014. Devido ao termo de sigilo assinado durante o processo de parceria com o IFPB Virtual, não foi possível anexar a transcrição das entrevistas realizadas.

4.2.4 Execução

Após a transcrição das informações fornecidas pelos docentes e a extração e processamento das informações dos fóruns em estudo, os dados coletados foram submetidos a uma análise estatística e um estudo da acurácia do arcabouço, visando responder as questões de pesquisa previamente levantadas.

4.2.5 Resultados e Discussão

Os resultados provenientes da comparação entre a abordagem proposta e a abordagem tradicional empregada pelo docente são apresentados e discutidos a seguir.

I. Análise Estatística dos Dados

Após a definição da hipótese a ser investigada nesta etapa da avaliação e antes da realização da análise inferencial, foi necessário verificar se os dados coletados no experimento atendiam os requisitos de normalidade, para que fosse definido se seriam utilizados testes paramétricos ou não-paramétricos. Assim sendo, foi realizado um estudo geral de algumas medidas para cada conjunto de dados.

Os gráficos de dispersão apresentados nas Figuras 4.3, 4.4 e 4.5, para os Estudos de Caso 1, 2 e 3, respectivamente, oferecem melhor entendimento do comportamento dos dados.

Idealmente, a distribuição de dados deveria possuir uma representação linear, pois as notas da participação simulada e da participação atribuída pelo docente deveriam ser equivalentes ou suficientemente próximas.

É possível notar que, para os Estudos de Caso 1 e 2, há uma dispersão menor entre os dados, havendo menor variabilidade entre os valores simulados e reais do que no Estudo de Caso 3, cujos dados apresentam-se mais esparsos.

Figura 4.3 Gráfico de dispersão entre o conjunto de dados obtidos pelo uso da abordagem proposta e tradicional, para o Estudo de Caso 1.

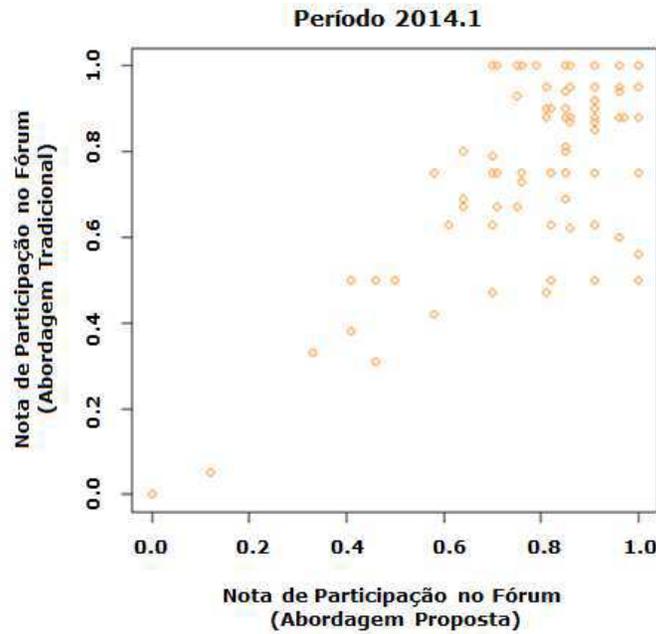


Figura 4.4 Gráfico de dispersão entre o conjunto de dados obtidos pelo uso da abordagem proposta e tradicional, para o Estudo de Caso 2.

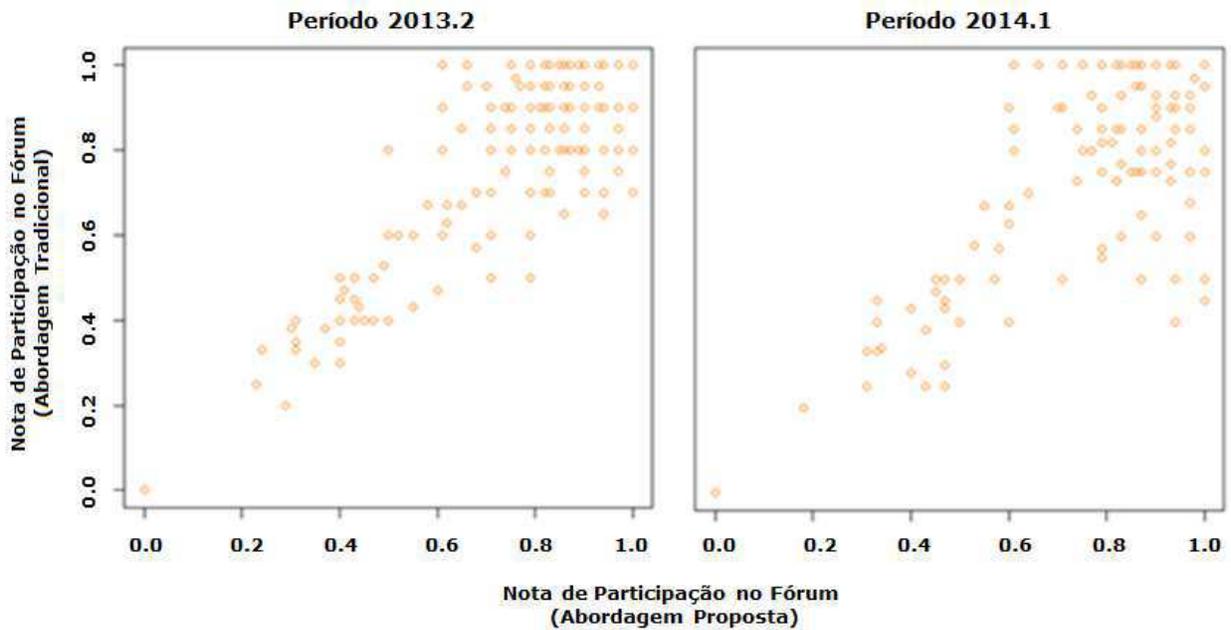
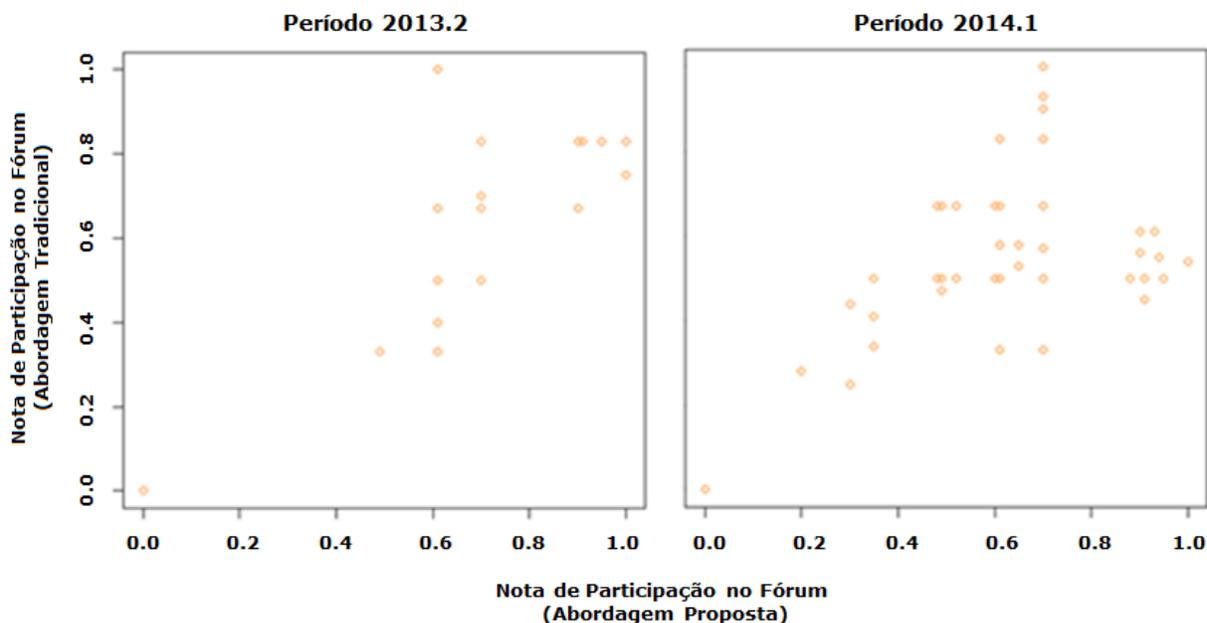


Figura 4.5 Gráfico de dispersão entre o conjunto de dados obtidos pelo uso da abordagem proposta e tradicional, para o Estudo de Caso 3.



É necessário notar que este último caso apresenta menor quantidade de instâncias em sua amostragem do que os demais. Além disso, a dispersão origina-se a partir da distância entre a nota simulada e a nota atribuída pelo docente. Assim, essa distância pode dever-se à representatividade dos conjuntos de palavras-chaves fornecidos para o cálculo das notas das mensagens. Quanto menor o número de conjuntos atribuídos, menor será a variabilidade de notas possíveis para atribuição às mensagens, *i.e.*, um fórum com apenas um conjunto de palavras-chaves, conforme representado pelo exemplo fictício a seguir,

cores: {*amarelo, vermelho, azul, verde, lilás, branco, preto*} (Ex. 01)

só fornecerá uma dentre 2 (duas) possíveis notas: 0 ou 1. Desta forma, no exemplo, caso o aluno tenha discutido sobre cores, citando pelo menos uma das palavras relevantes para este tema, sua nota de participação será 1. Se, em sua mensagem, não houver nenhuma menção às palavras contidas no único conjunto fornecido pelo professor, cores, sua nota será 0.

A definição de conjuntos significativos pode ser um processo complexo, a depender do tema do fórum ou, ainda, da experiência do docente em relação ao tema abordado no fórum. Dentre os casos analisados, o total de conjuntos de palavras-chaves referentes ao Estudo de Caso 3 foi o menor (apenas quatro conjuntos), o que pode afetar, conforme esclarecido anteriormente, o poder de avaliação da técnica utilizada para verificação das mensagens.

Visando-se a apoiar a análise visual dos dados, foram executados testes para verificação de normalidade, cujos resultados confirmaram o esperado: os dados não seguem uma distribuição normal, o que se deve à expectativa de que os dados sejam conformes a uma distribuição linear. Após conduzida a análise de normalidade, concluiu-se que não seria possível aplicar um teste paramétrico. Assim, foi utilizado o teste não paramétrico Wilcoxon-Mann-Whitney (SIGEL; CASTELLAN, 1988). A Tabela 4.6 contém os resultados de sua aplicação, tendo sido adotado um nível de significância de 5% ($\alpha = 0,05$).

Tabela 4.6 Resultados dos testes não-paramétrico Wilcoxon-Mann-Whitney para os três Estudos de Caso (nível de significância $\alpha = 0,05$).

Estudo de Caso	Período	p-valor
1	2014.1	0,7623
2	2013.2	0,0002
	2014.1	0,3902
3	2013.2	0,0023
	2014.1	0,0476

A hipótese nula para o teste Wilcoxon-Mann-Whitney indica que as medianas populacionais são iguais para as 2 (duas) amostras. Desta forma, aceitando-se a hipótese nula, tem-se que não há diferenças estatísticas entre os valores simulados e reais. Assim, a hipótese nula não pode ser rejeitada para o Estudo de Caso 1 nem para as turmas do período 2014.1 dos Estudos de Caso 2 e 3.

A experiência do docente pode exercer certa influência na não aceitação da

hipótese nula para as turmas do período 2013.2. Em geral, a escolha das palavras-chaves é feita a partir do conhecimento adquirido pelo docente após a repetição dos casos: a otimização de um conjunto de palavras ocorrerá à medida que o docente realizar a leitura de novas mensagens, uma vez que os alunos podem trazer novos termos e expressões, antes não considerados.

A leitura e a interpretação das mensagens pelo docente também pode ter sofrido influência de alguns fatores ao longo da discussão e, até mesmo, de um período para outro: algo considerado relevante anteriormente, pode não ter mais o mesmo valor, na avaliação geral da mensagem. Devido à proximidade da realização do estudo com o encerramento do período letivo 2014.1, as palavras selecionadas e a forma de leitura crítica das mensagens pelo docente podem refletir o caso mais recente (*i.e.*, as turmas do período 2014.1).

II. Estudo do Erro e Correlação entre os Dados

A análise de correlação e o estudo do erro são necessários para verificar a acurácia do emprego de instâncias do arcabouço e dos resultados dele obtidos, quando comparados aos resultados obtidos a partir da abordagem tradicional. Na Tabela 4.7, são apresentados os resultados da análise de correlação de Pearson, entre as notas da participação calculada por meio da abordagem proposta e atribuída pelo docente. É possível notar que, para todas as turmas dos estudos de caso, há sempre uma forte correlação positiva entre as notas (o fator de correlação varia entre 88% e 97%).

Tabela 4.7 Resultados para o teste de correlação entre as notas simuladas e as notas atribuídas pelo docente, para cada estudo de caso analisado.

Estudo de Caso	Período	Correlação entre Notas Simuladas e Atribuídas pelo Docente
1	2014.1	0,9512
2	2013.2	0,9703
	2014.1	0,9420
3	2013.2	0,9350
	2014.1	0,8807

Vale salientar que o teste de correlação não leva em consideração a distribuição da frequência das variáveis, apenas o quão forte é a associação entre elas. As pesquisas encontradas na literatura que descreveram métricas para o cálculo da nota de participação dos alunos (Seção 2.2.2, Capítulo 2) também realizaram um estudo da acurácia do modelo proposto baseando-se no cálculo da correlação de Pearson. Assim foi possível comparar os resultados obtidos no presente estudo com os descritos pelas pesquisas citadas.

Wu e Chen (2005) analisaram 5 (cinco) turmas, em um mesmo fórum de discussão, que possuíam entre 15 a 31 alunos. Os valores de correlação encontrados pelos autores para cada turma variam entre 62% e 92%. A utilização do cálculo totalmente automatizado pode ter impactado nos resultados obtidos pelos autores, os quais admitiram a necessidade do aumento da flexibilidade da abordagem por eles proposta, a fim de melhor acomodar as necessidades dos docentes.

Os resultados obtidos por Wee e Abrizah (2011) também foram favoráveis, ao encontrar fatores de correlação entre 83% e 99% entre o PaI (*Participation Indicator*¹) e o valor real da participação atribuído pelo docente, nos estudos realizados. No entanto, vale salientar que as amostras utilizadas por elas possuíam apenas 64 instâncias, distribuídas entre 2 (dois) cursos (*i.e.*, as notas comparadas correspondiam a apenas 64 alunos).

Além disso, com exceção do processo de extração de critérios estatísticos, como o total de mensagens enviadas e tamanho das mensagens, os demais indicadores foram atribuídos manualmente pelos docentes. Aumentou-se, desta forma, a proximidade da avaliação simulada e da avaliação tradicional realizada pelo professor.

O erro também foi considerado no estudo da acurácia do arcabouço descrito nesta pesquisa. A Tabela 4.8 contém o Erro Quadrático Médio (EQM) associado a cada estudo de caso considerado. Em todos os casos, o erro foi inferior a 3%, estando muito próximo a 0 (zero).

¹ Nota de participação calculada pela abordagem proposta pelas autoras.

Tabela 4.8 Resultados para o cálculo do EQM entre as notas simuladas e atribuídas pelo docente, para cada estudo de caso analisado.

Estudo de Caso	Período	EQM
1	2014.1	0,0174
2	2013.2	0,0106
	2014.1	0,0191
3	2013.2	0,0158
	2014.1	0,0256

4.3 Limitações da Abordagem Proposta

Para a realização da validação do arcabouço, foi necessária a busca de uma base de dados adequada. Assim, devido à dificuldade de obtenção dos dados, o escopo do arcabouço proposto foi reduzido para fins de validação, de forma a se adaptar à realidade dos dados obtidos. No entanto, um ponto positivo da utilização dos mesmos dados para os grupos de controle e experimental foi o fornecimento de melhor compreensão da acurácia do modelo, uma vez que foi possível confrontar a avaliação tradicional, realizada pelo docente de forma manual, e a avaliação semiautomática, proposta pelo arcabouço.

Uma possível ameaça à validação realizada foi o uso de apenas uma abordagem para a avaliação das mensagens. Acredita-se, entretanto, que a eficiência da utilização desta medida está relacionada ao peso atribuído ao indicador de qualidade e à experiência do próprio docente quanto ao tema abordado no fórum. O refinamento dos conjuntos de palavras-chaves deve ser considerado e realizado de forma constante pelo docente, de forma a refletir melhor sua metodologia de análise das mensagens. A proposta é que, ao longo dos períodos em que a mesma disciplina seja ofertada e o mesmo fórum temático seja aberto à discussão, os conjuntos de palavras-chaves sejam refinados e otimizados, adequando-se melhor ao tema em discussão.

Neste ponto, ainda é necessário citar a dificuldade encontrada pelos docentes para a atribuição de palavras-chaves. No Estudo de Caso 2, por exemplo, que se refere a uma disciplina da área de Ciências Humanas, por possuir um tema que requer a discussão de tópicos mais gerais, o uso de palavras-chaves para implementação do indicador de qualidade das mensagens pode comprometer a acurácia de uma avaliação semiautomática das mensagens. Isto se dá devido ao fato de que se torna difícil prever o comportamento e a natureza das informações trazidas pelo aluno para a discussão. Ao fornecer os conjuntos de palavras, o docente do referido estudo de caso comentou que *"o uso de palavras-chaves é um pouco complicado, porque é muito subjetivo"*. A afirmação do docente evidencia que, apesar de sua experiência – o docente realiza a avaliação de fóruns no tema estudado há 3 (três) semestres, para a mesma disciplina, a natureza do tema da discussão e o que é esperado do aluno deve ser levado em consideração ao se escolher uma (ou mais) técnica(s) para a avaliação das mensagens.

Outros fatores, além daqueles citados, também devem ser levados em consideração, tais como: a presença de plágio; a inclusão de anexos às mensagens, tais como vídeos, animações e imagens; e a categorização das mensagens. Sobre esta última métrica, o docente responsável pelos fóruns de discussão do Estudo de Caso 3 citou que *"(...) comentários de apoio ao colega, que corroborem ou concordem (com o que foi dito), podem agregar valor à discussão"*.

Em relação a plágios e à utilização de textos retirados da *Internet*, o uso de palavras-chaves também se torna limitado. Em todos os fóruns estudados, a discussão foi voltada a um tema proposto em um texto (acrescido, muitas vezes, de um arquivo multimídia, como imagens ou vídeos) e é a partir de sua leitura que os alunos devem realizar comentários e críticas. Um dos docentes afirmou que *"(A maioria dos alunos) lê o texto, faz um 'copiou, colou' e não se posiciona subjetivamente e, principalmente, criticamente. (...) Quando o aluno realiza esse tipo de postagem, a gente procura orientá-lo a ser mais autônomo e, é lógico,*

que ele vai perder certa pontuação, porque você se deteve à postura do texto e não trouxe algo novo, não construiu". Em se tratando da remoção de textos da Internet, o docente em questão complementa: "*Não duvidando da capacidade do aluno, mas surgem dúvidas quando 'jogamos' (o conteúdo das mensagens) em um site de busca e descobrimos de onde ele extraiu aquele texto".* Isto explica os casos extremos, em que a nota atribuída pelo arcabouço foi máxima, enquanto a nota dada pelo docente foi bem menor e vice-versa. Um exemplo em que esta situação pode ocorrer é o envio de uma mensagem que contenha pelo menos uma palavra-chave de cada conjunto de palavras relevantes ao tema fornecido pelo professor, em que o docente identificou plágio. Neste caso, o sistema automaticamente atribuirá a nota máximo, 1 (um), à mensagem. No entanto, embora a mensagem esteja aparentemente completa, por ter citado palavras relevantes ao tema, o professor, ao identificar o plágio, penalizou o aluno, atribuindo a nota mínima, 0 (zero).

Outra limitação do estudo realizado diz respeito ao tratamento de mensagens. Apesar de todas as mensagens terem recebido o mesmo tratamento (*i.e.*, remoção de *stop words*, retirada de sinais de acentuação e pontuação, e redução dos termos aos seus radicais), erros de sintaxe ou coloquialismos não foram tratados.

Foram investigadas outras formas de processamento dos dados de entrada, tais como a utilização de uma base de sinônimos para a adição automática de novos termos ao conjunto de palavras-chaves, o que poderia, eventualmente, aumentar a representatividade do conjunto fornecido pelo docente. No entanto, a língua portuguesa é bastante rica e possui formas e construções linguísticas que podem possuir diferentes significados, a depender do contexto. Assim, esta estratégia necessitaria de um estudo mais aprofundado para determinar uma forma eficiente de adição de sinônimos que não impactasse, de modo negativo, a avaliação das mensagens. Considerar sinônimos ou transformações das palavras-chaves fornecidas pelo docente pode não refletir, necessariamente, a realidade do tema.

4.4 Considerações sobre o Conteúdo do Capítulo

Neste capítulo, descreveu-se a aplicação da abordagem proposta aos estudos de caso, com o intuito de validá-la. Foram discutidas as etapas que compuseram o processo de validação, desde o estudo inicial realizado junto aos docentes do IFPB Virtual, com o intuito de coletar opiniões e informações sobre o processo de avaliação praticado na instituição, até o estudo da acurácia do arcabouço, ao confrontá-lo com dados obtidos a partir da avaliação tradicional realizada pelos docentes.

Como consequência da sondagem de opinião dos docentes, foram reafirmadas as conclusões formuladas a partir da sondagem inicial, descrita no Capítulo 3, assim como a partir das análises de estudos afins documentados na literatura da área. Foi possível, desta forma, confirmar a importância do uso de indicadores de participação no processo de avaliação da participação dos alunos em fóruns de EaD.

Os 3 (três) estudos de caso relatados neste capítulo possibilitaram analisar a acurácia do arcabouço proposto. Os resultados mostraram que, para todos os casos estudados, as notas calculadas por meio de instâncias do arcabouço e as notas atribuídas pelo docente possuem forte correlação. O estudo estatístico permitiu inferir que o uso de instâncias do arcabouço é aplicável.

Analisando-se as estatísticas calculadas, verificou-se que não houve diferenças nos resultados obtidos a partir da utilização do arcabouço e da abordagem tradicional, para 3 (três) das 5 (cinco) turmas analisadas.

Capítulo 5

Considerações Finais e Sugestões para Pesquisas Futuras

Neste capítulo, são formuladas as considerações finais da pesquisa. Na Seção 5.1, são apresentadas as principais conclusões formuladas a partir dos resultados apresentados no capítulo anterior. Nas Seções 5.2 e 5.3, apresentam-se as contribuições da pesquisa e listam-se sugestões para pesquisas futuras, respectivamente.

5.1 Considerações Finais

Nesta dissertação, foi proposto um arcabouço que visa a fornecer o auxílio a docentes de cursos de EaD durante o processo de avaliação da participação dos alunos em fóruns de discussão. Tal processo comumente envolve a adoção de diferentes metodologias e o uso de um ou mais fatores para a mensuração e a atribuição de notas às interações dos alunos.

Ainda não há um consenso, entre pesquisadores e docentes desta modalidade de ensino, sobre o paradigma que melhor representa o processo subjetivo adotado pelo docente para avaliar o diálogo e o comportamento dos alunos nos fóruns. No entanto, é um fato a necessidade de soluções para realizar este processo de avaliação de forma mais ágil, objetiva e imparcial.

Segundo um dos docentes consultados durante o estudo de validação da abordagem proposta:

"Pode perguntar a qualquer docente, que o trabalho mais custoso é a análise das interações no fórum. Se tivesse uma ferramenta que oferecesse um sumário destas interações, seria muito útil. Você evitaria o trabalho de ler as mensagens uma a uma e ponderar a sua avaliação. Talvez fosse até mais justo. Como a

avaliação dos alunos é feita de forma contínua, pode ser que a forma de avaliar sofra alterações ao passar dos dias em que o fórum esteja aberto."

Visando prover uma forma que represente a forma de pensar e agir do docente durante o processo de avaliação da participação, foi proposto um arcabouço semiautomático para avaliar a participação dos alunos em fóruns. O arcabouço foi idealizado e implementado após uma revisão do estado da arte, a qual forneceu uma perspectiva inicial de quais os principais indicadores qualitativos utilizados por docentes na atribuição de notas de participação aos alunos. Os resultados das pesquisas analisadas foram reforçados após a administração de um questionário de sondagem da opinião da comunidade acadêmica interessada.

A validação da abordagem proposta foi composta de 2 (duas) etapas: (i) a análise da opinião de docentes de EaD (representados por docentes da instituição de ensino em que a avaliação foi realizada) a respeito da utilização dos indicadores propostos pelo modelo conceitual; e (ii) o estudo da acurácia da solução proposta, a partir da avaliação da aplicabilidade de instâncias do arcabouço.

Como resultados da primeira etapa, além do delineamento do perfil típico dos docentes participantes do experimento, responsáveis pela avaliação da participação dos alunos, obteve-se a reafirmação da importância da adoção de indicadores de participação no processo de avaliação da participação. Os resultados obtidos corroboraram, também, a necessidade de métodos e ferramentas para a realização deste processo, uma vez que observou-se uma diversidade de metodologias de avaliação utilizadas pelos docentes consultados.

Na segunda etapa, foi analisado o comportamento de instâncias do arcabouço para 3 (três) estudos de caso, cada um representado por fóruns temáticos de áreas distintas. Por meio de uma análise estatística inicial, constatou-se que, para 3 (três) das 5 (cinco) turmas consideradas no estudo experimental, o uso do arcabouço para o cálculo da nota de participação não apresentou diferenças estatisticamente significativas em relação à atribuição

tradicional da nota de participação pelo docente. Uma análise mais específica da acurácia do arcabouço permitiu verificar uma forte correlação entre as notas simuladas e as notas atribuídas pelo docente, em todos os casos investigados (fator de correlação sempre superior a 85%). Além disso, analisando-se o cálculo do erro, pôde-se perceber que o erro médio foi sempre inferior a 3% em todos os casos analisados.

De acordo com os estudos apresentados e com os resultados de pesquisas na área, verificou-se que a avaliação da participação de alunos deve ser realizada sob uma perspectiva multidimensional. O uso de diferentes critérios, representados pelos indicadores de participação, é considerado de extrema relevância para os docentes da área. No entanto, nem sempre é possível utilizá-los, uma vez que a extração destes indicadores pode se tornar uma tarefa onerosa e incômoda para o docente. O uso de uma ferramenta computacional pode fornecer uma forma eficaz para contornar este problema.

No entanto, deve-se também levar em consideração, citando-se um dos docentes participantes do estudo experimental, que "*todos (os docentes) possuem sua estratégia de trabalho, sua didática*". Assim, é preciso considerar, antes de se desenvolver qualquer ferramenta para este fim, as diferentes necessidades dos professores, disciplinas, cursos e instituições. O fornecimento de um recurso de fácil adaptação deve ser considerado, ao se pensar na automatização do processo de avaliação da participação em cursos a distância.

5.2 Contribuições da Pesquisa

As contribuições resultantes desta pesquisa podem ser sumariadas como segue:

- Todo o embasamento teórico resultante da realização de uma revisão de literatura e o agrupamento de pesquisas na área, enfatizando os principais indicadores de participação encontrados, cuja relevância pôde ser corroborada pelo estudo da comunidade acadêmica interessada, poderá servir como suporte para a criação de novos arcabouços, com propósitos

mais específicos, ou refinamento da própria abordagem apresentada nesta dissertação;

- O arcabouço proposto consiste em uma tentativa de fornecimento de uma metodologia flexível para a avaliação da participação em fóruns de EaD. O arcabouço representa, assim, uma forma mais efetiva para auxiliar os docentes na tomada de decisão ao realizar esta avaliação. A realização dos estudos de caso permitiu confirmar a aplicabilidade do arcabouço, ao confrontá-lo com a abordagem tradicional utilizada por docentes que não possuíam acesso a ferramentas de apoio durante a avaliação;
- Uma característica importante da metodologia adotada no processo de avaliação da acurácia do arcabouço foi o emprego de uma técnica semiautomática para a análise semântica das mensagens. Embora ainda seja possível obter melhorias em relação ao método empregado, o estudo realizado mostrou resultados promissores sobre uso de técnicas de Inteligência Artificial neste contexto;
- Segundo relatos dos docentes participantes do estudo experimental, a sugestão de notas de participação pode ser útil no processo da avaliação da participação. Segundo um dos docentes consultados: *"a ideia de fornecer um 'pitaco' para a nota de participação, daria uma certa base, facilitando nosso trabalho de avaliação"*. Durante a realização das entrevistas, foi verificado, que o esforço e tempo empregados para a verificação das mensagens do fórum e na análise dos *logs* de interação, levam os docentes a não considerarem fatores que consideram relevantes para a avaliação. Portanto, foi possível observar o ganho que recomendações de notas (para cada mensagem, bem como a nota final de participação), e a tabulação dos dados quantitativos, apresentariam para o processo de avaliação de forma geral.

5.3 Sugestões para Pesquisas Futuras

Apesar dos estudos de caso, aliados à sondagem da opinião dos docentes

participantes dos experimentos, terem permitido verificar a aplicabilidade do arcabouço proposto, não foi possível realizar generalizações dos resultados encontrados. Assim, como sugestão de pesquisa futura, pode-se citar a utilização do arcabouço em turmas ativas, em que os docentes façam uso de diferentes indicadores de participação para a atribuição da nota. Uma sugestão para promover uma investigação integral do arcabouço proposto é a sua aplicação para o acompanhamento do desempenho dos alunos de forma parcial (*i.e.*, observação dos indicadores de participação ao longo de toda a discussão e não unicamente ao final dela).

Outra sugestão, mais simplificada, é a realização de novos estudos experimentais, com grupos de controle de maior representatividade, *i.e.*, amostras com maior número de instâncias e maior riqueza de interações trocadas entre os alunos. A realização de tal estudo forneceria a possibilidade de melhoria dos resultados da acurácia do arcabouço.

Um dos componentes mais relevantes para a análise de interações em fóruns é a análise semântica das mensagens. Assim, torna-se relevante pesquisas futuras que focalizem o aprimoramento de técnicas destinadas ao processamento textual das mensagens. Uma sugestão é a atribuição de pesos ou otimização da forma de coleta das palavras-chaves que melhor representem o raciocínio do docente ao avaliar o conteúdo das mensagens. A utilização de bases de dados que contenham o material da disciplina ou a construção de sistemas que possuam acesso direto à *Internet* podem auxiliar na detecção de plágio ou mensagens que apresentem conteúdo sem muita relevância, visto que não expressam de forma direta a opinião ou o pensamento do aluno. Esta sugestão pode ser implementada, por exemplo, pelo uso de técnicas de mineração de dados. Outra sugestão é a utilização de redes neurais para o prognóstico da nota de participação final dos alunos, considerando-se um conjunto de dados do mesmo fórum em turmas passadas.

Outra sugestão é a utilização de bases de dados para o aperfeiçoamento das palavras-chaves, *e.g.* a utilização de sinônimos para enriquecimento dos

conjuntos determinados pelo docente. No entanto, é necessário citar que qualquer modificação ou acréscimo à base de palavras que não seja realizado sob a supervisão do docente especialista pode afetar de forma significativa o resultado esperado. Alterações nos conjuntos de palavras-chaves podem não refletir o raciocínio do docente.

É importante reconhecer também que a avaliação do arcabouço foi realizada utilizando instâncias reduzidas, *i.e.*, um número mínimo dos indicadores de participação incorporados pelo arcabouço foram utilizados nos estudos de caso. Consequentemente, não foi possível realizar um estudo aprofundado da relevância que os indicadores de participação possuem no processo de avaliação da participação. A realização de estudos de caso em tempo real, *i.e.*, que façam uso de instâncias mais completas ao longo da discussão pode fornecer recursos para investigar melhor a importância dos fatores e indicadores de participação propostos pelo arcabouço.

Por fim, outra sugestão de pesquisa futura é a investigação de outros fatores que afetam a participação dos alunos, que vão além do escopo dos indicadores de participação listados neste documento, tais como a evasão, o conhecimento por parte dos alunos de que o fórum está sendo avaliado ou não, a presença do docente no fórum e o nível de conhecimento em informática e de familiaridade com o AVA. Como o foco desta pesquisa foi fornecer auxílio ao docente durante a tomada de decisão na avaliação da participação, não foram levados em consideração fatores relacionados aos alunos. Observar o comportamento dos alunos e considerar a participação na visão do discente pode impactar significativamente nos resultados de pesquisas no contexto da avaliação da participação.

Referências Bibliográficas

AGUILAR, D. A. G.; THERON, R.; PENALVO, F.J.G. Reveal the Relationships among Students Participation and Their Outcomes on E-Learning Environments: Case Study. *IEEE 13th International Conference on Advanced Learning Technologies (ICALT'13)*. Beijing, p. 443-447, 2013.

ALSTETE, J. W.; BEUTELL, N. J. Performance indicators in online distance learning courses: a case study of management education. *Quality Assurance in Education*, Vol. 12, No. 1, p. 6-14, 2004.

ANAYA, A. R.; BOTICARIO, J. G. A Data Mining Approach to Reveal Representative Collaboration Indicators in Open Collaboration Frameworks. *Educational Data Mining (EDM'2009)*, p. 210-219, 2009.

ANAYA, A. R.; BOTICARIO, J. G. Ranking learner collaboration according to their interactions. *First annual engineering education conference (EDUCON)*, 2010.

ANAYA, A. R.; BOTICARIO, J. G. Application of machine learning techniques to analyse student interactions and improve the collaboration process. *Expert Systems with Applications*, Vol. 38, No. 2, 2011.

ANDRESEN, M. Asynchronous discussion forums: success factors, outcomes, assessments, and limitations. *Educational Technology & Society*, Vol. 12, No. 1, p. 249-257, 2009.

AZEVEDO, B. F. T. *MineraFórum: um recurso de apoio para análise qualitativa em fóruns de discussão*. 2011. Tese (Doutorado em Informática na Educação) - PPGIE, UFRGS, Porto Alegre.

AZEVEDO, B; F. T.; BEHAR, P. A.; REATEGUI, E. B. Indicadores de relevância para análise de fóruns de discussão. *XXIII Simpósio Brasileiro de Informática na Educação'SBIE2012*. Rio de Janeiro, 2012.

BAEZA-YATES, R. A.; RIBEIRO-NETO, B. *Modern Information Retrieval*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., 1999.

BALDAN, M. A.; MENEZES, C. S. Um Ambiente para Construção de Perfis a Partir de Textos Pessoais. *XXIII Simpósio Brasileiro de Informática na Educação (SBIE'2012)*. Rio de Janeiro, 2012.

BARR, A.; FEIGENBAUM, E. *The Handbook of Artificial Intelligence*. Los Altos, California – EUA: William Kaufmann Inc., Vol. 1-2., 1981.

BAUER, M. W.; GASKELL, G. *Pesquisa qualitativa com texto, imagem e som: um manual prático*. Petrópolis – Brasil: Ed. Vozes, 2002.

BEEBE, R.; VONDERWELL, S.; BOBOC, M. Emerging Patterns in Transferring Assessment Practices from Face-to-face to Online Environments. *Electronic Journal of E-learning*, Vol. 8, No. 1, 2010.

BEER, C.; JONES, D.; CLARK, K. The indicators project identifying effective learning: Adoption, activity, grades and external factors. *Australasian Society for Computers in Learning in Tertiary Education*. Auckland, p. 60-70, 2009.

BIRCH, D.; VOLKOV, M. Students' perceptions of compulsory asynchronous online discussion. *ANZMAC*. Fremantle, Western Australia, 2005.

BLACKBOARD Inc. Blackboard Learning System. Sistema de gestão da aprendizagem. 1997.

BLISS, C. A.; LAWRENCE, B. From Posts to Patterns: A Metric to Characterize Discussion Board Activity in Online Courses. *Journal of Asynchronous Learning Networks*, Vol. 13, No. 2, p. 15-32, 2009.

BLOOM, B. S.; ENGELHART, M. D.; FURST, E. J.; HILL, W. H.; KRATHWOHL, D. R. *Taxonomy of educational objectives: The classification of educational goals*. Handbook I: Cognitive domain. New York: David McKay Company, 1956.

CABALLÉ, S.; DARADOUMIS, T.; XHAFI, F.; JUAN, A. Providing effective feedback, monitoring and Evaluation to on-line collaborative learning discussions. *Computers in Human Behavior*, Vol. 27, p. 1372-1381, 2011.

CACCIAMANI, S.; CESARENI, D.; MARTINI, F.; FERRINI, T.; FUJITA, N. Influence of participation, facilitator styles, and metacognitive reflection on knowledge

building in online university courses. *Computers & Education*, Vol. 58, p. 874-884, 2012.

CALVANI, A.; FINI, A.; MOLINO, M.; RANIERI, M. Visualizing and monitoring effective interactions in online collaborative groups. *British Journal of Educational Technology*, Vol. 41, No. 2, p. 213-226, 2009.

CASPI, A.; CHAJUT, E.; SAPORTA, K. Participation in class and in online discussions: gender differences. *Computers & Education*, Vol. 50, No. 3, p. 718-724, 2008.

CESARENI, D.; MARTINI, F. Costruire conoscenza in un forum universitario. *Rassegna di Psicologia*, Vol. 22, p. 89-112, 2005.

CENSUP. Censo da Educação Superior 2001 - Divulgação dos principais resultados do Censo da Educação Superior 2001. Disponível em: <http://download.inep.gov.br/download/centso/2001/Superior/Sinopse_Superior_2001.zip>. Acesso em: Junho, 2014.

CENSUP. Censo da Educação Superior 2013 - Divulgação dos principais resultados do Censo da Educação Superior 2013. Disponível em: <http://download.inep.gov.br/educacao_superior/centso_superior/apresentacao/2014/coletiva_centso_superior_2013.pdf>. Acesso em: Outubro, 2014.

CHAPMAN, S. SimMetrics: a Java & C # .NET library of similarity metrics. Disponível em: <<http://sourceforge.net/projects/simmetrics/>>. Último acesso: Outubro, 2014.

CHAN, C. K. K.; CHAN, Y.Y. Students' views of collaboration and online participation in Knowledge Forum. *Computers & Education*, Vol. 57, p. 1445-1457, 2011.

CLARK, H.; SCHAEFER, E. Contributing to discourse. *Cognitive Science*, Vol. 13, No. 2, p. 259-294, 1989.

COLLAZOS, C. A.; GUERRERO, L. A.; PINO, J. A.; RENZI, S.; KLOBAS, J.; *et al.* Evaluating collaborative learning processes using system-based measurement.

Educational Technology & Society, Vol. 10, No. 3, p. 257–274, 2007.

DAVIES, J.; GRAFF, M. Performance in e-learning: online participation and student grades. *British Journal of Educational Technology*, Vol. 36, No.4, p. 657-663, 2005.

DOUGIAMAS, M. MOODLE - Modular Object-Oriented Dynamic Learning Environment. Sistema de gestão da aprendizagem. 2001.

DRINGUS, L. P.; ELLIS, T. Building the SCAFFOLD for Evaluating Threaded Discussion Forum Activity: Describing and Categorizing Contributions. *ASEE/IEEE Frontiers in Education Conference*. Savannah – GA, 2004.

DRINGUS, L. P.; ELLIS, T. Temporal transitions in participation flow in an asynchronous discussion forum. *Computers & Education*, Vol. 54, p. 340-349, 2010.

ELIA, M.; CHAMOVITZ, I. CQMsg - Classificador e Qualificador de Mensagens: um instrumento para apoio à avaliação de fóruns temáticos. *XX Simpósio Brasileiro de Informática na Educação*. Santa Catarina – Brasil, 2009.

ELLIS, A. Personality type and participation in networked learning environments., Vol. 40, p. 101–114, 2003.

ERLIN, YUSOF, N.; RAHMAN, A. Integrating Content Analysis and Social Network Analysis for Analyzing Asynchronous Discussion Forum. *International Symposium on Information Technology (ITSim'08)*. 2008.

ERLIN, YUSOF, N.; RAHMAN, A. A. Messages Segmentation of Asynchronous Text-Based Discussion in Social Network Analysis. *II International Conference on Education Technology and Computer (ICETC'2010)*. 2009.

FALAKMASIR, M.; JAFAR, H. Using educational data mining methods to study the impact of virtual classroom in e-learning. *III International Conference on Educational Data Mining*. Pittsburgh, PA – USA. 2010.

GARRISON, D. R.; ANDERSON, T. *E-Learning in the 21st Century*. London – UK: RoutledgeFalmer. 2003.

GARRISON, D. R.; ANDERSON, T.; ARCHER, W. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, Vol. 2, No. 2-3, p. 87-105, 2000.

GARRISON, D. R.; ARBAUGH, J. B. Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, Vol. 10, p. 157-172, 2007.

GHOSH, A.; KLEINBERG, J. Incentivizing Participation in Online Forums for Education. *14th ACM Conference on Electronic Commerce*, p. 525-542, 2013.

GOMES, A. V.; OLIVEIRA, D. S. S.; MEDEIROS, A. C. GENOME: Uma ferramenta avaliativa de participação virtual. *XXIII Congresso Brasileiro de Informática na Educação*. Rio de Janeiro – Brasil. 2012.

GREENLAND, S. J. Using log data to investigate the impact of (a)synchronous learning tools on LMS interaction. *Australasian Society for Computers in Learning in Tertiary Education*. 2011.

HAKKARAINEN, K.; LALLIMO, J.; TOIKKA, S.; WHITE, H. Cultivating collective expertise within innovative knowledge-practice networks. LUDVIGSEN, S.; LUND, A.; RASMUSSEN, I.; SÄLJÖ, R. (Eds.), Learning across sites. *New tools, infrastructures and practices*, pp. 69-85, 2011. Oxon: Routledge.

HARRY, K.; JOHN, M.; KEEGAN, D. *Distance education: New perspectives*. Routledge, 2013.

HAYES-ROTH, F. Knowledge-based systems - The state of the art in the US. *Knowledge Engineering Review*, Vol. 1, p. 18-27, 1985.

HE, W. Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, Vol. 29, p. 90-102, 2013.

HENRI, F. Computer Conferencing and Content Analysis. In Kaye, A. R. (Ed), Collaborative learning through computer conferencing. *The Najaden Papers*, p. 117-136, 1992. Berlin.

HRASTINSKI, S. What is online learner participation? A literature review. *Computers & Education*, Vol. 51, p. 1755–1765, 2008.

HRASTINSKI, S. A theory of online learning as online participation. *Computers & Education*, Vol 52, p. 78-82, 2009.

HRASTINSKI, S.; JALDEMARK, J. How and why do students of higher education participate in online seminars? *Education & Information Technologies*, Vol. 17, p. 253–271, 2012.

JALDEMARK, J.; LINDBERG, J. O.; OLOFSSON, A. D. Sharing the distance or a distance shared: Social and individual aspects of participation in *ICT-supported distance-based teacher education*. CHAIB, M. e SVENSSON, A. K.; *ICT in Teacher Education: Challenging Prospects*, pp. 142–160, 2006.

JYOTHI, S.; MCAVINIA, C.; KEATING, J. A visualisation tool to aid exploration of students' interactions in asynchronous online communication. *Computers & Education*, Vol. 58, p. 30–42, 2012.

KAY, R. H. Developing a comprehensive metric for assessing discussion board effectiveness. *British Journal of Educational Technology*, Vol. 37, No. 5, p. 761–783, 2006.

KERR, D.; ISELI, M.; MOUSAVI, H. Automatic Short Essay Scoring Using Natural Language Processing to Extract Semantic Information in the Form of Propositions. *GATES3 / OPP1003019 / CRESST REPORT*. University of California - Los Angeles, 2013.

KHANNA, S.; KAUSHIK, A.; BARNELA, M. Expert Systems Advances in Education. *National Conference on Computational Instrumentation*, p. 109-112. India, 2010.

LACHI, R. L.; ROCHA, H. V.; OEIRAS, J. Y. Y. Avaliação de cursos a distância: uso de indicadores para assegurar qualidade. XVII Simpósio Brasileiro de Informática na Educação, 2006.

LIN, G.; Social presence questionnaire of online collaborative learning: Development and validity. *Association for Educational Communications and*

Technology, 2004.

LIN, F.; HSIEH, L.; CHUANG, F. Discovering genres of online discussion threads via text mining. *Computers & Education*. Vol. 52, No. 2, p. 481-495, 2009.

MACFADYEN, L. P.; DAWSON, S. Mining LMS data to develop an 'early warning system' for educators: A proof of concept. *Computers & Education*. Vol. 54, No 2, p. 588-599, 2010.

MARTIN, J. R. English text: Systems and structure. Amsterdam: Benjamin Press. 1992.

MAZZOLINI, M.; MADDISON, S. When to jump in: the role of the instructor in online discussion forums. *Computers & Education*. Vol. 49, No. 2, p. 193-213, 2007.

MCKERLICH, R.; RIIS, M.; ANDERSON, T.; EASTMAN, B. Student Perceptions of Teaching Presence, Social Presence, and Cognitive Presence in a Virtual World. *MERLOT Journal of Online Learning and Teaching*, Vol. 7, No. 3, p. 324-336, 2011.

MEC - Ministério da Educação, Secretaria de Educação a Distância. Indicadores de qualidade para cursos de graduação a distância. Brasília, 2007.

MOORE, G. Three types of interaction. *The American Journal of Distance Education*. Vol. 3, No. 2, p. 1-6, 1989.

MOORE, M. G. Recent contributions to the theory of distance education. *Open Learning*. Vol. 5, p. 10-15, 1990.

NANDI, D.; CHANG, S.; BALBO, S. A conceptual framework for assessing interaction quality in online discussion forums. Same places, different spaces: *Australasian Society for Computers in Learning in Tertiary Education*. Auckland, 2009.

NANDI, D.; HAMILTON, M.; CHANG, S.; BALBO, S. Evaluating quality in online asynchronous interactions between students and discussion facilitators. *Australasian Journal of Educational Technology*. Vol. 28, No. 4, p. 684-702, 2012.

NEJAD, N. Z.; HAJIHEYDARI, N. An investigation into factors influencing learners'

participation in E-learning. *VI National and III International conference of e-Learning and e-Teaching (ICELET'2012)*. 2012.

NISTOR, N. Stability of attitudes and participation in online university courses: Gender and location effects. *Computers & Education*, Vol. 68, p. 284-292, 2013.

NISTOR, N. e NEUBAUER, K. From participation to dropout: Quantitative participation patterns in online university courses. *Computers & Education*, Vol. 55, p. 663-672, 2010.

OLIVEIRA JR.; R. L. e ESMIN, A. A. A. Monitoramento automático de mensagens de fóruns de discussão usando técnica de classificação de texto semi-supervisionado. *XXIII SBIE - Simpósio Brasileiro de Informática na Educação*. Rio de Janeiro – Brasil, 2012.

PETERS, O. Distance Teaching and Industrial Production: A Comparative Interpretation in Outline. *Distance Education: International Perspectives*. Londres/Nova Iorque: Croomhelm/St. Martin. 1983.

PORTER, M. F. An Algorithm for Suffix Stripping. *Readings in information retrieval*. Vol. 14, No. 3, p. 130-137, 1980.

PTSTEMMER – A stemming toolkit for the Portuguese language. Version 2.0. Disponível em: <<http://www.code.google.com/p/ptstemmer/>>. Acesso em: Outubro, 2014.

RABBANY, R.; TAKAFFOLI, M.; ZAIANE, O. R. Analyzing Participation of Students in Online Courses Using Social Network Analysis Techniques. *IX International Conference on Educational Data Mining*. Netherlands, 2011.

RAGHAVAN, P.; CATHERINE, R.; IKBAL, S.; KAMBHATLA, N.; MAJUMDAR, D. Extracting problem and resolution information from online discussion forums. *Management of Data*. p. 77, 2010.

ROMERO, C.; LÓPEZ, M.; LUNA, J.; VENTURA, S. Predicting students' final performance from participation in on-line discussion forums. *Computers & Education*. Vol. 68, p. 458-472, 2013.

RYABOV, I. The Effect of Time Online on Grades in Online Sociology Courses. *MERLOT Journal of Online Learning and Teaching*, Vol. 8, No 1, 2012.

SALAZAR, F. A. S. R. *Um estudo sobre o papel de medidas de similaridade em visualização de coleções de objetos*. Dissertação de Mestrado. USP, São Paulo. 2012.

SCHRIRE, S. A model for evaluating the process of learning in asynchronous computer conferencing. *Journal of Instruction Delivery Systems*. Vol. 17, No. 1, p. 6–12, 2003.

SEETHAMRAJU, R. Effectiveness of Using Online Discussion Forum for Case Study Analysis. *Hindawi Publishing Corporation - Education Research International*, Vol. 2014, 2014.

SELF, J. A. Dormobilea: Vehicle for metacognition. T. W. Chan & J. A. Self (Eds.), *Emerging computer technologies in education*, p. 1–20, 1994.

SHAW, R. A study of the relationships among learning styles, participation types, and performance in programming language learning supported by online forums. *Computers & Education*, Vol. 58, p. 111–120, 2012.

SHIU, S. e PAL, S. K. *Foundations of Soft Case-Based Reasoning*. Wiley Series on Intelligent Systems, Wiley-Interscience, 1ª Ed., 2004.

SIGEL, S. e CASTELLAN, Jr.; N. J. *Nonparametric Statistics for The Behavioral Sciences*. McGraw-Hill, 1988.

SWAN, K.; SHEA, P.; RICHARDSON, J.; ICE, P.; GARRISON, D. R.; CLEVELAND-INNES, M.; *et al.* Validating a measurement tool of presence in online communities of inquiry. *E-Mentor*, Vol. 2, No. 24, p. 1–12, 2008.

TELEDUC (2002). Disponível em: <<http://teleduc.nied.unicamp.br>>. Acesso em: Novembr, 2014.

TIRNEY, P. A qualitative analysis framework using natural language processing and graph theory. *The International Review of Research in Open and Distance Learning*. Vol. 13, No. 5, p. 173-189, 2012.

TOBARRA, L.; ROBLES-GÓMEZ, A.; ROS, S.; HERNÁNDEZ, R.; CAMINERO, A. C. Analyzing the students' behavior and relevant topics in virtual learning communities. *Computers in Human Behavior*, Vol. 31, p. 659–669, 2014.

UZUNER, S. Educationally valuable talk: A new concept for determining the quality of online conversations. *Journal of Online Learning and Teaching*, Vol. 3, No. 4, p. 400–410, 2007.

VENÂNCIO, V.; LOPES, R. D. Avaliação da Aprendizagem em Sistemas Interativos: uma revisão comparativa focada no SBIE, WIE e WAvalia. *XXIV Simpósio Brasileiro de Informática na Educação (SBIE'2013)*. 2013.

VONDERWELL, S.; ZACHARIAH, S. Factors that influence participation in online learning. *Journal of Research on Technology in Education*, Vol. 38, No. 2, p. 213–230, 2005.

WEE, M.; ABRIZAH, A. An Analysis of an Assessment Model for Participation in Online Forums. *Computer Science and Information Systems*. Vol. 122, No. 1, p. 121-140, 2011.

WENGER, E. *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press. 1998.

WISE, A. F.; PERERA, N.; HSIAO, Y.; SPEER, J.; MARBOUTI, F. Microanalytic case studies of individual participation patterns in an asynchronous online discussion in an undergraduate blended course. *Computers & Education*. Vol. 15, p. 108-117, 2012.

WU, Y. B.; CHEN, X. Assessing Student Learning with Automated Text Processing Techniques. *Journal of Asynchronous Learning Networks*. Vol. 9, No. 3, 2005.

YOO, J.; KIM, J. Predicting Learner's Project Performance with Dialogue Features in Online Q&A Discussions. *Intelligent Tutoring Systems Conference (ITS'12)*. 2012.

YUKSELTURK, E. An Investigation of Factors Affecting Student Participation Level in an Online Discussion Forum. *The Turkish Online Journal of Educational*

Technology. Vol. 9, No. 2, 2010.

ZHANG, J.; SCARDAMALIA, M.; LAMON, M.; MESSINA, R.; REEVE, R. Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Educational Technology Research & Development*. Vol. 55, No. 2, p. 117–145, 2007.

ZORRILLA, M.; GARCÍA, D.; ÁLVAREZ, E. An Approach to Measure Student Activity in Learning Management Systems. *International Conference on Computer Supported Education (CSEDU'10)*, p. 21-28, 2010.

Apêndice A

Questionário sobre Técnicas para Avaliação da Participação de Alunos em EaD

Este questionário faz parte de uma pesquisa de mestrado, realizada pela aluna do curso de pós-graduação em Ciência da Computação, Danielle Chaves. O objetivo deste questionário é a identificação das principais métricas utilizadas pelos professores de Educação a Distância (EaD) durante a avaliação de alunos nesta modalidade de ensino. O questionário apresenta, em uma etapa inicial, um conjunto de perguntas mais gerais, visando traçar um perfil dos respondentes. Na segunda e última etapa, são feitos alguns questionamentos a respeito de métricas específicas utilizadas para avaliar alunos em fóruns de discussão. Ao final, por favor, se achar necessário, acrescente algum comentário sobre o questionário, sugerindo métricas que não foram apresentadas ou fornecendo comentários sobre a avaliação da participação de alunos em EaD de modo geral ou, ainda, se preferir, envie um e-mail para daniellemedeiros@copin.ufcg.edu.br. Antecipadamente, agradecemos a sua colaboração!

* Obrigatório

Primeira Etapa

1. Qual é sua área de atuação? *

Marcar apenas uma oval.

Ciências Exatas e Tecnológicas

Ciências Biológicas e da Saúde

Ciências Agrárias

Ciências Humanas e Sociais

Outro: _____

2. Por favor, informe a(s) instituição(ões) a(s) qual(is) você está vinculado:

3. Qual é sua experiência, em anos, como docente de cursos de EaD? *

Marcar apenas uma oval.

- Menos de 1
- 1 - 2
- 3 - 4
- 5 - 6
- Mais de 6

4. Qual(is) recurso(s) você tipicamente utiliza para avaliar a participação dos alunos ao longo do curso? *

Marque todas que se aplicam.

- Funcionalidades oferecidas por ferramenta de gerenciamento de aprendizagem (frequência de acesso, uso das funcionalidades, etc.)
- Atividades realizadas ao longo do curso
- Fóruns de discussão
- Envio de mensagens para professor e/ou monitor
- Outro: _____

5. Em relação aos fóruns de discussão, você considera importante a verificação das mensagens enviadas pelos alunos como forma de avaliar a sua participação. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

Segunda Etapa

Opine a respeito das seguintes afirmações, indicando o grau de importância de cada item para sua avaliação da participação dos alunos em fóruns de discussão. Por exemplo, na Afirmação 1, o grau de importância refere-se à relevância do uso de categorias atribuídas às mensagens como meio de avaliar a participação do aluno.

1. Considero o uso de categorias atribuídas às mensagens (e.g., "Esclarecimento", "Questionamento", "Exemplo", etc.), uma forma efetiva de resumir seu conteúdo. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

1.1. Importância: *

Marcar apenas uma oval.

- | | | | | | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| <input type="radio"/> |

2. Entendo ser desnecessário fornecer informações sobre o que é esperado do aluno em relação ao conteúdo das mensagens (e.g., o que deve ou não ser postado, como deve ser o seu comportamento em relação aos demais colegas). *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

2.1. Importância: *

Marcar apenas uma oval.

- | | | | | | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| <input type="radio"/> |

3. Creio que, ao início da discussão, deve-se fornecer critérios ou padrões para postagens de mensagens, de forma a direcionar melhor os alunos e manter o foco no assunto em discussão.

* *Marcar apenas uma oval.*

- Concordo totalmente

- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

3.1. Importância: *

Marcar apenas uma oval.

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>										

4. O impacto das mensagens enviadas pelo aluno em relação aos demais colegas é um fator contribuinte para sua participação, uma vez que, quanto mais relevante for o conteúdo das mensagens postadas por ele, maior pode ser a repercussão entre os demais alunos. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

4.1. Importância: *

Marcar apenas uma oval.

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>										

5. Acho importante analisar a iniciativa do aluno, de acordo com a frequência de envio de mensagens ao fórum e sua atitude em buscar novas informações de forma a enriquecer a discussão. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

5.1. Importância: *

Marcar apenas uma oval.

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>										

6. Avaliar a qualidade das mensagens é necessário, uma vez que ela representa a riqueza de informações trazida na mensagem postada pelo aluno. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

6.1. Importância: *

Marcar apenas uma oval.

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>										

7. Verificar a quantidade de mensagens que o aluno definiu como lidas em um tópico de discussão é algo irrelevante para a avaliação de sua participação. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

7.1. Importância: *

Marcar apenas uma oval.

0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>										

8. O tamanho das mensagens é algo que impacta na participação do aluno ao longo da discussão. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

8.1. Importância: *

Marcar apenas uma oval.

- | | | | | | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| <input type="radio"/> |

9. A quantidade de mensagens respondidas pelo aluno é relevante na determinação do nível de sua participação durante a discussão. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

9.1. Importância: *

Marcar apenas uma oval.

- | | | | | | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| <input type="radio"/> |

10. É válida, também, a verificação do total de mensagens enviadas pelo aluno (em discussões iniciadas e em resposta às mensagens dos colegas). *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente

Discordo totalmente

10.1. Importância: *

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

11. É importante verificar a regularidade de envio de mensagens, de forma a definir o nível de atividade do aluno ao longo de toda a discussão. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

11.1. Importância: *

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

12. Não acho necessária a utilização de uma limitação de tempo para o encerramento da discussão, a partir da atribuição de datas de início e fim da discussão. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

12.1. Importância: *

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

13. A utilização de palavras-chaves permite que o professor defina um conjunto de termos relevantes ao tópico, antes de iniciar a discussão, e as utilize como forma de verificar se os argumentos contidos em cada mensagem são relevantes ou não. *

Marcar apenas uma oval.

- Concordo totalmente
- Concordo parcialmente
- Não concordo, nem discordo
- Discordo parcialmente
- Discordo totalmente

13.1. Importância: *

Marcar apenas uma oval.

- | | | | | | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| <input type="radio"/> |

Por favor, adicione comentários, caso os considere necessários.

33. Caso deseje, informe seu e-mail, a fim de receber o resultado da pesquisa.

Apêndice B

Documentação Final de Parceira entre LIA-UFMG e IFPB Virtual



DECLARAÇÃO

Eu, Danielle Chaves de Medeiros, aluna de mestrado do Programa de Pós-Graduação em Ciência da Computação, matrícula de número 0112015808-15M, declaro para os devidos fins que mantereirei as informações acessadas nos servidores de cursos EaD do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) sob os princípios de sigilo e confidencialidade. Atesto ainda, que os dados serão utilizados apenas para validar a pesquisa e não serão repassados à terceiros sem autorização prévia do IFPB.

Campina Grande, 26 de março de 2014.

Danielle Chaves de Medeiros

Danielle Chaves de Medeiros



DECLARAÇÃO

Eu, Joseana Macedo Fechine Régis de Araújo, orientadora de doutorado da aluna Alana Marques de Moraes (matrícula de número 0112015801-1D) e da aluna de mestrado Danielle Chaves de Medeiros (matrícula de número 0112015808-15M) no Programa de Pós-Graduação em Ciência da Computação (COPIN), declaro para os devidos fins que mantereí as informações acessadas nos servidores de cursos EaD do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) sob os princípios de sigilo e confidencialidade. Atesto ainda, que os dados serão utilizados apenas para validar as pesquisa e não serão repassados à terceiros sem autorização prévia do IFPB.

Campina Grande, 26 de março de 2014.

A handwritten signature in black ink, appearing to be 'AJ', is positioned above a horizontal line that spans the width of the signature area.

Joseana Macedo Fechine Régis de Araújo



**MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA
PARAÍBA**

CARTA DE ANUÊNCIA

Declaramos, para os devidos fins, que aceitaremos as pesquisadoras Alana Marques de Moraes e Danielle Chaves de Medeiros, pós-graduandas em Ciência da Computação pela Universidade Federal de Campina Grande (UFCG), nesta instituição para desenvolverem seus respectivos projetos de pesquisa, sob orientação da Profa. Joseana Macêdo Fachine Régis de Araújo, cujo objetivo é estudar e desenvolver métodos e metodologias para a avaliação de alunos em Educação a Distância, com o auxílio de técnicas de Inteligência Artificial, no IFPB.

Ressaltamos que os dados coletados serão mantidos em absoluto sigilo, de acordo com a Resolução do Conselho Nacional de Saúde (CNS/MS) 466/2012, que trata da Pesquisa envolvendo Seres Humanos. Salientamos ainda que tais dados só poderão ser utilizados tão somente na realização deste estudo. Assim, a aceitação está condicionada ao cumprimento das pesquisadoras aos requisitos da resolução citada anteriormente, comprometendo-se as pesquisadoras a utilizarem os dados e materiais coletados exclusivamente para os fins da pesquisa.

João Pessoa, em 18 de agosto de 2014.



Prof. João Batista de Oliveira Silva
Reitor do IFPB

Apêndice C

Questionário sobre Avaliação da Participação de Alunos em Fóruns de EaD no IFPB Virtual

Este questionário faz parte de uma pesquisa sobre a avaliação da participação de alunos em fóruns de discussão, realizada pela aluna de mestrado, Danielle Chaves, do curso de Computação - UFCG. O objetivo de sua administração é o delineamento do perfil dos docentes e a coleta de informações sobre o processo de avaliação realizado pelos mesmos em fóruns temáticos nos cursos disponibilizados pelo IFPB Virtual. Para seu preenchimento, é necessário que o docente tenha conhecimento dos conceitos de fatores de influência na participação do aluno e indicadores de participação, de forma que ele possa expressar sua opinião de maneira clara e precisa. Desde já, agradeço sua contribuição!

Primeira Etapa – Delineamento do Perfil dos Docentes

Esta seção apresenta uma série de questionamentos que visam traçar o perfil geral dos docentes a distância do IFPB Virtual.

1. Você é do sexo:

Marcar apenas uma oval.

Feminino

Masculino

2. Indique sua faixa etária:

Marcar apenas uma oval.

Entre 20 e 30 anos

Entre 31 e 40 anos

Entre 41 e 50 anos

Mais de 51 anos

3. Você é:

Marcar apenas uma oval.

Destro

Canhoto

Ambidestro

4. Você utiliza corretivos visuais (óculos ou lentes de contato)?

Marcar apenas uma oval.

Sim

Não

5. Indique seu nível de conhecimento de Sistemas Computacionais:

Marcar apenas uma oval.

Principiante

Intermediário

Experiente

6. Em qual(is) plataforma(s) computacional(ais) você possui experiência de uso?

Marcar todas as que se aplicam.

Smartphone

Tablet

PC

Notebook

Outro(s): _____

7. Em qual(is) sistema(s) operacional(is) você possui experiência de uso?

Marcar todas as que se aplicam.

GNU/Linux

Microsoft Windows

MAC OS

Outro(s): _____

8. Indique seu grau de escolaridade:

Marcar apenas uma oval.

- Nível Técnico
- Graduação
- Pós Graduação (Especialização)
- Pós Graduação (Mestrado)
- Pós Graduação (Doutorado)

9. Qual é sua experiência, em anos, como docente de cursos de EaD?

Marcar apenas uma oval.

- Menos de 1
- 1 - 2
- 3 - 4
- 5 - 6
- Mais de 6

10. Em qual(is) plataforma(s) de ensino a distância você possui experiência de uso?

Marcar todas as que se aplicam.

- Blackboard
- Moodle
- Piazza
- TelEduc
- Outra(s): _____

11. Informe com que frequência você utiliza a plataforma de ensino a distância de sua instituição:

Marcar apenas uma oval.

- Diariamente, várias vezes ao longo do dia
- Diariamente, uma ou duas vezes ao longo do dia
- Várias vezes ao longo da semana

De uma a três vezes ao longo da semana

12. Em relação aos fóruns de discussão, você considera importante a verificação das interações entre os alunos como forma de avaliar a sua participação.

Marcar apenas uma oval.

Concordo Totalmente

Concordo Parcialmente

Nem concordo, nem discordo

Discordo parcialmente

Discordo totalmente

Segunda Etapa - Sondagem de Opinião Sobre a Avaliação da Participação em Fóruns

Esta seção contém questionamentos acerca dos indicadores considerados por você ao avaliar a participação do aluno no fórum em questão. Adicionalmente, pretende-se verificar sua opinião sobre os indicadores de participação considerados para este fim. As próximas questões estão relacionadas ao indicador qualitativo de avaliação das mensagens do aluno.

1. Você considera o tempo gasto pelo aluno no fórum ou no próprio Moodle ao avaliar sua participação?

Marcar apenas uma oval.

Sim

Não

Para cada item abaixo, julgue a sua importância, de forma geral, para a avaliação da participação do aluno em fóruns.

Quanto maior for o valor atribuído, maior será a importância dada ao indicador. Marcar apenas uma oval.

1.1. Tempo gasto pelo aluno na leitura das mensagens do fórum.

0 1 2 3 4 5 6 7 8 9 10

1.2. Tempo gasto pelo aluno acessando o material da disciplina.

0 1 2 3 4 5 6 7 8 9 10

1.3. Tempo gasto pelo aluno na composição de mensagens.

0 1 2 3 4 5 6 7 8 9 10

1.4. Tempo gasto pelo aluno na edição de suas mensagens já postadas.

0 1 2 3 4 5 6 7 8 9 10

2. Você considera o total de contribuições realizadas pelo aluno ao avaliar sua participação?

As contribuições do aluno estão relacionadas ao total de suas mensagens já enviadas. Marcar apenas uma oval.

- Sim
 Não

Para cada item abaixo, julgue a sua importância, de forma geral, para a avaliação da participação do aluno em fóruns.

Quanto maior for o valor atribuído, maior será a importância dada ao indicador. Marcar apenas uma oval.

2.1. Total de mensagens enviadas pelo aluno ao fórum.

0 1 2 3 4 5 6 7 8 9 10

2.2. Total de tópicos (*threads*) iniciadas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

2.3. Total de respostas enviadas pelo aluno às mensagens dos demais.

0 1 2 3 4 5 6 7 8 9 10

2.4. Total de respostas às mensagens enviadas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

2.5. Total de mensagens lidas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

2.6. Total de mensagens editadas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

3. Ao avaliar a participação do aluno no fórum, você considera a sua frequência de acesso?

Marcar apenas uma oval.

Sim

Não

Para cada item abaixo, julgue a sua importância, de forma geral, para a avaliação da participação do aluno em fóruns.

Quanto maior for o valor atribuído, maior será a importância dada ao indicador. Marcar apenas uma oval.

3.1. Frequência de acesso ao Moodle.

0 1 2 3 4 5 6 7 8 9 10

3.2. Frequência de acesso ao fórum de discussão.

0 1 2 3 4 5 6 7 8 9 10

4. Você considera a opinião do aluno no processo de avaliação de sua participação?

Marcar apenas uma oval.

Sim

Não

Julgue a importância do item abaixo, de forma geral, para a avaliação da participação do aluno em fóruns.

Quanto maior for o valor atribuído, maior será a importância dada ao indicador. Marcar apenas uma oval.

4.1. Notas atribuídas pelo aluno às demais mensagens.

0 1 2 3 4 5 6 7 8 9 10

5. No processo de verificação da participação, você avalia as mensagens enviadas pelo aluno?

Marcar apenas uma oval.

- Sim
 Não

5.1. Caso tenha respondido positivamente o questionamento anterior, qual a métrica que v. utiliza no processo de avaliação das mensagens?

Marcar todas as que se aplicam.

- Média das notas atribuídas às mensagens enviadas pelo aluno.
 Nota atribuída à última mensagem enviada pelo aluno.
 Nota atribuída à primeira mensagem enviada pelo aluno.
 Nota atribuída à mensagem mais relevante enviada pelo aluno.
 Outra: _____

Para cada item abaixo, julgue a sua importância, de forma geral, para a avaliação da participação do aluno em fóruns.

Quanto maior for o valor atribuído, maior será a importância dada ao indicador. Marcar apenas uma oval.

5.2. Avaliação das mensagens enviadas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

5.3. Tamanho das mensagens enviadas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

5.4. Categorias das mensagens enviadas pelo aluno.

0 1 2 3 4 5 6 7 8 9 10

Apêndice D

Formulário de Entrevista com Docentes

Considerações Iniciais

Os avisos a serem dados antes da entrevista são os seguintes, preferencialmente nessa ordem:

- Esta entrevista faz parte da pesquisa de pós-graduação em Ciência da Computação da aluna Danielle Chaves, oferecida pela UFCG;
- O objetivo desta pesquisa é criar um arcabouço para modelagem da avaliação da participação de alunos em fóruns de discussão de cursos de EaD. Esse arcabouço visa proporcionar uma forma de apoiar a decisão do docente ao avaliar a participação do aluno, fornecendo um mecanismo mais efetivo para a mensuração da qualidade das interações;
- O objetivo desta entrevista é coletar informações de docentes de cursos a distância do IFPB Virtual;
- O arcabouço para a modelagem da avaliação da participação de alunos em EaD deve ser previamente explicado, de forma a esclarecer a metodologia utilizada, assegurando-se a compreensão do docente a respeito da adaptação do modelo conceitual;
- Os dados coletados nesta entrevista serão utilizados para delinear um perfil do docentes em EaD e verificar as suas visões e necessidades ao avaliar a participação dos alunos por meio de suas interações em fóruns;
- A sessão da entrevista será gravada;
- Os dados serão analisados em conjunto, sem identificar os respondentes por nome, as menções serão feitas apenas por curso e/ou disciplina.

Formulário de Entrevistas

Entrevistado N°

Data da Entrevista:

Sexo:

Experiência:

Disciplina:

Fórum:

1. Descreva o tema discutido no fórum.
2. A discussão no fórum possuía algum direcionamento específico, como discussão de definições ou conceitos, apresentação de exemplos dentro de um tema específico, etc.?
3. Você, ao longo do acompanhamento da discussão, ofereceu diretrizes a serem seguidas pelos alunos? Em caso afirmativo, especifique-as.
4. Quais os fatores analisados por você ao avaliar a participação dos alunos no fórum? (e.g., total de contribuições, qualidade das interações, frequência de acesso ao ambiente, comportamento dos alunos)
5. Caso tenha utilizado a avaliação das mensagens como critério de avaliação, qual a abordagem utilizada para verificar o conteúdo das mensagens? Como foi realizado o cálculo final para atribuição da nota a este critério? média das notas atribuídas às mensagens enviadas pelo aluno; nota atribuída à última mensagem enviada pelo aluno; nota atribuída à primeira mensagem enviada pelo aluno; nota atribuída à mensagem mais relevante enviada pelo aluno)
6. Você identificou perfis de alunos ao analisar as interações no fórum? (e.g., alunos mais participativos, alunos que contribuíram de forma mais significativa para a discussão, alunos que apenas observaram a discussão, mas não contribuíram)
7. Você poderia identificar conjuntos de palavras-chaves ou expressões relevantes ao tema da discussão?
8. Analisando a estrutura do modelo conceitual apresentado, identifique, se presentes, os fatores que você considerou para avaliar os alunos no fórum, bem como os indicadores utilizados. Adicionalmente, indique o peso (em porcentagem) que você atribuiu a cada fator.
9. Você acredita que a aplicação do modelo seria interessante durante a avaliação da participação no fórum? Aponte sugestões ou observações, caso considere necessário, a respeito da solução apresentada.

Apêndice E

Análise de Funções de Similaridade para Verificação do Conteúdo de Mensagens em Fóruns de Discussão

Danielle C. Medeiros¹, José Eustáquio R. de Queiroz¹, Joseana M. F. R. Araújo¹

¹Departamento de Sistemas e Computação
Universidade Federal de Campina Grande (UFCG) – Campina Grande, PB – Brasil
daniellemedeiros@copin.ufcg.edu.br, {rangel, joseana}@dsc.ufcg.edu.br

Abstract. *In distance learning courses, evaluating the effectiveness and the efficiency of learning activities, such as students' participation throughout the courses, can be time consuming and susceptible to the strong influence of human subjectivity. In such a task, AI techniques can be useful to assist the teacher. A pilot study of similarity functions applied to this context is presented in this article, in order to determine its applicability for the evaluation of messages in discussion forums. Although there is a need for further studies to determine evaluation indicators, it was statistically validated the effectiveness and efficiency of the use of similarity measures.*

Resumo. *Em cursos de EAD, avaliar a eficácia e a eficiência das atividades de aprendizagem, tal como a participação dos alunos ao longo do curso, pode ser uma tarefa onerosa e passível de forte influência da subjetividade humana. Técnicas de IA podem ser úteis para auxiliar o professor na realização de tarefas dessa natureza. Neste artigo, é apresentado um estudo da aplicação de funções de similaridade ao contexto de EaD, com o propósito de analisar sua aplicabilidade ao apoio à avaliação de mensagens em fóruns de discussão. Embora exista a necessidade de estudos mais aprofundados sobre a definição de indicadores de avaliação aplicáveis a esse contexto, a eficácia e a eficiência da aplicação de medidas de similaridade foi validada estatisticamente.*

1. Introdução

No Brasil, há diversos estudos com foco na avaliação da aprendizagem nas diferentes modalidades de ensino. Embora a quantidade de pesquisas neste sentido tenha aumentado, ainda há certa escassez de estudos, tanto empíricos, quanto teóricos, no âmbito da Educação a Distância (EaD) (Leite *et al.* 2011). Em EaD, quando não é possível o contato presencial entre professor e alunos, os meios para realizar a avaliação do aluno, em especial sua participação ao longo do curso, encontram-se em funcionalidades disponíveis em Ambientes Virtuais de Aprendizagem (AVA). O Moodle (Dougiamas 2001) e o Blackboard (Blackboard Inc. 1997) são exemplos de AVA

amplamente utilizados em cursos *online*.

Buscando prover a interação entre os alunos, essas ferramentas disponibilizam algumas funcionalidades, tais como fóruns de discussão assíncronos, que possibilitam, a partir de um tema determinado pelo professor, a interação entre os alunos, que compartilham ideias e conhecimentos. Os participantes do fórum têm acesso ao histórico das mensagens postadas e recebidas, podendo revisá-las, caso a mensagem seja de sua autoria, acrescentando comentários às mensagens postadas pelos colegas. Assim, esta funcionalidade tornou-se um elemento importante para promover o compartilhamento de experiências, tão importante em cursos não presenciais (Calvani *et al.* 2009).

No entanto, um dos obstáculos impostos pelo uso de fóruns de discussão como forma de avaliação da participação dos alunos é o grande dispêndio de tempo por parte do professor, dependendo do fluxo de mensagens presentes no fórum. Muitas vezes, há um grande número de mensagens a serem avaliadas pelo docente (Andresen 2009), o que pode tornar tal atividade impraticável.

Em sua maioria, os AVA comumente utilizados oferecem informações estatísticas, tais como a quantidade de mensagens que um determinado discente enviou ao fórum ou o número de comentários deixados em resposta a mensagens de colegas. A utilização deste tipo de informação, em geral, não é suficiente para avaliar o nível de atividade dos alunos, já que os dados estatísticos acabam levando em consideração apenas a quantidade e não a qualidade das postagens (Erlin *et al.* 2009). Uma forma de minimizar este problema consiste na adoção de técnicas de automatização das formas de avaliação, visando a fornecer meios para auxiliar o professor no monitoramento das atividades dos alunos e nos processos decisórios, ao avaliar a participação de cada aluno.

Para automatizar atividades de avaliação em cursos *online*, faz-se necessária, inicialmente, a realização de um estudo dos indicadores comumente utilizados pelos professores. Hrastinski (2008) sugere como indícios de atividade participativa dos alunos a utilização do tempo de acesso ao ambiente virtual, da frequência de leitura de postagens e da quantidade de mensagens postadas. Andresen (2009) menciona alguns estudos que tratam do sucesso e efetividade da adoção desses indicadores na avaliação da participação em cursos *online*. No entanto, esses indicadores não refletem a qualidade das informações discutidas nem fornecem meios para verificar se as postagens se mantêm no âmbito do assunto proposto.

Um dos grandes obstáculos para a verificação automática do conteúdo em fóruns é a subjetividade associada à avaliação realizada pelo professor. Convencionalmente, o docente é capaz de verificar, por exemplo, se o conteúdo das mensagens é pertinente ao tópico em discussão, se as informações trazidas apresentam argumentos originais ou se seu conteúdo foi extraído integralmente da *web*. Neste contexto, o uso de técnicas de Inteligência Artificial (IA) pode constituir-se uma ferramenta adequada para a resolução de tais problemas.

No âmbito da Inteligência Artificial, existem várias estratégias para a representação do conhecimento. Uma das técnicas aplicáveis à verificação textual é o cálculo do grau de similaridade entre termos, que possibilita verificar se há relação entre dois conjuntos de dados. Dentre as funções de similaridade mais populares estão: a Similaridade dos Cossenos, a Distância Euclidiana, o Índice de Similaridade de Jaccard e os algoritmos Dice's Coefficient e Overlap (Tan *et al.* 2005).

No presente artigo, apresenta-se, neste sentido, um estudo piloto sobre o uso de funções de similaridade como indicador para a verificação da participação dos alunos em fóruns assíncronos

de EaD. Para tanto, foram utilizadas mensagens extraídas de um curso *online* de Computação Gráfica. Cada mensagem postada foi avaliada por um professor especialista, recebendo um valor entre 0 (zero) e 1 (um). O professor descreveu, ainda, um conjunto de palavras-chaves pertinentes ao assunto discutido no fórum, o qual foi utilizado como indicador de participação a ser estudado. Assim, foi efetuado o cálculo da similaridade entre o conjunto de palavras previamente definido pelo docente e cada mensagem do fórum, utilizando-se 4 (quatro) algoritmos: Similaridade dos Cossenos, Distância Euclidiana, Índice de Coeficiente de Jaccard e Dice's Coefficient. O objetivo do estudo foi determinar a existência de diferenças estatísticas entre as funções de similaridade estudadas e, por extensão, a verificação daquela que produzia os melhores resultados. Para examinar a eficiência das funções de similaridade, cada valor obtido a partir de uma função foi confrontado com os valores reais atribuídos pelo especialista.

Dentre as funções analisadas, foi possível verificar que os melhores resultados foram obtidos com a utilização da função de Similaridade dos Cossenos. Além disto, a partir da observação do conteúdo das mensagens e do comportamento dos dados simulados, concluiu-se, também, que utilizar unicamente um indicador não é suficiente para avaliar a participação do aluno, uma vez que este não captura a subjetividade aplicada na avaliação dos discentes.

2. Revisão de Literatura

A utilização de métodos para analisar mensagens em fóruns, com o objetivo de verificar a participação de alunos em cursos *online*, têm sido um dos grandes desafios para professores desta modalidade de ensino. Lin *et al.* (2009) apresentaram uma forma de definir tópicos a partir do conteúdo de fóruns de discussão, fazendo uso de técnicas de mineração de dados. Para isso, eles propuseram a classificação das informações enviadas ao fórum (e.g., anúncio, afirmação, pergunta, explicação, etc.). O foco do estudo era a verificação da eficiência do uso de mineração de dados para tal fim. Assim, os resultados obtidos foram confrontados com a análise subjetiva realizada pelo professor. Para os autores, a presença do professor ao longo da discussão pode contribuir positivamente para a evolução da aprendizagem dos alunos.

No estudo realizado por Júnior e Esmin (2012), o ambiente proposto em (Junior *et al.* 2011), para realização do monitoramento de fóruns, foi reestruturado, com a utilização de um algoritmo de aprendizado semi-supervisionado. No estudo anterior, o ambiente classificava as mensagens em Positivo ou Negativo. As mensagens que recebiam o classificador Negativo continham dúvidas, conteúdo indevido ou insatisfação. Todas as demais mensagens eram classificadas como Positivo. Na pesquisa atual, os autores restringiram a classificação das mensagens em duas categorias Dúvidas e Esclarecimentos a Dúvidas. Para isso, eles utilizaram um vetor de características, em que cada palavra de uma mensagem constitui uma dimensão deste vetor. As mensagens também passaram por um tratamento para remoção de palavras não relevantes para a diferenciação entre documentos (e.g., artigos, preposições, etc.) e algumas delas foram classificadas manualmente. Ao final, apesar dos bons resultados obtidos, eles concluem que a criação de ambientes similares ao desenvolvido, em muitos casos, pode ser complexa.

A avaliação da participação do aluno também foi assunto abordado por Gomes *et al.* (2012). Os autores descreveram a ferramenta GENOME, cujo objetivo era determinar a participação virtual do aluno. Para isso, os autores fizeram uso do conhecimento do docente para definir alguns requisitos, como o período de avaliação da participação; as disciplinas em que a

participação deve ser avaliada; além de possibilitar a configuração de parâmetros para o cálculo da nota e a verificação dos dados da participação ao longo do curso. Eles concluíram que o uso de um mediador para realizar a análise da participação dos alunos diminuiu o esforço realizado pelo docente, facilitando a organização e gerenciamento das notas no ambiente virtual.

Romero *et al.* (2013) sugeriram a utilização de um módulo para o Moodle para prever a aprovação ou reprovação dos alunos em determinada disciplina, a partir de sua participação em fóruns de discussão. Para isso, eles fizeram uso de técnicas de mineração de dados e de indicadores de participação quantitativos, qualitativos e sociais. Alguns dos indicadores contemplavam dados como a quantidade de palavras e o total de sentenças contidas em cada mensagem, além da média de avaliação das mensagens fornecida pelo próprio professor. Para realizar a previsão, eles fizeram uso de dois métodos de mineração: classificação e *clustering*. Os autores verificaram que as mensagens que possuíam conteúdo pertinente ao tema principal do fórum demonstraram maior eficiência no contexto de seu estudo. Como uma de suas conclusões, os autores apontaram como os principais indicadores para verificação do uso do fórum: a verificação de palavras, a avaliação das mensagens, a quantidade de mensagens enviadas e os indicadores sociais. Ao final, eles afirmaram que seria útil o uso de uma abordagem para avaliação automática das mensagens, uma vez que o processo manual é monótono e exige um tempo considerável.

Elia e Chamovitz (2009) apresentaram uma ferramenta de apoio aos estudos sobre a avaliação de mensagens em fóruns de discussão. A ferramenta CQMsg foi produto de pesquisas sobre fóruns categorizados e seu objetivo era a criação de um protocolo de avaliação, adaptável às necessidades de cada avaliador. Esta ferramenta classificava as mensagens em categorias, de acordo com uma taxonomia de classificação definida pelo criador do fórum. Para possibilitar que a ferramenta fosse utilizada de forma mais conveniente, os alunos foram incentivados a construir, ao início de suas mensagens, sentenças que refletissem o seu conteúdo. Assim, os avaliadores poderiam avaliar as mensagens, seguindo um conjunto de elementos definidos previamente pelo professor. No entanto, a análise das avaliações das mensagens foi feita de forma puramente estatística. Os autores afirmaram, ao final, que estão investigando o uso de técnicas de IA, como a lógica nebulosa, para este fim.

De acordo com a revisão bibliográfica efetuada, pôde-se constatar a adoção de indicadores, dentre eles a verificação do conteúdo de mensagens de fóruns, como forma de avaliar a participação do aluno. Observou-se, também, que, apesar de fornecer meios para organizar e facilitar esta avaliação, ainda é de inteira responsabilidade de um supervisor humano, a triagem e análise dos dados fornecidos, utilizando-os como apoio a sua avaliação final. É possível verificar, portanto, a necessidade de se estudar uma forma de avaliar as mensagens automaticamente, visando a facilitar a atividade de análise das mensagens realizada pelo professor.

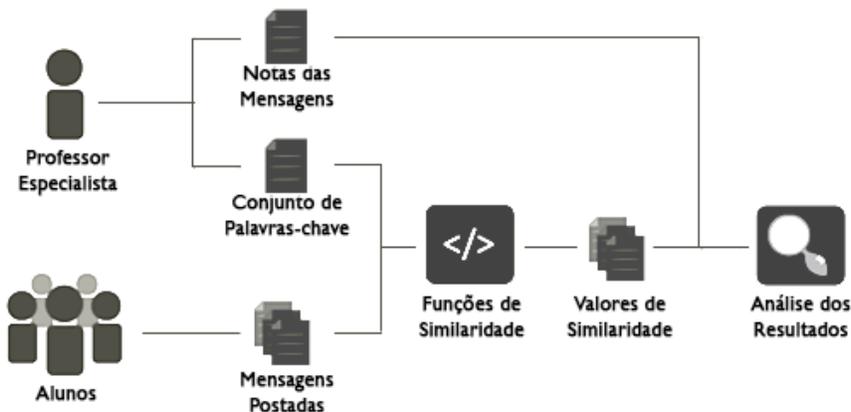
3. Abordagem Proposta

No estudo-piloto realizado, pretendeu-se analisar, dentre as funções de similaridade mais utilizadas para mineração de dados, qual, se aplicável, seria a mais adequada para avaliar mensagens de fóruns em cursos de EaD. Para tanto, foi realizado um experimento, seguido de um processamento estatístico dos dados coletados via experimentação.

A pesquisa teve como objetivo validar o uso de funções de similaridade para a determinação automática do nível de participação de alunos em cursos de EaD, utilizando

mensagens trocadas por alunos em fóruns de discussão. Assim sendo, foram utilizadas as mensagens de um fórum de discussão, em um curso de EaD de curta duração, da disciplina de Computação Gráfica. Um indicador de participação estudado consistiu na similaridade entre as palavras-chaves (termos comuns e relevantes para o tema discutido no fórum) definidas pelo professor especialista e as mensagens enviadas pelo aluno. O estudo realizado pode ser visualizado a partir do diagrama em blocos da Figura 1, descrita mais detalhadamente na seção 3.1.

Figura 1. Diagrama em blocos da abordagem proposta.



3.1. Estudo Empírico

A amostra utilizada para a análise foi obtida de fóruns de discussão de um ambiente virtual de ensino de um curso *online* de Computação Gráfica, consistindo de pares de mensagens e sua respectiva avaliação por um professor especialista.

Para este estudo, o professor especialista definiu também um conjunto de palavras-chaves que julgou relevantes para a discussão e para a atribuição do nível de participação do aluno (exemplos são ilustrados na Tabela 1). Observa-se, a partir da tabela, que várias palavras-chaves apresentam sufixos semelhantes. Portanto, foi realizado um tratamento anterior à inserção destas palavras na base de conhecimento, de forma que apenas os sufixos foram utilizados, o que reduziu consideravelmente o conjunto de palavras definido pelo especialista.

Tabela 1. Exemplo de algumas palavras-chaves do conjunto definido pelo professor especialista, de acordo com o tema definido para o fórum.

Tema do Fórum	O que há de novo na Computação Gráfica?
Palavras-chaves	
<i>captura; capturar; capturado; capturada; capturados; capturadas</i>	
<i>processamento; processar; processado; processada; processados; processadas</i>	
<i>análise; analisar; analisado; analisada; analisados; analisadas</i>	
<i>edição; editor; editado; editada; editados; editadas</i>	

Para a realização do experimento, considerou-se a atribuição de nota:

- Máxima, 1 (um), às mensagens contendo todas as palavras-chaves descritas pelo

professor;

- Mínima, 0 (zero), às mensagens sem nenhuma das palavras-chaves descritas pelo professor;
- Intermediária, entre 0 (zero) e 1 (um), às mensagens contendo algumas das palavras-chaves.

As medidas de similaridade investigadas foram: Similaridade dos Cossenos, Distância Euclidiana, Índice de Coeficiente de Jaccard e Dice's Coefficient (Tan *et al.* 2005).

Para cada função de similaridade, foi gerado um conjunto de valores entre 0 (zero) e 1 (um), representando o valor da similaridade de cada mensagem em relação ao conjunto de palavras-chaves determinado pelo especialista.

Assim, em função do objetivo do estudo foram formulados alguns questionamentos, apresentados a seguir, com o intuito de proporcionar melhor entendimento do que se pretendia investigar.

P1. Há diferença estatística entre os valores simulados obtidos pelas funções de similaridade estudadas, quando confrontados com os valores reais determinados pelo especialista?

P2. Na existência de diferenças entre os valores observados, qual das funções estudadas apresenta resultados mais próximos aos valores reais?

Para responder tais contestações, foi realizada uma análise estatística. Os dados presentes na amostra utilizada para o experimento foram compostos por 37 mensagens extraídas de fóruns de discussão assíncronos, disponibilizados em um curso de Computação Gráfica ministrado no Moodle, de uma universidade particular do estado da Paraíba - Brasil, durante um período de um mês.

Os dados estudados para verificação estatística consistiram em: uma variável dependente do tipo quantitativa contínua, contida no intervalo (0,1), que determinava o valor da similaridade entre a mensagem e o conjunto de palavras-chaves definido pelo especialista (indicador adotado); e uma variável independente do tipo qualitativa nominal, que representava a função de similaridade.

Como o experimento realizado utilizou os resultados gerados a partir da aplicação de funções matemáticas determinísticas, não houve necessidade de repetição de experimentos. O fator único estudado foi o Valor de Similaridade. Este fator possuiu quatro níveis categóricos: Função de Similaridade dos Cossenos, Função de Distância Euclidiana, Função da Similaridade de Jaccard e Função Dice's Coefficient (Tan *et al.* 2005).

Para apoiar o experimento, foram utilizadas as seguintes ferramentas computacionais: ambiente de desenvolvimento integrado Eclipse (Foss 2012), para auxiliar o desenvolvimento do experimento, e o software R Statistics (Ihaka e Gentleman 1993), para realização da análise estatística.

O experimento foi dividido em 3 (três) etapas, a saber: (i) tratamento da amostra; (ii) execução dos algoritmos de similaridade; e (iii) análise dos resultados. Na primeira etapa da experimentação, foi realizada uma análise dos dados da amostra e o tratamento automático do texto, consistindo na retirada de pontuações e acentuações gráficas e na redução de palavras com

mesmo sufixo. A segunda etapa consistiu na execução dos algoritmos de similaridade listados anteriormente para o conjunto de mensagens da amostra, resultando na coleta de quatro conjuntos de valores de similaridade, um para cada algoritmo executado. A terceira e última etapa consistiu na análise estatística dos dados simulados. Após uma análise descritiva inicial, verificou-se que o comportamento de alguns grupos de dados não seguia uma distribuição normal e não possuía homogeneidade de variância. Assim, determinou-se que seria utilizado o teste não paramétrico Mann-Whitney, para estudar estatisticamente os dados, confrontando-se cada conjunto de valores simulados com os valores reais, determinados pelo professor, a fim de verificar se havia diferenças estatísticas entre cada par.

3.2. Resultados

Sobre os dados coletados, foi realizada uma análise estatística exploratória. Os resultados obtidos na análise são descritos a seguir.

Na Tabela 2, é apresentada a sumarização de algumas medidas dos dados simulados, bem como dos dados disponibilizados pelo especialista. Observando apenas os valores da tabela, pode-se verificar que o algoritmo de Similaridade dos Cossenos se destaca dentre os demais, por possuir os valores mais próximos aos reais.

Tabela 2. Sumarização de medidas para cada nível do fator estudado, além dos valores reais determinados pelo especialista.

Simulados	Valores	Minímo	Q1	Mediana	Média	Q3	Máximo
	Similaridade dos Cossenos	0,2425	0,4201	0,4537	0,4534	0,5145	0,6860
Distância Euclidiana	0,8339	0,8467	0,8503	0,8521	0,8578	0,8871	
Similaridade de Jaccard	0,0588	0,1765	0,2059	0,2154	0,2647	0,4706	
Dice's Coefficient	0,1111	0,3000	0,3415	0,3456	0,4186	0,6400	
Reais	Especialista	0,0000	0,3000	0,5000	0,4649	0,5000	1,000

A realização de análise descritiva dos dados possibilitou a verificação da necessidade de utilização de um teste não paramétrico, visto que, após a análise detalhada dos dados, estes se mostraram não normais e heterogêneos. O teste escolhido foi Mann-Whitney, cujos resultados são apresentados na Tabela 3, para um nível de significância de 5% ($\alpha = 0,05$).

Tabela 3. Resultados do teste Mann-Whitney, para cada função de similaridade estudada e valores reais determinados pelo especialista.

Algoritmos	p-valor encontrado
Similaridade dos Cossenos – Notas do Especialista	0,1821
Distância Euclidiana – Notas do Especialista	0,1481E-13
Similaridade de Jaccard – Notas do Especialista	0,1274E-08
Dice's Coefficient – Notas do Especialista	0,0045

Buscando reafirmar as conclusões advindas da aplicação do teste não paramétrico, calculou-se o Erro Quadrático Médio (*Mean Squared Error* – MSE), cujos resultados encontram-

se na Tabela 4. A função de Similaridade dos Cossenos apresenta uma taxa de erro de cerca de 3%, obtendo o melhor resultado em relação às demais funções.

Tabela 4. Erro Quadrático Médio (MSE) para cada função de similaridade estudada.

Algoritmos	Erro Quadrático Médio	Intervalo de Confiança ($\alpha = 0,05$)
Similaridade dos Cossenos	0,0323	(0,0146; 0,0490)
Distância Euclidiana	0,1863	(0,1318; 0,2408)
Similaridade de Jaccard	0,0948	(0,0667; 0,1220)
Dice's Coefficient	0,0478	(0,0307; 0,0648)

3.3. Avaliação da Validação

Em relação à validade do experimento, o fato de que o estudo foi realizado utilizando-se mensagens de fóruns com baixa atividade, devido ao curto período em que estiveram ativos, pode ser citado como uma ameaça à validação externa. Este tipo de validação está relacionado à equivalência entre as conclusões geradas pelo o estudo da amostra em relação a sua população original.

Em geral, os fóruns de discussão em ambientes de EaD podem gerar um volume relativamente grande de postagens, dependendo da quantidade de alunos e do tempo em que permanecem abertos. Dessa forma, os dados podem se comportar de forma diferente em ambientes com maior fluxo de postagens.

Para a validade de construção, que diz respeito às medidas utilizadas, pode-se citar a utilização de um único indicador de avaliação (palavras-chaves), que corresponde a uma medida objetiva, não levando em consideração a subjetividade do docente aplicada na avaliação de discentes e a influência do fator humano. Ambos possuem papel fundamental no ambiente de avaliação real.

A validação interna corresponde a verificar se o efeito ocorrido nas variáveis dependentes foi resultado de manipulação das variáveis independentes. No contexto do experimento realizado, pode-se citar a natureza textual dos dados da amostra, que, sendo extraídos de um ambiente real e mesmo após tratamento, ainda pode conter erros.

4. Discussão dos Resultados

Após a análise estatística, à luz dos questionamentos previamente formulados, foram verificados indícios de que há diferença estatística entre a utilização das funções estudadas. Observando-se os resultados apresentados na Tabela 3, pode se concluir que os algoritmos de Distância Euclidiana, Similaridade de Jaccard e Dice's Coefficient apresentaram diferença estatística quando comparados com os valores reais, i.e., essas funções não apresentam resultados que representem adequadamente a população na qual o experimento está inserido.

É possível verificar que não existem diferenças significativas entre a aplicação da função de Similaridade dos Cossenos e os valores atribuídos no ambiente real por um especialista. Assim,

do ponto de vista estatístico, os melhores resultados foram obtidos a partir da utilização deste algoritmo. Tal afirmação é reforçada por esta função também apresentar, as menores média e mediana para o erro.

Desta forma, foram obtidas respostas para ambas as indagações propostas, enquanto os resultados demonstram que é possível a aplicação de funções de similaridade na verificação da participação de alunos, com resultados positivos. O uso desta técnica pode representar uma forma efetiva de avaliar as mensagens automaticamente, proporcionando a economia do tempo e esforços necessários para a avaliação tradicional destas mensagens realizada pelo professor.

No entanto, é importante ressaltar que, apesar da confirmação de que o uso do indicador estudado pode ser utilizado de forma satisfatória para a avaliação das mensagens dos fóruns, o uso de apenas um indicador para o problema principal da pesquisa (a avaliação da participação do aluno em cursos de EaD) não é representativo. Este fato pode ser constatado quando os valores reais são comparados com os simulados, uma vez que não é possível capturar a subjetividade existente na avaliação realizada pelo professor. O professor, da forma convencional, é capaz de analisar fatores como cópias de textos existentes na *web*, apresentação de vídeos e imagens relevantes ao tema, conversas individuais com o tutor, dentre outros fatores. Neste contexto, outros indicadores também podem ser utilizados, como os indicadores sociais.

5. Considerações Finais

O estudo-piloto confirmou que é possível empregar funções de similaridade no contexto da avaliação da participação de alunos em cursos a distância. Em se tratando do estudo destas funções, o algoritmo da Similaridade dos Cossenos apresentou resultados superiores aos demais e confirmou os benefícios trazidos pelo emprego de técnicas da IA para o tipo de problema abordado. Portanto, pôde-se verificar a viabilidade de implantar técnicas de IA, como Raciocínio baseado em Casos/Regras, usando medidas de similaridade.

Por se tratar de um estudo inicial, apenas um indicador de avaliação da participação foi considerado. E, embora o erro encontrado tenha sido relativamente pequeno, ainda existem limitações para sua aplicação de forma isolada, por ser um indicador objetivo. Durante a avaliação de alunos, não apenas em cursos a distância, mas também presencial, há forte presença da subjetividade. Estudos mais aprofundados sobre como é realizado o processo de avaliação pelos professores desta modalidade de ensino já estão em andamento. Pretende-se, com isso, verificar a viabilidade de empregar novos indicadores, associados ao uso de medidas de similaridade, como apresentado neste trabalho. Como proposta de aprimoramento dos resultados da pesquisa, sugere-se realizar o estudo de forma a generalizar as conclusões, adicionando dados de cursos de diferentes áreas e com diferentes níveis de atividade.

Referências

- Andresen, M. (2009) Asynchronous discussion forums: success factors, outcomes, assessments, and limitations. Em: *Educational Technology & Society*, p. 249–257. Volume 12, Edição 1.
- Blackboard Inc. (1997) Blackboard Learning System. Sistema de gestão da aprendizagem.

- Calvani, A., Fini, A, Molino, M. e Ranieri, M. (2009) Visualizing and monitoring effective interactions in online collaborative groups. Em: *British Journal of Educational Technology*, p. 213–226. Volume 41, Edição 2.
- Dougiamas, M. (2001) MOODLE - Modular Object-Oriented Dynamic Learning Environment. Sistema de gestão da aprendizagem.
- Elia, M. e Chamovitz, I. (2009) CQMMsg - Classificador e Qualificador de Mensagens: um instrumento para apoio à avaliação de fóruns temáticos. Em: *Anais do XX SBIE - Simpósio Brasileiro de Informática na Educação*. Santa Catarina, Brasil.
- Erlin, Yusof, N. e Rahman, A. (2009) Students' interactions in online asynchronous discussion forum: A social network analysis. Em: *International Conference on Education Technology and Computer*, p. 25-29.
- FOSS - Free and Open-source Software (2012). Eclipse, 4.2 Juno. Software development environment.
- Gomes, A. V., Oliveira, D. S. S. e Medeiros, A. C. (2012) GENOME: Uma ferramenta avaliativa de participação virtual. Em: *Anais do Congresso Brasileiro de Informática na Educação*. Rio de Janeiro, Brasil.
- Hrastinski, S. (2008) What is online learner participation? A Literature Review. Em: *Computers & Education*, p. 1755-1765. Volume 51, Edição 4.
- Ihaka, R. e Gentleman, R. (1993). R Statistics, versão 2.15.1. Software environment for statistical computing and graphics.
- Júnior, R. L. O., Esmin, A. A. A., Coelho, T. A., Araujo, D. L., Alosno, L. e Giroto, R. (2011) Uma ferramenta de monitoramento automático de mensagens de fóruns em ambientes virtuais de aprendizagem. Em: *Anais do XXII SBIE - Simpósio Brasileiro de Informática na Educação*. Sergipe, Brasil.
- Júnior, R. L. O. e Esmin, A. A. A. (2012) Monitoramento automático de mensagens de fóruns de discussão usando técnica de classificação de texto semi-supervisionado. Em: *Anais do XXIII SBIE - Simpósio Brasileiro de Informática na Educação*. Rio de Janeiro, Brasil.
- Leite, E. A. M., Sales, G. L., Sousa, L. L. R. e Joye, C. R. (2011) Avaliação assistida, feedbacks e questionários do moodle. Em: *Anais do XXII SBIE - Simpósio Brasileiro de Informática na Educação*. Sergipe, Brasil.
- Lin, F., Hsieh, L. e Chuang, F. (2009) Discovering genres of online discussion threads via text mining. Em: *Computers & Education*, p. 481-495. Volume 52, Edição 2.
- Romero, C., López, M., Luna, J. e Ventura, S. (2013) Predicting students' final performance from participation in on-line discussion forums. Em: *Computers and Education*, p. 458-472. Volume 68.
- Tan, P., Steinbach, M. e Kumar, V. (2005) Introduction to Data Mining. Addison-Wesley Longman Publishing Co., 1a Edição