

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO
DOUTORADO EM INFORMÁTICA NA EDUCAÇÃO**

REINALDO DE JESUS DA SILVA

**SISTEMA DE RECOMENDAÇÃO DE OBJETO DE APRENDIZAGEM BASEADO EM
POSTAGENS EXTRAÍDAS DO AMBIENTE VIRTUAL DE APRENDIZAGEM**

**Porto Alegre
2016**

REINALDO DE JESUS DA SILVA

**SISTEMA DE RECOMENDAÇÃO DE OBJETO DE APRENDIZAGEM BASEADO EM
POSTAGENS EXTRAÍDAS DO AMBIENTE VIRTUAL DE APRENDIZAGEM**

Tese apresentada ao Programa de Pós-Graduação em Informática na Educação do Centro Interdisciplinar de Novas Tecnologias na Educação da Universidade Federal do Rio Grande do Sul, como requisito para obtenção do título de Doutor em Informática na Educação.

Orientadora:

Dra. Rosa Maria Vicari

**Porto Alegre
2016**

CIP - Catalogação na Publicação

Silva, Reinaldo de Jesus da
SISTEMA DE RECOMENDAÇÃO DE OBJETO DE APRENDIZAGEM
BASEADO EM POSTAGENS EXTRAÍDAS DO AMBIENTE VIRTUAL
DE APRENDIZAGEM / Reinaldo de Jesus da Silva. --
2016.
111 f.

Orientadora: Dra. Rosa Maria Vicari.

Tese (Doutorado) -- Universidade Federal do Rio
Grande do Sul, Centro de Estudos Interdisciplinares
em Novas Tecnologias na Educação, Programa de Pós-
Graduação em Informática na Educação, Porto Alegre, BR-
RS, 2016.

1. Ambiente Virtual de Aprendizagem. 2. Objeto de
Aprendizagem. 3. Mineração de Texto. 4. Recomendação.
I. Vicari, Dra. Rosa Maria, orient. II. Título.

REINALDO DE JESUS DA SILVA

**SISTEMA DE RECOMENDAÇÃO DE OBJETO DE APRENDIZAGEM BASEADO
EM POSTAGENS EXTRAÍDAS DO AMBIENTE VIRTUAL DE APRENDIZAGEM**

Tese apresentada ao Programa de Pós-Graduação em Informática na Educação do Centro Interdisciplinar de Novas Tecnologias na Educação da Universidade Federal do Rio Grande do Sul, como requisito para obtenção do título de Doutor em Informática na Educação.

Aprovada em: 02/09/2016.

Profa. Dra. Rosa Maria Vicari - Orientadora

Prof. Dr. Luís Carlos Costa Fonseca - UEMA

Prof. Dr. Eliseo Berni Reategui - PPGIE/UFRGS

Profa. Dra. Daisy Schneider - UFRGS

Aos meus pais e irmãos, minha fortaleza.

À minha esposa Rita de Cássia Montenegro da Silva, razão do meu viver.

A Bianca e Renan, filhos amados, alegrias de minha vida, razões de minha luta, sempre.

“O que sabemos é uma gota, o que ignoramos é um oceano”.

Isaac Newton

AGRADECIMENTOS

A Deus, pela minha vida.

À minha família, apoio constante em tudo que faço.

À minha mãe e meu pai, pelo amor e dedicação recebidos durante toda a minha vida, pois tenho certeza de que o sonho, não só deles, mas de todos os pais, é de ver os seus filhos conquistando mais vitórias.

À minha esposa Rita de Cássia Montenegro da Silva, pela compreensão, carinho, apoio e incentivo.

Aos meus filhos, Bianca e Renan, pela compreensão na minha ausência.

À minha cunhada Ana Cláudia Montenegro Costa, pelo incentivo e pela força durante o desenvolvimento desta pesquisa.

À minha orientadora, Profa. Rosa Maria Vicari, que se fez presente quando o apoio foi necessário nos momentos difíceis dessa jornada, e, sobretudo, pelas orientações e pela capacidade de dividir conhecimentos e experiências.

Ao Prof. Dr. José Valdeni de Lima, pelas orientações e ponderações como Coordenador Geral do DINTER UFRGS-UFMA-UEMA.

Aos professores Henrique Mariano e Ademir Martins, pela exemplar condução operacional do DINTER UFRGS-UFMA-UEMA.

À CAPES, pela bolsa de doutoramento para o desenvolvimento da tese.

Aos colegas do CINTED, pelo suporte técnico e logístico no desenvolvimento da tese.

E a todos aqueles que, de alguma forma, contribuíram para a elaboração desta tese.

RESUMO

Os fóruns de discussões apresentam-se com umas das ferramentas de interação utilizadas nos ambientes virtuais de aprendizagem (AVAs). Esta pesquisa tem como objetivo propor um sistema computacional para recomendação de Objeto de Aprendizagem (OA), levando em consideração as postagens feitas de dentro dos fóruns de um Ambiente Virtual de Aprendizagem (AVA). A metodologia utilizada foi a pesquisa qualitativa, dos tipos descritiva e explicativa. Esse sistema identifica as palavras-chave nos fóruns de um AVA; usam as palavras-chave como indícios dos interesses dos usuários; classifica (atributos pesos) as palavras mais relevantes (*Hot Topics*); submete a um mecanismo de busca (repositório), neste trabalho foram usados os motores de busca, para fins de teste e oferece os resultados da busca aos usuários. As contribuições deste sistema para os sujeitos participantes desta pesquisa são: recomendação automática de OA para os alunos e professores; aplicação de mineração de dados para sistema gestão educacional; técnica de mineração de textos, utilizando algoritmo TF*PDF (*Term Frequency * Proportional Document Frequency*) e integração do AVA com repositório digital. Para validar o sistema de recomendação de OA em um AVA foi desenvolvido protótipo do sistema com uma amostra, contendo vinte e cinco alunos e cinco professores de duas turmas das disciplinas de Modelagem de Banco de Dados e Interface de Usuários e Sistemas Computacionais do curso de Engenharia de Computação da Universidade Estadual do Maranhão. O estudo realizado sobre o tema, e relatado nessa tese, tem como foco a recomendação de OA nos fóruns de um AVA. A avaliação e validação realizadas, através de protótipo do sistema com professores e alunos evidenciaram que o sistema de recomendação de *Web Services* (RECOAWS) proposto atende às expectativas e pode apoiar professores e alunos, nas suas atividades pedagógicas, dentro dos fóruns.

Palavras-chave: Ambiente Virtual de Aprendizagem. Objeto de Aprendizagem. Mineração de Texto. Recomendação.

ABSTRACT

Discussion forums get present with one of interaction tools used in virtual learning environments (VLEs). This research aims to propose a computational system for Learning Object recommendation (LO), taking into account the posts made from within the forums of a Virtual Learning Environment (VLE). The methodology used was a qualitative study of descriptive and explanatory types. This system identifies the keywords in the forums of a VLE; It uses the keywords as evidence of the interests of users; ranks (attributes weights) the most relevant words (Hot Topics); It submits to a search engine (repository), this work were used search engines for testing purposes and provides the search results to users. The contributions of this system to the participants in this study are: automatic recommendation of LO for students and teachers; data mining application to educational management system; text mining techniques, using TF * PDF algorithm (Term Frequency * Proportional Document Frequency) and integration of the VLE with digital repository. To validate the LO recommendation system in a VLE was developed prototype system with a sample, with twenty-five students and five teachers from two classes of database modeling disciplines and User Interface and Computational Systems of Engineering course Computing of the State University of Maranhão. The study on the subject, and reported in this thesis is focused on LO recommendation in the forums of a VLE. The evaluation and validation performed by the prototype system with teachers and students showed that the Web Services recommendation system (RecoaWS) proposed meets expectations and can support teachers and students in their educational activities within the forums.

Keywords: Virtual Learning Environment. Learning object. Text mining. Recommendation.

LISTA DE ILUSTRAÇÕES

Figura 1 - Interação do usuário com o sistema de recuperação através de diferentes tarefas.....	31
Figura 2 - Visões lógicas intermediárias utilizáveis para sistema de RI.....	32
Figura 3 - Ciclo de Vida de um serviço Web SOAP.....	40
Figura 4 - Diagrama de Casos de Uso.....	58
Figura 5 - Diagrama de sequência.....	59
Figura 6 - Arquitetura no processo recomendação de OA no AVA.....	62
Figura 7 - Bloco recoa copiado para o diretório <i>blocks</i> do Moodle.....	68
Figura 8 - Verificação de <i>plugins</i> do Moodle.....	68
Figura 9 - Bloco recoa instalado.....	69
Figura 10 - Interface do RECOAWS.....	70
Figura 11 - Frequência de acesso ao AVA.....	74
Figura 12 - Localização do bloco recoa dentro do AVA/Moodle.....	75
Figura 13 – Palavras-chave extraídas pelo sistema RECOAWS, através do Algoritmo TF*PDF.....	76
Figura 14 - Links sugeridos pelo sistema RECOAWS.....	77
Figura 15 - OAs recomendado pelo sistema RECOAWS que se mostraram adequados para serem utilizados na disciplina.....	77
Figura 16 - Integração do sistema RECOAWS com AVA facilitou a busca de OA para disciplina.....	78
Figura 17 - Satisfação em utilizar o sistema RECOAWS no fórum do AVA.....	79
Figura 18 - Localização do bloco recoa dentro do AVA.....	80
Figura 19 – Palavras-chave extraídas pelo sistema RECOAWS, através do Algoritmo TF*PDF.....	80
Figura 20 - Links sugeridos pelo sistema RECOAWS.....	81
Figura 21 - OAs disponibilizados pelo sistema RECOAWS foram adequados para serem utilizados na disciplina.....	82
Figura 22 - Integração do sistema RECOAWS com AVA facilitou a busca de OA para disciplina.....	83
Figura 23 – Frequência de OA para alunos no AVA.....	84
Figura 24 - Satisfação em utilizar o sistema RECOAWS no fórum do AVA.....	84

LISTA DE QUADROS

Quadro 1 - Especificações deste <i>Web Service</i>	60
Quadro 2 - Lista de Prioridades.....	61
Quadro 3 - Categorias dos resultados da <i>API Bing Search</i>	64
Quadro 4 - Parâmetros da <i>API Bing Search</i>	65
Quadro 5 - Formato retornados pela <i>API Bing Search</i>	65
Quadro 6 - Elementos que compõem os arquivos retornados pela <i>API Bing Search</i>	65
Quadro 7 - Etapas para adicionar o bloco do moodle no curso.....	69
Quadro 8 - Tópicos a serem discutidos nos fóruns do AVA	73

LISTA DE TABELAS

Tabela 1 - Dados sobre utilização do Moodle.....	27
Tabela 2 - Sumário dos trabalhos relacionados.....	49
Tabela 3 - Localização do bloco recoa no AVA, através dos fóruns.....	85
Tabela 4 - Palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns.....	86
Tabela 5 - Os links estão de acordo com as discussões dentro dos fóruns de um AVA.....	87
Tabela 6 – Disponibilização de OA pelo sistema RECOAWS foram adequados para serem utilizados pela disciplina.....	87
Tabela 7 – Integração do sistema RECOAWS com o AVA facilitou a busca de OA.	88
Tabela 8 – Satisfação em utilizar Sistema RECOAWS nos fóruns de um AVA.....	88
Tabela 9 - Resultados da análise descritiva da amostra dos professores.....	109
Tabela 10 - Resultados da análise descritiva da amostra dos alunos.....	109
Tabela 11 - Resultados da localização do bloco recoa no AVA, através dos fóruns.	110
Tabela 12 - Resultados das palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns.....	110
Tabela 13 – Resultados dos <i>links</i> sugeridos estão de acordo com tema da discussão	110
Tabela 14 - Resultados da integração do sistema RECOAWS ao AVA/MOODLE facilita a busca de OA.....	111
Tabela 15 - Resultados do grau de satisfação em utilizar o sistema RECOAWS no AVA/MOODLE.....	111

LISTA DE ABREVIATURAS E SIGLAS

API	<i>Application Programming Interface</i>
AVA	Ambiente Virtual de Aprendizagem
BIOE	Banco Internacional de Objetos Educacionais
DCMI	<i>Dublin Core Metadata Initiative</i>
DM	<i>Data Mining</i>
EAD	Educação a distância
EDM	<i>Education Data Mining</i>
FBC	Filtragem Baseada em Conteúdo
FEB	Federação de Repositórios Educa Brasil
FC	Filtragem Colaborativa
HTTP	<i>Hypertext Transfer Protocol</i>
HTTPS	<i>Hypertext Transfer Protocol Secure</i>
IEEE	<i>Institute of Electrical and Eletronic Engineers</i>
JSON	<i>JavaScript Object Notation</i>
KDD	<i>Knowledge Discovery in Databases</i>
LOM	<i>Learning Objects Metadata</i>
LTSC	<i>Learning Technology Standart Commitee</i>
IMS	<i>Metadata of Instructional Management System</i>
MOODLE	<i>Modular Object-Oriented Dynamic Learning Environment</i>
OA	Objeto Aprendizagem
OBAA	Objetos de Aprendizagem Baseados em Agentes
REST	<i>REpresentational State Transfer</i>
RI	Recuperação de Informação
SOAP	<i>Simple Object Access Protocol</i>
SR	Sistema de Recomendação
TLS	<i>Transport Layer Security</i>
TIC	Tecnologias de Informação e Comunicação
TF-IDF	<i>Term frequency - Inverse Document Frequency</i>
TF-PDF	<i>Term frequency - Ponderal Document Frequency</i>
UEMA	Universidade Estadual do Maranhão
UNISINOS	Universidade do Vale do Rio dos Sinos
UFRGS	Universidade Federal do Rio Grande do Sul
<i>UDDI</i>	<i>Universal Description, Discovery and Integration</i>
XML	<i>Extensible Markup Language</i>
<i>WSDL</i>	<i>Web Services Description Language</i>
WS	<i>Web Service</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Objetivos	19
1.1.1	Objetivo Geral.....	19
1.1.2	Objetivos Específicos.....	20
1.2	Apresentação do Trabalho	20
2	REFERENCIAL TEÓRICO	22
2.1	Objeto de Aprendizagem	22
2.2	Ambiente Virtual de Aprendizagem	24
2.3	Repositório Digital	28
2.4	Sistema de Recuperação de Informação	30
2.4.1	A visão lógica de um documento.....	31
2.4.2	Modelo Booleano.....	33
2.4.3	Modelo Vetorial.....	33
2.4.4	Modelo Probabilístico.....	34
2.5	Sistema de Recomendação	35
2.7	Arquitetura Orientada a Serviços	38
2.7.1	<i>Simple Object Access Protocol - SOAP</i>	39
2.7.2	<i>Representational State Transfer - REST</i>	40
2.8	Mineração de Texto	41
2.9	Algoritmo <i>Term Frequency – Inverse Document Frequency</i> (TF-IDF)	43
2.10	Algoritmo <i>Term Frequency * Proportional Document Frequency</i> (TF*PDF)	44
2.11	Considerações do Capítulo	45
3	TRABALHOS RELACIONADOS	47
3.1	Contribuições	50
3.2	Considerações do Capítulo	51
4	METODOLOGIA DA PESQUISA	52
4.1	Sujeitos da Pesquisa	53
4.2	Instrumento de Coleta de Dados	53
4.3	Procedimentos para Coleta de Dados	54
4.3	Procedimentos para Análise de Dados	55
4.4	Recursos Tecnológicos	56
4.5	Considerações do Capítulo	56
5	PROTOTIPAGEM DO SISTEMA DE RECOMENDAÇÃO DE OA DENTRO DO FÓRUM DE UM AVA	57

5.1	Modelagem do Sistema RECOAWS	58
5.2	Arquitetura do Sistema RECOAWS	61
5.3	Serviço de Recomendação	62
5.4	Integração do Sistema RECOAWS com AVA	66
5.5	Usuários	70
5.6	Conclusão do Capítulo	71
6	RESULTADOS E DISCUSSÕES	72
6.1	Resultados	74
6.1.1	Alunos.....	74
6.1.2	Professores.....	79
6.2	Análise de discussões dos resultados	85
7	CONCLUSÕES E TRABALHOS FUTUROS	90
	REFERÊNCIAS	93
	APÊNDICE A – QUESTIONÁRIO PROFESSORES	98
	APÊNDICE B – QUESTIONÁRIO ALUNOS	99
	APÊNDICE C – DECLARAÇÃO DE LIBERAÇÃO DO CCT/UEMA	100
	APÊNDICE D – DECLARAÇÃO DE LIBERAÇÃO DO CURSO DE ENGENHARIA DE COMPUTAÇÃO	101
	APÊNDICE F - CÓDIGO FONTE PRINCIPAL MÉTODO DO BLOCO RECOAWS	102
	APÊNDICE G - SERVIÇO WEB RECOAWS	103
	ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)	108
	ANEXO B - RESULTADOS DAS AMOSTRAS CALCULADAS ATRAVÉS DO SOFTWARE BIOESTAT	109

1 INTRODUÇÃO

Esta tese tem como objetivo o desenvolvimento de um sistema de recomendação de objetos de aprendizagem (OA), baseado em postagens extraídas de dentro dos fóruns de um ambiente virtual de aprendizagem (AVA), que é um espaço *online* destinado a organizar e coordenar as atividades de ensino e de aprendizagem, ou seja, são as salas de aula *online*, que agregam ferramentas para a criação, a tutoria e a gestão de atividades (SILVA, 2011).

A motivação para a realização desta pesquisa surgiu do desenvolvimento dos projetos Federação Educa Brasil (FEB) e Objetos de Aprendizagem, baseados em Agentes (OBAA). Esses projetos vêm sendo utilizados por várias Instituições de Ensino e Pesquisa no Brasil. Entretanto, a recuperação efetiva dos objetos está condicionada à disposição do usuário em ir à interface de consulta da FEB, digitar algumas palavras-chave e navegar entre muitos documentos recuperados. Além disso, a ferramenta de busca da FEB, assim como vários outros disponíveis na internet, só considera uma consulta pontual (*ad hoc*), não levando em consideração o histórico de interação do utilizador, mas é importante ressaltar que isso não é um problema específico dessas ferramentas, pois não foram projetadas para esta finalidade (BRASIL, 2010).

Para a elaboração desta pesquisa, escolheu-se como metodologia a pesquisa qualitativa. O trabalho foi dividido em duas etapas. Na primeira, procurou-se responder aos objetivos específicos, mediante pesquisa descritiva e explicativa, resultando em um protótipo do sistema. Na segunda etapa, analisou-se os resultados obtidos com avaliação do protótipo do sistema. O sistema de recomendação foi denominado inicialmente de RECOAWS (Recomendação, Objeto de Aprendizagem e *Web Services*), o sujeito da pesquisa é composto por vinte e cinco alunos e cinco professores de duas turmas das disciplinas de Modelagem de Banco de Dados e Interface de Usuários e Sistemas Computacionais do curso de Engenharia de Computação da Universidade Estadual do Maranhão.

Como embasamento teórico, há dois pontos principais: sistema de recomendação de OAs através do algoritmo TF*PDF (*Term Frequency * Proportional*

Document Frequency) e arquitetura orientada a serviços para busca dos OAs nos repositórios.

No que tange aos sistemas de recomendação, segundo Ricci *et al.* (2011), esses sistemas tentam prever quais são os produtos ou serviços mais adequados, com base nas preferências e restrições do usuário. Para completar essa tarefa, eles coletam dos usuários suas preferências, que podem ser explícitas, através das avaliações de produtos; ou implícitas, inferidas através da interpretação das ações do usuário.

Quanto ao algoritmo TF*PDF, proposto em 2001 por Bun e Ishizuka, o TF*PDF é uma abordagem que busca atribuir pesos mais significativos aos termos mais frequentes na coleção de documentos e tem se apresentado como uma excelente ferramenta no que se refere à mineração de texto para a detecção de tópicos, tanto em um documento quanto em um conjunto de documentos.

A mineração de textos, também chamado de mineração de dados textuais, permite transformar grande parte desses conteúdos não estruturados em conhecimento útil para as organizações. A definição geral de mineração de texto inclui todos os tipos de processamento de texto que tratam de encontrar, organizar e analisar informações (KONCHADY, 2006).

Para arquitetura orientada a serviço, Zimmermann (2012) destaca que o serviço *web* ou *web service* são um meio simples, de baixo custo e multiplataforma para compartilhamento de funções e dados entre aplicações corporativas; sendo assim, utiliza-se *Web services* para busca de OAs no repositório digital. De acordo com a norma IEEE 1484 (2002, p.3), um objeto de aprendizagem é definido como “qualquer entidade, digital ou não, que possa ser usada para fins de ensino, aprendizagem ou treinamento”. As principais vantagens são as possibilidades de reutilização, a possibilidade de adaptação de um objeto, a ilustração de conceitos abstratos e o aspecto interativo que possibilita melhoras nos processos de ensino e de aprendizagem.

De maneira geral, na Educação, o uso de OAs torna-se bastante vantajoso, pois propicia o aumento da circulação de informações em ambientes educacionais, tudo com auxílio do desenvolvimento de tecnologias digitais, como: computadores, dispositivos móveis, TV a cabo e internet que contemplam as mais diversas necessidades de seus usuários.

Nesse sentido, para que esses OAs não tenham utilidade apenas uma única vez, é necessária que haja uma padronização na sua catalogação, e mais, é importante também que se possa recuperar esse conteúdo. Consequentemente, os OAs devem ser “armazenados em repositórios e reutilizados em diversos ambientes de aprendizagem. Para isso, torna-se necessária a descrição de seu conteúdo de uma forma padronizada, permitindo o intercâmbio de informações” (VICARI *et al*, 2009, p. 7).

A disseminação de repositórios educacionais de OAs, nesse contexto, propicia o surgimento de um volume expressivo de recursos pedagógicos, que contribuirá para avanços significativos nas práticas educacionais, possibilitando ao professor organizar, avaliar quais recursos pedagógicos, ou seja, que OAs, poderão ser utilizados e qual a melhor estratégia. Ao dispor de todo esse material, em um repositório digital, os envolvidos nesse processo de ensino-aprendizagem, cooperam entre si. Dessa forma, os OAs podem ser armazenados e a reutilização é, para tanto, fundamental para a escolha e a criação de um padrão de metadados.

Dentre as diversas organizações e grupos de pesquisa, que criaram padrões para metadados, destacam-se: OBAA (Objetos de Aprendizagem Baseados em Agentes) (VICARI *et al*, 2010); LOM (*Learning Objects Metadata*) do *Learning Technology Standart Committee of Institute of Electrical and Eletronic Engineers* (IEEE/LTSC) (IEEE, 2002); o *Metadata of Instructional Management System* (IMS) *Global Consortium* (IMS, 2002); e o *Dublin Core Metadata Initiative* (DCMI, 2014) os quais se baseiam na taxionomia para garantir a consistência entre os metadados dos diferentes autores desenvolvedores dos OAs.

Ao longo dos anos, pesquisas estão sendo desenvolvidas na área de recomendação de informação, no sentido de proporcionar aos colaboradores e usuário de sistema educacional a distância. Entre elas, podem-se destacar: Manouselis; Vuorikari e Van Assche (2007); Betancur; Moreno e Ovalle (2009); Yan e WU (2009); Ferreira e Raabe, (2010); Vieira e Nunes (2011); Casali (2011); Rodrigues (2012), entre outras, utilizando métodos e técnicas que auxiliem nos processos de ensino e de aprendizagem dentro do AVA.

A utilização de um sistema de recomendação, integrado ao AVA para recomendar OAs dentro de um fórum, ainda é um obstáculo a ser ultrapassado, visto que se tem uma infinidade de OAs disponibilizados por instituições públicas e

privadas, prontos para serem utilizados pelos seus professores e alunos, mas que não sabem ou não têm ferramentas adequadas para localizá-los e recuperá-los para o reuso.

Com base nos pressupostos evidenciados, considera-se relevante desenvolver um sistema de recomendação de OAs, a partir de um repositório digital, levando-se em consideração os registros deixados pelos alunos e professores, dentro dos fóruns de um AVA, no contexto educacional. Uma das vantagens dos sistemas de recomendação é justamente prevenir o usuário de informações indesejáveis, já que, atualmente, a *web* se configura uma verdadeira “tempestade de informações”. Nesse aspecto, tem-se a pretensão de oferecer aos usuários apenas o conteúdo que seja realmente relevante e correlacionado ao contexto em que este está inserido.

Um AVA provê, em geral, ferramentas tais como: bate-papo (*chats*), vídeo aulas, lista de discussão, mural, enquete, portfólio e fóruns. Em relação aos fóruns, eles consistem em espaços para discussões e trocas de ideias sobre assuntos definidos por seus participantes e podem permitir uma experiência de aprendizagem favorável ao processo pedagógico (ABAWAJY, 2012). Sendo assim, um dos desafios dos professores e alunos dentro AVA é a recomendação de OAs. Nessa perspectiva, a problemática que norteou esta tese foi entender: “como um mecanismo de recomendação de OAs, armazenados em repositórios, pode oferecer conteúdo útil a usuários de um AVA, a partir de postagens registradas nos fóruns?”. Partindo desse pressuposto, abordar-se-á na próxima seção o objetivo geral.

1.1 Objetivos

Os objetivos deste trabalho foram divididos em geral e específicos.

1.1.1 Objetivo Geral

O objetivo geral desta pesquisa é propor um sistema computacional para recomendação de OA, baseado em postagens extraídas de dentro dos fóruns de um AVA.

1.1.2 Objetivos Específicos

Para que o objetivo geral possa ser atingido, são propostos os seguintes objetivos específicos:

Realizar o levantamento bibliográfico dos principais modelos e técnicas de sistema de recomendação de informação;

Identificar e analisar os fundamentos da área de recuperação da informação, mais especificamente mineração textual e sistemas de recomendação;

Apresentar arquitetura de sistema de recomendação baseado em postagens extraídas do fórum de um AVA, por mineração textual;

Demonstrar um sistema de recomendação de OA baseado em uma arquitetura *Web Services*;

Apresentar a integração do sistema recomendação ao AVA;

Disponibilizar ao usuário de um AVA, através dos fóruns, acesso ao sistema de recomendação;

Verificar se as palavras-chave, extraídas pelo sistema de recomendação, estão de acordo com as discussões realizadas no fórum de um AVA;

Investigar se o sistema de recomendação disponibilizou OA de acordo com discussões no fórum de um AVA;

Avaliar se os OAs recomendados pelo sistema estão de acordo com as postagens realizadas nos fóruns;

Analisar a satisfação dos usuários em relação ao Sistema de Recomendação.

1.2 Apresentação do Trabalho

Este trabalho está dividido em sete capítulos, dispostos da seguinte forma: o primeiro capítulo apresenta os conceitos de AVA, algoritmo TF*PDF, OA e *Web Service*; o objetivo geral, que permitiu delimitar o estudo e facilitando a sua orientação; os objetivos específicos, permitindo e determinando as etapas que devem ser cumpridas para se alcançar o objetivo geral; e elementos teóricos que clarificam as ideias da pesquisa. O segundo capítulo expõe os tópicos relacionados

às tecnologias de Sistema de Recomendação, destacando as técnicas de filtragens; no terceiro capítulo são apresentados os trabalhos relacionados da tese; o quarto capítulo apresenta a metodologia adotada na pesquisa; já o quinto capítulo apresenta a prototipagem do sistema implementado no decorrer desta pesquisa; o sexto capítulo exhibe a pesquisa realizada, assim como os seus resultados e discussões. E, por fim, no sétimo, as conclusões e trabalhos futuros da pesquisa.

2 REFERENCIAL TEÓRICO

Para a compreensão do que vem a ser OA, AVA, repositórios digitais, sistema de recomendação, sistema de recuperação de informação, arquitetura orientada a serviços, dentre outros recursos para mineração de texto, faz-se necessário um “estado-da-arte” que desperte, no leitor, o interesse em descobrir as possibilidades desses recursos digitais.

2.1 Objeto de Aprendizagem

A internet oferece a seus usuários a possibilidade de criar ambientes de aprendizagem inovadores e desafiantes ao propiciar o acesso a fontes de informações dificilmente acessíveis por outros meios. A dificuldade de reutilização dos conteúdos, para fins educacionais, decorre da ineficiência na busca por conteúdos específicos, para a solução do problema, relacionados ao desenvolvimento de OAs.

Nesse contexto, a operacionalização do processo de desenvolvimento de OA apresenta algumas inconsistências. De acordo com a análise realizada, tais eventos são ocasionados por diversos fatores, dentre os quais podem-se elencar: a falta de padronização de objeto de aprendizagem entre as instituições de ensino superior; a ausência de integração do AVA com repositórios de conteúdo digital; a dificuldade para a recuperação e a recomendação de OAs dentro de um AVA; e a ausência de padronização no processamento de comunicação entre as bases de dados.

Esse fato levou a comunidade científica a desenvolver novos recursos para auxiliar o ensino e a aprendizagem, como, por exemplo, o desenvolvimento de OAs que visa facilitar a reutilização, a distribuição e a personalização de conteúdos educativos na internet, promovendo a portabilidade e a interoperabilidade dos recursos de aprendizagem.

Com o advento da web, surgem mais possibilidades para colaboração e compartilhamento de recursos didáticos e digitais, com finalidade educacional, que podem ser utilizados em qualquer área do conhecimento. Nessa perspectiva de

utilização de recursos digitais no processo de ensino e aprendizagem, Wiley (2000) considera que o conceito de OA pode ser compreendido da seguinte forma:

[...] objetos de aprendizagem são geralmente entendidos como sendo entidades digitais distribuídos parcialmente através da Internet, isto significa que, qualquer número de pessoas pode acessar e utilizar simultaneamente (em oposição aos meios de instrução tradicionais, tais como uma sobrecarga ao utilizar uma fita de vídeo, que só pode existir em um lugar ao mesmo tempo). Além disso, aqueles que incorporam objetos de aprendizagem podem colaborar e beneficiar imediatamente de novas versões. Estas são as diferenças significativas entre objetos de aprendizagem e outras mídias de instrução que tem existido anteriormente (WILEY, 2000, p. 3, tradução nossa).

Do conceito apresentado por Wiley (2000), é fácil perceber características relacionadas aos OAs, como: estar disponível em formato digital; poder ser acessado por várias pessoas ao mesmo tempo (através da Internet) e a possibilidade dos usuários colaborarem com novas versões. Ainda é possível perceber que o autor, ao propor a definição, restringe os objetos de aprendizagem apenas ao formato digital.

Conforme Haughey e Muirhead (2005), os OAs são concebidos para ajudar os professores a desempenhar as seguintes funções: introduzir novos temas e habilidades; proporcionar um reforço de competências existentes; estender a aprendizagem através de novos meios para a apresentação de material curricular; ilustrar os conceitos que não são facilmente explicados através de métodos tradicionais de ensino; dar suporte a novos tipos de oportunidades de aprendizagem que não estão disponíveis em um ambiente de sala de aula; e proporcionar atividades de enriquecimento para alunos motivados.

Os OAs são armazenados em repositórios digitais, no ambiente virtual, e são descritos por meio de metadados. Os metadados representam e tornam acessíveis os objetos de aprendizagem, os quais constituem, assim, recursos pedagógicos que podem ser reutilizados em diversos ambientes e por qualquer pessoa.

Para Fabre; Tamusiunas e Tarouco (2003), os principais benefícios oriundos da utilização dos metadados para catalogar OAs são: acessibilidade – a sua busca é facilitada, permitindo, assim, o acesso ao OA; Interoperabilidade – correspondem à capacidade de um OA poder operar em diferentes plataformas e/ou sistemas;

durabilidade – condições de utilização do OA mesmo com as evoluções de *software* e *hardware*.

O aprendizado pode se tornar muito mais rico, através da reunião de um conjunto de possibilidades pedagógicas, como artigos, *softwares* educativos, revistas e jornais *on-line*, jogos, vídeos, projetos, entre outros. Tendo oportunidade de acessar todos esses recursos na internet, os estudantes podem tornar-se participantes ativos na busca do conhecimento, no momento em que forem desafiados a criarem um espaço com informações relevantes de acordo com a temática em evidência.

Finalmente, ensinar, utilizando recursos digitais, exige uma forte dose de atenção do professor. Diante de tantas possibilidades de busca, de compartilhamento e cooperação, o professor se depara com outra realidade, uma nova abordagem do ensinar e de aprender no contexto virtual. A próxima seção, apresenta o ambiente virtual de aprendizagem.

2.2 Ambiente Virtual de Aprendizagem

Um AVA é um espaço *online* que permite a interação assíncrona e síncrona entre alunos, professores e tutores, que se realiza através de ferramentas que variam de acordo com cada ambiente. Os ambientes, em sua maioria, apresentam modelo básico, no qual as estrutura das páginas já estão definidas, e um conjunto adicional de recursos que pode ser acrescentado á estrutura do curso.

A interação dos participantes dos processos de ensino e aprendizagem (professores e alunos), através de um AVA, ocorreu devido o avanço do ensino a distância na década de 90, principalmente com uso das Tecnologias da Informação e Comunicação (TICs), em razão do advento da internet e o surgimento dos novos recursos tecnológicos.

Os recursos desses ambientes digitais de aprendizagem são basicamente os mesmos existentes na Internet (correio, fórum, bate-papo, conferência, banco de recursos, e outros), com a vantagem de propiciarem a gestão da informação segundo critérios preestabelecidos de organização, definidos de acordo com as características de cada software. Estes possuem bancos de informações

representados em diferentes mídias (textos, imagens, vídeos, hipertextos) e interligados com conexões constituídas por links internos ou externos ao sistema.

Para Almeida (2003), os AVAs buscam oferecer suportes de atividades mediados pelas TICs. Eles oferecem suporte e gestão à sala de aula virtual, integrando diversas tecnologias. Uma característica comum em ambientes computacionais, utilizados no contexto educacional, é a capacidade de coletar e armazenar uma grande quantidade de dados sobre os estudantes.

Esses dados são bastante amplos, variando desde registros de acesso, interações diversas com o sistema, até dados com ricos significados semânticos, tais como as mensagens em *chats* e fóruns. No AVA, os fóruns são uma poderosa ferramenta de comunicação e consistem em espaços de discussões e trocas de ideias entre seus participantes. Através deles, cada participante pode fazer comentários sobre os temas discutidos, possibilitando, assim, o entendimento mútuo.

Os fóruns constituem uma ferramenta assíncrona, isto é, independente de tempo, que permitem a conversa de todos com todos, cada qual no seu tempo, possibilitando a criação de um ambiente centrado na interação *online*. Kenski (2002) destaca que a interação, com o conhecimento e com as pessoas, é algo fundamental para o aprendizado. Para a transformação de um determinado grupo de informações em conhecimentos, é preciso que as discussões sejam trabalhadas, discutidas, comunicadas.

Segundo Abawajy (2012) as principais características que podem influenciar e diferenciar os vários tipos de fóruns de discussão são:

1. grau de interação - o sucesso ou fracasso de uma discussão *online* assíncrona está relacionado com a qualidade das interações, predominantemente aluno-aluno, e a profundidade da aprendizagem, que ocorre nessas discussões;

2. requisitos de participação – a participação em fórum de discussão fornece ao aluno oportunidades para aprendizagem ativa, através da leitura de outros comentários, da postagem de suas próprias questões e do fornecimento de resposta a outras postagens;

3. volume e frequência de postagens - discussões assíncronas fornecem ao aluno tempo de reflexão e permitem que eles compartilhem suas próprias perspectivas ao analisarem o ponto de vista dos outros. Além disso, a quantidade de mensagens, geradas em discussões, é um fator muito importante

para tornar um fórum atraente e gerenciável. Assim, é preciso observar que, quando o número de mensagens aumenta, o fórum pode causar problema ao aluno para identificar conteúdo relevante;

4. atividade de discussão - está relacionada com a concepção do próprio fórum, no sentido de definir se a participação do aluno é obrigatória ou não, se há atribuição de nota, bem como os temas de discussão;

5. resposta - a característica comum em fóruns de discussão é que os alunos postam mensagens em um local permanente, onde podem ser lidas e comentadas; além do mais, as postagens dos professores/tutores, na forma de resposta, são importantes para motivar os alunos a contribuírem na discussão. No entanto, respostas pouco frequentes, inexistentes, atrasadas, irrelevantes ou negativas, podem render discussões inúteis, impedindo ou dificultando a aprendizagem eficaz. Desse modo, os alunos podem deixar de contribuir se eles não recebem nenhuma resposta imediata ou comentários de outros participantes. Para Scherer (2009), os fóruns são espaços abertos para alunos e professores inserirem estratégias de comunicação e diálogo, fazendo com que eles se movimentem na busca por entendimento e produção do saber.

Atualmente, existem diversos AVAs, sejam comerciais e/ou gratuitos, disponíveis no mercado e dentre eles, podem-se destacar: AulaNet, Claroline, eFront, Atutor, Docebo, Dokeos, Ilias, Openelms, Moodle, Sakai, Teleeduc, Rooda e Planeta Rooda (SILVA, 2011). Observa-se que existem vários AVAs que estão sendo utilizadas para o ensino tanto na EAD quanto no ensino presencial. Dessa forma, com base nos pressupostos evidenciados, considera-se relevante investigar um sistema para recomendação de OA, baseado em postagens extraídas dos fóruns de um AVA no contexto educacional.

Segundo os autores Penteadó, Gluz e Galafassi (2011), 98% das instituições de ensino públicas utilizam AVA/*Moodle*. Os autores destacam ainda que há um quadro semelhante também encontrado nas instituições privadas. Como pode-se observar, os autores destacam o uso do *Moodle* nas instituições de ensino.

O *Moodle* é um sistema gerenciador de cursos de aprendizagem de código aberto, utilizado como auxílio aos educadores, para criar efetivas comunidades de aprendizagem *online* (MOODLE ORG, 2015; COLE, 2007), ou seja, é um sistema de distribuição livre, sob licença *open-source*, em todo o mundo e em diferentes

instituições. A arquitetura do *Moodle*, sua implementação, interoperabilidade e intencionalidade são consideradas boas e tem uma comunidade muito forte – está presente em mais de 230 países, em mais de 76.645 sites registrados e 87.681.684 usuários cadastrados.

A utilização do *Moodle*, como AVA, vem crescendo rapidamente. Segundo dados extraídos do site oficial do ambiente, até maio de 2016, 230 Países e 76.645 sites usam a plataforma. A Tabela 1 mostra dados sobre a utilização do *Moodle* até maio de 2016.

Tabela 1 - Dados sobre utilização do Moodle

INDICADOR	QUANTITATIVO
Sites registrados	76.645
Números de Países	230
Números de Cursos	9.786.258
Números de Usuários	87.681.684
Inscrições	260.449.952
Postagens no fórum	178.560.245
Recursos utilizados	88.797.386
Questões do quiz	458.456.038

Fonte: <http://moodle.org/stats/>, acesso em: 29 de maio 2016

De acordo com Silva (2011, p.123), o “*Moodle* requer criatividade, planejamento e atenção aos novos recursos e possibilidades que são oferecidos pelo crescente processo de digitação da informação”. Existem novos recursos que podem ser incorporados ao *Moodle* são eles: *SlideShare*, *Authorstream*, *Youtube*, *Picasa*, *Hot Potatoes*, *Webquest*, entre outros. Esses recursos possibilitam ao usuário do AVA o desenvolvimento do OA, que pode acontecer de diversas formas, como por exemplo, *chats*, fóruns de discussão, *wikis*, dentre outras. Em relação aos fóruns, eles consistem em espaços para discussões e trocas de ideias sobre assuntos definidos por seus participantes e podem permitir uma experiência de aprendizagem favorável ao processo pedagógico (ABAWAJY, 2012).

A produção tradicional de cursos de aprendizagem *online* é uma atividade trabalhosa em que o desenvolvedor (geralmente o professor) tem de escolher os conteúdos, que serão apresentados, decidir a estrutura dos conteúdos e determinar os elementos dos conteúdos mais apropriados para cada tipo de potenciais usuários do curso. Dada a complexidade dessas decisões, um projeto curto é dificilmente possível, mesmo quando é feito cuidadosamente. Além disso, na maioria dos casos,

é necessário evoluir e possibilitar modificações no conteúdo dos cursos, estrutura e navegação, baseados nas informações de uso dos alunos, tudo isso para proporcionar a eles condições de permanência e avanço em um curso *online*.

A base de dados do *Moodle* registra todas as interações referentes ao uso do ambiente, à medida que o usuário o utiliza (ROMERO; VENTURA e GARCÍA, 2008). As interações referentes à visualização dos módulos e realização dos exercícios, por exemplo, são armazenadas em tabelas correspondentes (grupos de tabelas que guardam informações sobre os módulos ativos e grupos de tabelas que guardam informações referentes aos questionários), dado a estrutura modular do curso.

O *Moodle* possui um banco de dados extenso, e com muitas tabelas, demandando muito estudo para melhor compreendê-lo. Ele registra as informações em um banco de dados que contém mais de 200 tabelas. Como exemplo, segue alguns nomes de algumas tabelas agrupadas: *mdl_assignment*, *mdl_course*, *mdl_course_categories* e *mdl_course_completions*. Portanto, entre os AVAs, disponíveis no mercado, nesta pesquisa, adota-se o AVA/Moodle para o desenvolvimento do protótipo do sistema Recoaws (SILVA *et al*, 2015). Na próxima seção, apresenta-se o repositório digital.

2.3 Repositório Digital

Com o crescimento da oferta de OAs, na internet, surgiu a necessidade de criação de ferramentas que pudessem gerenciar o grande volume de arquivos, catalogá-los, utilizando padrões internacionais de metadados para garantir a interoperabilidades do conteúdo e acesso do público aos OAs, a partir da divulgação das comunidades que criam e compartilham esses recursos. Estas ferramentas são comumente chamadas de repositórios e, na literatura, são definidas como sistemas disponíveis na internet que fornecem, principalmente, facilidades de depósito e acesso aos objetos digitais. Nesse sentido, os repositórios agregam uma variedade de facilidades, em grande parte relacionada ao gerenciamento dos objetos digitais neles depositados.

Sendo assim, os repositórios, além de gerenciar os documentos digitais, têm como finalidade a preservação destes e são sistemas flexíveis que, apesar de terem

“origem acadêmica, estão sendo utilizados para outros fins, tais como bibliotecas jurídicas, bancos de objetos educacionais ou acervos” (SHINTAKU;MEIRELLES, 2010, p.17-18).

Logo, quando o depósito do OA é realizado em um repositório de uma instituição dificilmente esse conteúdo será perdido, pois existe toda uma infraestrutura e o interesse dessas instituições em manter os repositórios, para registro da memória intelectual do que é produzido pela comunidade científica.

Um repositório institucional de acesso aberto constitui um serviço de informação científica - em ambiente digital e interoperável - dedicado ao gerenciamento da produção científica e/ou acadêmica de uma instituição (universidades ou institutos de pesquisa). Contempla a reunião, armazenamento, organização, preservação, recuperação e, sobretudo, a ampla disseminação da informação científica produzida na instituição (LEITE *et al.*, 2012, p. 7).

Com o repositório digital, sendo criado na perspectiva de acesso aberto e isento de qualquer taxa, acontece a democratização da produção científica, pois o conhecimento produzido nas universidades pode ser utilizado pela sociedade em geral, como professores e alunos do ensino fundamental, médio e universitário. Outra vantagem do acesso livre corresponde a uma maior visibilidade da universidade e de seus pesquisadores, diante da comunidade científica, da forma como aconteceu no repositório Lume da UFRGS, que esteve durante o ano de 2015, como um dos mais acessados da América Latina. Desse modo, os pesquisadores da UFRGS estão tendo os seus trabalhos citados por pesquisadores de outros países.

Porém, quando uma universidade ou instituições de ensino se envolvem em projetos, voltados para divulgação do seu patrimônio cultural e científico, elas devem utilizar repositório digital, em vez de outras coleções de arquivos (como simples páginas *web*) existentes na Internet. Para Heery e Anderson (2005), existem algumas características importantes que definem se uma ferramenta pode ser considerada um repositório digital: o depósito do conteúdo deve ser realizado pelo criador (autor de uma dissertação ou professor orientador) ou proprietário; sua arquitetura deve ter recursos para administrar o armazenamento do conteúdo e os metadados; possuir rotinas para depósito, edição, busca por conteúdos e contador de visitas por OA; e ser confiável e gerenciar bem todo o conteúdo armazenado.

Mediante os apontamentos listados acima, pode-se postular que um repositório digital é um meio pelo qual as universidades podem divulgar a sua produção científica, atraindo mais visibilidade para a instituição, assim como para os seus pesquisadores.

Nesse sentido, o repositório digital deve possuir a capacidade de gerenciar o depósito, catalogação de metadados, controlador de visitas e propiciar o autoarquivamento por parte dos autores.

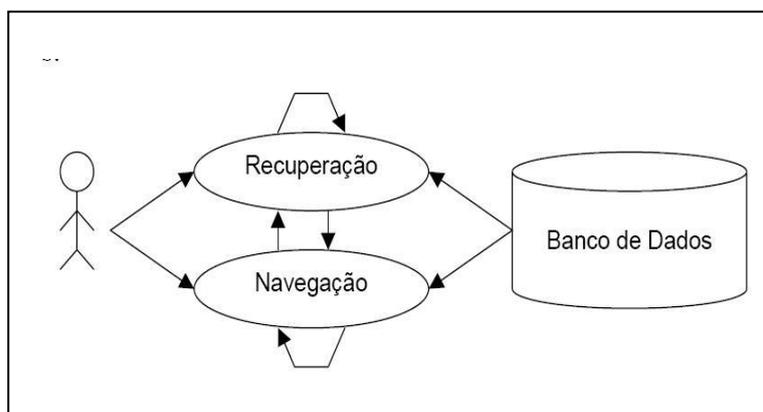
Portanto, é importante mencionar que a principal característica de um repositório digital está relacionada com o seu objetivo de oferecer OAs para serem utilizados por docentes e discentes nos processos de ensino e de aprendizagem.

2.4 Sistema de Recuperação de Informação

A Recuperação de Informação (RI) visa fornecer aos seus usuários acesso às informações de seu interesse com qualidade. Essas informações podem ser recuperadas em uma página *Web*, catálogo *online*, repositório digital, objetos multimídia, entre outros (BAEZA-YATES;RIBEIRO-NETO,2011).

A área de RI remete à representação, armazenamento, organização e acesso a itens de informação. A representação e a organização dos itens deve prover, para o usuário, acesso fácil àquelas informações pelas quais ele se interessa. Segundo Baeza-Yates;Ribeiro-Neto, (1999, p. 3), a “recuperação efetiva de informações relevantes é diretamente afetada pelas tarefas com as quais o usuário está envolvido (*user tasks*), bem como pela visão lógica dos documentos adotada pelo sistema de RI”. A Figura 1 ilustra a interação do usuário nas diferentes tarefas identificadas.

Figura 1 - Interação do usuário com o sistema de recuperação através de diferentes tarefas



Fonte: Baeza-Yates; Ribeiro Neto 1999, p. 4.

O usuário de um sistema de RI precisa traduzir suas necessidades de informação em uma consulta na linguagem fornecida pelo sistema. Um sistema de RI recupera essa informação, através de um motor de busca, pesquisando ou consultando informação do seu interesse.

2.4.1 A visão lógica de um documento

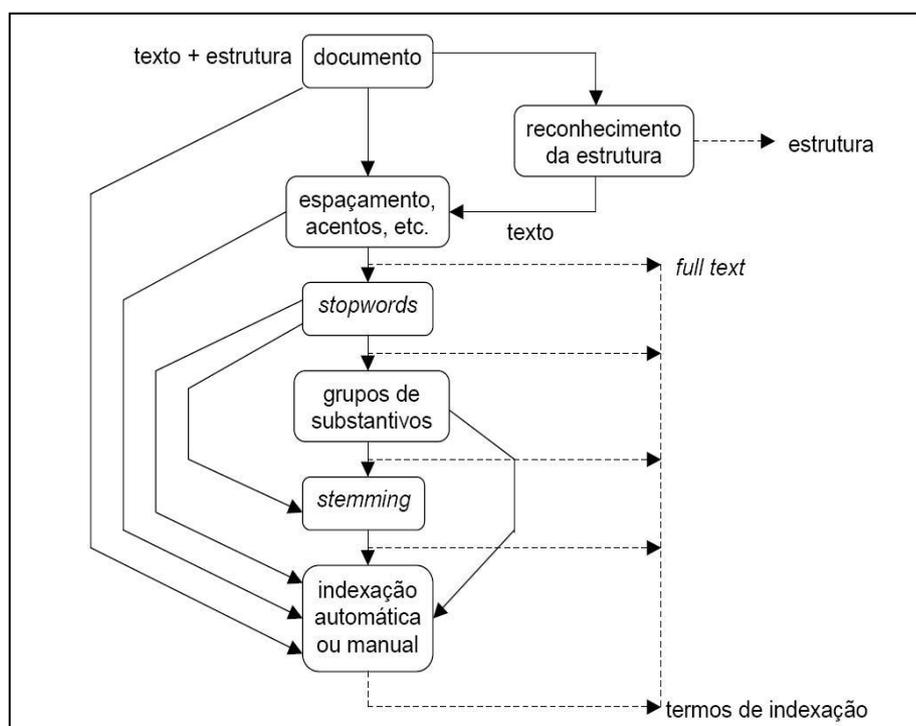
Por questões históricas e de limitações computacionais, os documentos de uma coleção são frequentemente representados por um conjunto de termos indexadores ou por palavras-chave. Essas palavras podem ser extraídas diretamente de um documento de texto ou especificadas por um indivíduo. Desconsiderando sua origem, quer sejam extraídas automaticamente, quer sejam geradas por um especialista, elas fornecem uma visão lógica do documento.

Os computadores modernos têm possibilitado a representação de documentos por seu conjunto completo de palavras. Nesse caso, Baeza-Yates e Ribeiro-Neto (1999) dizem que o sistema de recuperação utiliza uma visão lógica baseada no texto completo (uma representação). Entretanto, mesmo os computadores mais modernos, no caso de coleções, precisam reduzir o conjunto de palavras-chave representativas; isso pode ser acompanhado da remoção de *stopwords* (como artigos e conetivos), do uso de *stemming* (que substituem palavras

flexionadas por seus respectivos radicais), da identificação de substantivos (que eliminam adjetivos, advérbios e verbos) e da compressão das estruturas lógicas internas (índices e índices invertidos, por exemplo). Essas operações, chamadas de *operações textuais*, reduzem a complexidade da representação dos documentos, além de possibilitarem a conversão de um texto completo em um conjunto de termos indexadores.

Para Baeza-Yates e Ribeiro-Neto (1999, p. 5), “o documento completo é claramente a mais integral visão lógica de um documento, mas seu uso implica custos computacionais elevados”. Um conjunto resumido de categorias (geradas por um especialista humano) provê a visão lógica mais concisa de um documento, mas seu uso pode levar a uma recuperação de baixa qualidade. Em alguns casos, visões lógicas intermediárias (de um documento) podem ser utilizadas por um sistema de RI, como ilustra a Figura 2. Portanto, adotando algumas das representações intermediárias, o sistema de RI pode reconhecer estruturas internas normalmente encontradas em um documento (capítulos, seções, subseções, e outras).

Figura 2 - Visões lógicas intermediárias utilizáveis para sistema de RI



Fonte: Yates; Ribeiro-Neto, 1999, p. 6

Conforme ilustrado na Figura 2, o problema de representar logicamente um documento é contínuo. Essa representação pode variar desde o texto completo até uma especificação de alto nível, feita por um especialista humano. A escolha de uma representação, mais ou menos elaborada, depende da necessidade do sistema de RI a ser implementado. Serão apresentados os três modelos de RI: booleano, vetorial e probabilístico.

2.4.2 Modelo Booleano

O modelo booleano é o modelo de recuperação simples, baseado na teoria de conjunto e na álgebra booleana. Nesse modelo, um documento é representado por um conjunto de termos indexados. As buscas são realizadas através de expressões booleanas, compostas por termos ligados através dos conectivos lógicos AND, OR e NOT (e, ou e não) e, como resultado, apresenta-se os documentos que satisfaçam às restrições lógicas da expressão de busca.

2.4.3 Modelo Vetorial

O modelo Vetorial representa documentos, como um vetor de termos, em que cada termo possui um valor de importância associado, ou seja, um peso. A consulta feita pelo usuário, da mesma forma, é representada por um vetor. Segundo Baeza-Yates;Ribeiro-Neto (2011), cada consulta possui um vetor resultado, construído através do cálculo da similaridade baseada no ângulo (cosseno) entre o vetor, que representa o documento, e o que representa a consulta. Dessa forma, os vetores dos documentos podem ser comparados com o vetor da consulta e o grau de similaridade, entre cada um deles, pode ser identificado.

Esse modelo reconhece que o “uso de pesos binários é muito limitado e propõe um *framework* pelo qual é possível mensurar uma similaridade parcial”. (Baeza-Yates;Ribeiro-Neto, 1999, p. 27). Isso é alcançado pela associação de pesos não binários aos termos indexados das consultas e dos documentos. Esses pesos são utilizados, então, para medir o grau de similaridade entre cada documento armazenado no sistema e a consulta do usuário. Classificando os documentos recuperados, em ordem decrescente desse grau de similaridade, o modelo vetorial

leva em conta documentos que se igualam apenas parcialmente com a consulta. O principal efeito resultante é que o conjunto de documentos ranqueados é bem mais preciso que o apresentado pelo modelo booleano.

No modelo vetorial, cada documento é representado como um vetor de termos. Cada termo possui um valor associado, que indica o grau de importância (peso - *weight*) do documento.

2.4.4 Modelo Probabilístico

O modelo probabilístico trata do problema da recuperação de informação de uma visão probabilística. A partir de uma consulta do usuário, o conjunto dos documentos é dividido em quatro grupos, sendo eles: o conjunto de documentos relevantes, o conjunto dos documentos recuperados, o conjunto dos documentos relevantes recuperados e o conjunto dos documentos não relevantes e não recuperados. O que se busca, nesse modelo, é saber a probabilidade de um documento D ser ou não relevante para uma consulta. Essa informação pode ser obtida assumindo-se que a distribuição de termos, na coleção, seja capaz de informar a relevância provável para um documento.

Então, pode-se pensar no processo de busca como um procedimento de especificação das propriedades de um conjunto resposta ideal (o que é análogo à interpretação do problema de *clustering*). O problema é descobrir exatamente o que essas propriedades são, pois tudo que se sabe é que elas são termos indexados dos quais a semântica pode ser utilizada para caracterizar essas propriedades.

O modelo probabilístico é baseado no seguinte pressuposto: dada uma consulta q e um documento d_j em uma coleção, o modelo probabilístico tenta estimar a probabilidade do usuário considerar o documento d_j interessante (i.e. relevante). O modelo assume que essa probabilidade de relevância depende apenas da representação da consulta do documento. Além disso, o modelo assume que existe um subconjunto de documentos que o usuário prefere definir como a resposta para a consulta q . Esse conjunto de resposta ideal é chamado de R e deve maximizar a probabilidade global de relevância para o usuário. Documentos pertencentes ao conjunto R são ditos relevantes para a consulta; documentos fora dele são ditos não relevantes. Segundo Baeza-Yates;Ribeiro-Neto (1999), essa pressuposição é

bastante problemática porque não define explicitamente como computar as probabilidades de relevância. Na verdade, nem mesmo o espaço amostral, utilizado para a definição dessas probabilidades, é dado.

Dada uma consulta q , o modelo probabilístico associa a cada documento d_j , como uma medida de sua similaridade a consulta, a razão $P(d_j \text{ relevante para } q)/P(d_j \text{ não relevante para } q)$ que computa as probabilidades do documento d_j ser relevante para a consulta q . Tomando as probabilidades de relevância como o ranque, minimiza-se a probabilidade de um julgamento errôneo.

2.5 Sistema de Recomendação

Segundo Burke (2007, p. 377), os sistemas de recomendação (SR) são “agentes de informação personalizados que fornecem recomendações e sugestões de itens que possam ser de utilidade para o usuário”. Em um contexto de comércio eletrônico, estes produtos podem ser itens de compra. Já em um contexto de biblioteca digital, podem ser textos ou outros meios de comunicação relevantes para os interesses do usuário.

Muitas empresas, tais como *Amazon.com*, *YouTube*, *Netflix*, incorporam sistemas de recomendação como uma forma de personalizar seu conteúdo para usuários, objetivando a fidelização e, conseqüentemente, aumentando sua lucratividade. Na *Amazon.com*, sistemas de recomendação são utilizados para personalizar a loja *online* para cada usuário, por exemplo, mostrando programação de títulos a um engenheiro de *software* e brinquedos de bebê para uma nova mãe (CAZELLA; REATEGUI, 2005).

Os sistemas de recomendação surgiram como uma resposta à dificuldade das pessoas de realizarem escolhas em meio a grande variedade de produtos e serviços e as várias alternativas apresentadas. Segundo Ricci *et al.* (2011, p.2):

O desenvolvimento dos Sistemas de Recomendação se deu a partir de uma observação bastante simples: os indivíduos muitas vezes dependem de recomendações fornecidas por outros na tomada de rotina, decisões diárias. Por exemplo, ao selecionar um filme para assistir, os indivíduos tendem a ler comentários de críticos de cinema para auxiliar em suas escolhas.

Os sistemas de recomendação tentam prever quais são os produtos ou serviços mais adequados, com base nas preferências e restrições do usuário. Para completar essa tarefa, eles coletam dos usuários suas preferências, que podem ser explícitas, através das avaliações de produtos, ou implícitas, inferidas através da interpretação das ações do usuário (Ricci *et al.*, 2011).

Dessa forma, procuram facilitar a busca por conteúdo interessante ao usuário, prevendo itens que possam ser relevantes a ele. Nesse sentido, os Sistemas de Recomendação têm se mostrado um meio valioso para lidar com o problema da sobrecarga de informações. Esses sistemas apontam o usuário para o sentido de novas informações, ainda não exploradas, que possam ser de interesse para as suas tarefas.

Os sistemas geram recomendações, usando vários tipos de conhecimentos e dados sobre o usuário, tais como itens disponíveis e transações anteriores armazenadas em bancos de dados personalizados. Essas recomendações podem ser úteis ou não para o usuário. Para descobrir a relevância de uma sugestão, o usuário pode fornecer, imediatamente ou numa etapa seguinte, um retorno implícito ou explícito. Todas essas ações do usuário e feedbacks podem ser armazenados no banco de dados (perfil) e utilizados para a geração de novas recomendações nas interações próximas do sistema do usuário (SILVA *et al.*, 2015).

Nesse sentido, o sistema de recomendação envolve a construção de um modelo ou perfil de interesses do usuário, o que varia de acordo com a técnica de recomendação utilizada, sendo as principais delas a Filtragem Colaborativa, a Recomendação Baseada em Conteúdo, o Modelo Híbrido, entre outras. A seguir, são detalhados esses três modelos.

Por Filtragem Baseadas em Conteúdo (FBC) – recomendam, ao usuário, itens que sejam similares ao que ele preferiu no passado. Por exemplo, se um usuário classificou, positivamente, um filme como sendo do gênero romance, então, o sistema pode recomendar outros filmes do mesmo gênero. No caso de documentos textuais, por exemplo, basta que o conteúdo seja similar, isto é, mostrar presença de termos comuns para que dois textos sejam considerados similares (GARIN *et al.*, 2006).

Por Filtragem Colaborativa (FC) - consiste na recomendação de produtos ao usuário ativo de itens que outros usuários, com gostos similares aos seus, tinham no passado. Por exemplo, analisa-se a vizinhança do usuário, a partir da regra: "Se um usuário gostou de televisão e rádio, o outro usuário, que gostou de televisão, também poderá gostar de rádio". Esta é a razão pela qual refere-se a filtragem colaborativa como "pessoa-a-pessoa" correlação. Sendo assim, quando um sistema recomendar filmes, essa recomendação só será feita, a outro usuário, se este filme tiver sido bem avaliado por algum usuário com perfil semelhante ao seu. Usuários com perfil semelhante, neste caso, são usuários que, no passado, avaliaram os mesmos filmes de maneira similar (GARIN *et al*, 2006).

Demográfica - este tipo de sistema recomenda itens com base no perfil demográfico do usuário. Estes produtos recomendados podem ser produzidos para diferentes nichos demográficos, combinando os votos de usuários nesses nichos. Por exemplo, o usuário é redirecionado para *sites* específicos com base em seu idioma ou país.

Baseadas em Conhecimento - recomendam itens com base no conhecimento e sugerem produtos à base de inferências sobre as necessidades de um usuário e suas preferências. Este conhecimento pode funcionar, explicitamente, sobre algumas funcionalidades do produto para atender a necessidade do usuário e, em última instância, como o item é útil para ele.

Baseadas em comunidade - este tipo de sistema recomenda itens com base nas preferências dos usuários "amigos". Esta técnica segue o ditado popular "Diga-me quem são os seus amigos, que eu lhe direi quem você é". A evidência sugere que as pessoas tendem a confiar mais em recomendações de seus amigos do que em recomendações de indivíduos semelhantes, mas anônimos (RICCI, 2011).

Híbridas - combinam duas ou mais técnicas de recomendação para obter um melhor desempenho, tentando fortificá-las e superar suas desvantagens. Por exemplo, os métodos de FC têm problema para recomendar produto para o item novo, e a FBC recomenda itens para usuários incomuns, ou seja, pode recomendar itens independentes da quantidade de usuários. Sendo assim, FC é combinada com alguma outra técnica FBC numa tentativa de evitar o problema de classificação na inserção de um novo item. (CAZELLA; REATEGUI, 2005). A utilização do SR, nesta

pesquisa, tem como objetivo recomendar OAs provenientes de repositórios digitais, utilizando sistema de recuperação de informação.

Sendo assim, constata-se que os Sistemas de Recuperação e Recomendação de Informação trabalham com um grande volume de informação. Mas quais as diferenças dessas duas tecnologias? Como já apresentado, os Sistemas de Recuperação de Informação surgiram a partir da necessidade de se obter informações em um curto espaço de tempo; enquanto os Sistemas de Recomendação de Informação foram uma consequência da grande quantidade de conteúdo, o que acabava gerando uma sobrecarga de informação.

Por isso, os Sistemas de Recomendação são ferramentas de software e técnicas que fornecem sugestões de itens que sejam úteis para um usuário. Segundo Ricci *et al.* (2011), esses sistemas tentam prever quais são os produtos ou serviços mais adequados, com base nas preferências e restrições do usuário. Nos sistemas de recuperação da informação, por exemplo, o usuário descreve a sua necessidade de informação, através de uma consulta (query), a interação é provocada pelo usuário, baseando-se na percepção de uma necessidade do momento; já enquanto os sistemas de recomendação de Informação, por exemplo, mantêm um perfil dos interesses dos usuários, referem-se às preferências dos usuários e entregam as informações para as pessoas que realmente necessitam.

O fato é que esta pesquisa fazer uso de um motor de busca não significa que o sistema proposto seja apenas um novo sistema de Recuperação de Informação; mais que isso, o sistema obtém informações do perfil do usuário de maneira automática e implícita, sem a sua ação direta nesse processo. Além disso, uma vez que os *links* são recomendados, o usuário irá avaliá-los, e essa avaliação irá compor o seu perfil para futuras recomendações personalizadas.

2.7 Arquitetura Orientada a Serviços

A arquitetura orientada a serviços tem como características propor soluções para integração entre sistemas. Segundo Sommerville (2011, p. 355), “um *web service* é uma representação-padrão para algum recurso computacional ou de informações que pode ser usado por outros programas”, em que um serviço é um meio de entregar resultado ao solicitante sem que este seja responsável pela

execução. Eles foram propostos para permitirem a integração entre sistemas diferentes, da mesma organização ou não, por isso têm como forte característica a independência de plataforma.

A utilização desses serviços, para o desenvolvimento de sistemas, deu surgimento a uma nova disciplina da engenharia de *software*, ou seja, a arquitetura orientada a serviços. Nela os sistemas são projetados para utilizarem recursos, disponibilizados por *web services*, como por exemplo, um serviço de consulta à previsão do tempo ou de acesso aos contracheques da empresa. Esta abordagem traz vantagens no que tange à modularidade e reusabilidade, visto que cada serviço tem seu escopo bem delimitado e as responsabilidades são encapsuladas e divididas.

Com o objetivo de prover todas estas características, a W3C definiu um padrão baseado em *Extensible Markup Language* (XML). Para XML (2015), o desenvolvimento de *web services* é composto basicamente pelos seguintes protocolos: para encapsulamento de mensagens - *Simple Object Access Protocol* (SOAP); um para a descrição do serviço – *Web Services Description Language* (WSDL) e outro para descoberta – *Universal Description, Discovery and Integration* (UDDI). Em contraproposta a este padrão, surgiram os serviços baseados em REST, que utilizam exclusivamente o protocolo *Hypertext Transfer Protocol* (HTTP) para a troca de mensagens e fornecimento de recursos. De acordo com Liu (2015), trouxeram inúmeras facilidades para nossa vida e trabalho. Na próxima seção, a pesquisa abordará estas duas tecnologias orientadas a serviços.

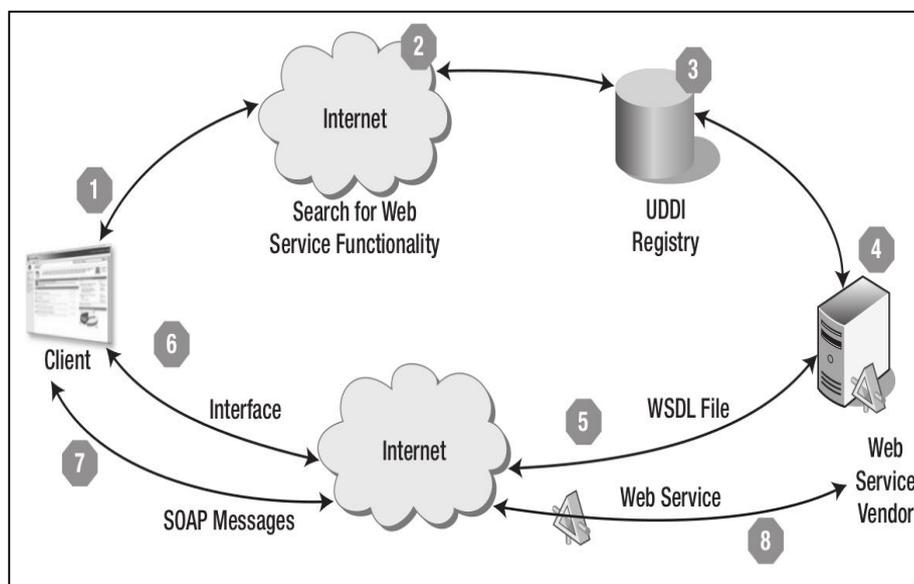
2.7.1 *Simple Object Access Protocol* - SOAP

SOAP é a tecnologia baseada em XML, especificada pelo *World Wide Web Consortium*, para o desenvolvimento de *web services*. Segundo o W3C (2015), é um protocolo destinado à troca de informações estruturadas em um ambiente descentralizado. Foi projetado para ser um protocolo de comunicação da camada de aplicação; logo, pode utilizar HTTP ou SMTP para transportar suas mensagens, através da Internet.

Todas as mensagens SOAP devem ser obrigatoriamente encapsuladas por um envelope, que deve conter um cabeçalho (*header*) e um corpo (*body*), sendo que

este último possui, de fato, a informação a ser transmitida.

Figura 3 - Ciclo de Vida de um serviço Web SOAP



Fonte: Peiris (2007)

Em um projeto de *web service* SOAP, é comum a utilização de outros dois protocolos complementares, o *WSDL* e o *UDDI*, ambos baseados em *XML*. O primeiro especifica a interface do serviço, métodos, tipo de dados, mensagens trocadas, entre outros; enquanto o segundo funciona como um catálogo de serviços e fornece um meio para a descoberta de novos *web services*. Na Figura 3, é possível observar como se dão as interações entre estes protocolos, ou seja, o cliente busca um serviço, acessando a um *UDDI*; em seguida, baixa o *WSDL* do provedor escolhido e, por último, acessa um *web service*, enviando mensagens SOAP.

2.7.2 Representational State Transfer - REST

REST é um padrão arquitetural para aplicações web definido por Fielding (2000). Segundo o autor, o REST ignora os detalhes de sintaxe de protocolo e de implementação, focando nas regras de negócio, na comunicação entre os componentes e suas restrições. Suas características básicas são a sua essência

cliente-servidor, a comunicação sem estado, a utilização de *cache*, a interface uniforme e a utilização de camadas.

Assim como no SOAP, as aplicações REST são independentes de plataforma e de linguagem e executam em cima do protocolo HTTP. No entanto, não há um formato específico para a representação dos dados, sendo possível utilizar tanto o XML, quanto *JSON*. Diante do exposto, Richardson (2008) propõe a utilização destas aplicações como serviços web, o que foi chamado de *web services RESTful*. Estes serviços utilizam-se da infraestrutura disponível na web, através do protocolo HTTP, e devem atender às propriedades abaixo: todos os recursos devem ser endereçáveis por uma URI (*Uniform Resource Identifier*) com nomes significativos; os serviços não devem guardar estado, ou seja, todas as informações necessárias para o processamento devem estar contidas na requisição; deve ter uma interface uniforme e utilizar os verbos HTTP para diferenciar as operações; e, deve prover *links* para permitir a navegação por outros serviços correlacionados.

No que se trata da uniformização da interface, como o autor cita, devem ser utilizados os métodos *GET*, *PUT*, *POST*, *DELETE*, *HEAD* e *OPTIONS*, do protocolo HTTP. É possível atender a todas as operações de um *CRUD*, ou sejam, *Create*, *Read*, *Update and Delete*, utilizando os verbos da seguinte forma: *GET* para recuperar, *PUT* para criar um novo, *POST* atualizar e *DELETE*, para deletar um recurso. Os demais métodos devem ser utilizados para recuperar metadados, no caso do *HEAD*, ou para identificar a que chamadas um recurso suporta, com o *OPTIONS*.

De acordo com Wagh (2012), estas características dão vantagens ao REST no que diz respeito a um tráfego menor de dados, flexibilidade e maior simplicidade no desenvolvimento dos serviços. No entanto, segundo o autor, o SOAP é mais seguro, por oferecer um protocolo específico para este fim, chamado de WS-Security.

2.8 Mineração de Texto

Os avanços tecnológicos, na área de compartilhamento e armazenamento de dados, fizeram com que o volume de informações, no formato digital, crescesse em proporções antes inimagináveis. Segundo Kuechler (2007, p. 86) “80% desses dados não estão em formato estruturado, sendo que uma grande parte deles são

textos”. Essas informações incluem: e-mails, arquivos eletrônicos gerados por softwares editores de texto, páginas *Web*, campos textuais em banco de dados, entre outros. Entretanto, esse formato de documentos foi criado para serem visualizados por seres humanos e não são adequados para a manipulação das informações neles contidas por sistemas computacionais.

Em geral, estes conteúdos são muito relevantes para as organizações, pois, segundo Han (2011), constituem um importante repositório organizacional, que envolve notícias, artigos científicos, livros, bibliotecas digitais, e-mails e páginas *Web*, projetos, estratégias e o próprio conhecimento adquirido.

Para Rezende, Marcacini e Moura (2011), a organização inteligente dessas coleções textuais é de grande interesse para a maioria das instituições, pois agiliza os processos de busca e recuperação da informação. Entretanto, o volume de dados textuais armazenados é tal que extrapola a capacidade humana de, manualmente, analisá-lo e compreendê-lo por completo.

Para Lopes (2004, p. 7), a Mineração de Texto (MT) “pode também ser definido como um conjunto de técnicas e processos que se prestam a descobrir conhecimento inovador nos textos”. Esta nova tecnologia está sendo empregada, atualmente, em projetos de diversas áreas. Dessa forma, pode-se entender a mineração de texto como sendo a aplicação de um conjunto de diferentes técnicas em dados não estruturados, com o objetivo de obter informações que, muitas vezes, podem não estar presentes, de maneira explícita, nestes documentos. Inspirado pelo *Data Mining* ou mineração de dados, que procura descobrir padrões emergentes em banco de dados estruturados, a mineração de textos tem como objetivo a extração de conhecimento úteis em dados não estruturados ou semiestruturados.

Para Konchady (2006), tanto mineração de dados quanto mineração de texto buscam por informações escondidas e empregam algoritmos semelhantes de Inteligência Artificial, aprendizagem de máquina, e estatística. Contudo, enquanto a mineração de dados lida com dados estruturados, a mineração de texto lida com dados não estruturados. Sendo assim, esta última representa uma extensão da área de *Data Mining*, focada na análise de textos.

O processo de descoberta de conhecimento em textos é uma evolução natural da recuperação de informações, já que os sistemas de recuperação de informação passaram a adotar algumas técnicas de análise de informações e de

aprendizado de máquina, muitas das quais provenientes da área de descoberta de conhecimento em bases de dados. Assim, ao invés do usuário ter que analisar quais dos documentos retornados são realmente relevantes, o próprio sistema faria essa análise e retornaria as informações de forma condensada e resumida (WIVES, 2002).

2.9 Algoritmo *Term Frequency – Inverse Document Frequency* (TF-IDF)

O algoritmo TF-IDF é um algoritmo popular para a indexação, é a medida TF-IDF que extrai as palavras-chave que aparecem com frequência em um documento, mas que não aparecem, com frequência, no restante dos outros documentos. TF (*Term Frequency*) mede o número de vezes que uma palavra aparece no documento atual, que pode refletir a frequência da palavra; enquanto o IDF (*Inverse Document Frequency*) reflete o número de documentos em que a palavra ocorre. Conseqüentemente, o valor IDF pode indicar o conteúdo de informação de uma palavra. A principal abordagem utilizada para o cálculo de peso de um termo é o proposto por Salton (1988) TF-IDF. Sendo TF igual a:

$$\text{TF}(t_k, d_j) = \frac{f_{k,j}}{\max_z f_{z,j}}$$

Fonte: adaptado de Salton (1988, p.518)

Em que máximo é calculado sobre as frequências f_z, j de todos os termos t_k que ocorrem no documento d_j . E $f_{k, j}$, é a frequência do termo k no documento d_j .

A frequência inversa do documento (IDF) foi proposta, inicialmente, por Jones (1972). A ideia, por trás deste algoritmo, é de que um termo de consulta, que ocorre em muitos documentos, não é um bom representante e, assim, deve ser dado menor peso do que os que ocorrem em apenas alguns documentos. Sendo assim, N é igual ao número de documentos da coleção e n_i é igual a quantidade de documentos em que o termo ocorre pelo menos uma vez.

$$\text{idf}_i = \log \frac{N}{n_i}$$

Fonte: Adaptado de Salton (1988, p.518)

A essência do funcionamento do método TF-IDF consiste em determinar o quanto é relevante uma palavra em relação a um conjunto de documentos. As palavras, que são comuns em um pequeno grupo de documentos, ou em apenas um documento, tendem a ter pesos TF-IDF mais elevadas do que as palavras comuns a todos os documentos.

2.10 Algoritmo *Term Frequency * Proportional Document Frequency* (TF*PDF)

O algoritmo TF*PDF foi proposto em 2001 por Bun e Ishizuka, o TF*PDF é uma abordagem que busca atribuir pesos mais significativos aos termos mais frequentes na coleção de documentos. Diferente da atribuição convencional de peso trabalhado no método TF-IDF, no algoritmo TF*PDF, o peso de um termo é linearmente proporcional à frequência, e exponencialmente proporcional à relação do documento que contém o termo. Sendo assim, o algoritmo PDF é representado pela fórmula a seguir: **PDF = $\exp(ni/N)$** Bun; Ishizuka (2002).

Desde sua proposta inicial, em 2001 por Bun e Ishizuka, o TF*PDF tem se demonstrado uma excelente ferramenta na mineração de texto para a detecção de tópicos em um documento ou em um conjunto deles. Diversas pesquisas, nos últimos anos, têm demonstrado a eficiência deste algoritmo como aquelas realizadas por: Jahnvi; Radhika (2012); Zhe *et al* (2012); Ren *et al* (2011); Kaur; Gupta (2012); Ma (2011) entre outros.

Entretanto, o objetivo do uso da ponderação de termos, nesta pesquisa, é de identificar os principais tópicos debatidos no fórum. Além do mais, com a extração das *stopwords*, é possível eliminar grande parte dos termos que não apresentam significância ao contexto da discussão. Nesse sentido, para que se possa identificar as palavras-chave (*hashtags*), faz-se necessário utilizar um algoritmo que atribua pesos mais significativos aos termos mais frequentes na coleção de documentos, ou seja, nos tópicos do fórum.

Pelo fato de dar menor peso aos termos que são frequentes dentro de um documento, mas não tão frequente na coleção, o TF-IDF não se torna adequado para a resolução do problema em questão. Sendo assim, esta pesquisa faz uso do algoritmo TF*PDF para mineração de texto a fim de extrair os principais temas debatidos nos fóruns de aprendizagem.

Nesse sentido, para se calcular a frequência do termo (TF), foi adotado uma abordagem diferente. Como não é feito cálculo de similaridade, entre os documentos, a coleção de documentos (os diferentes tópicos do fórum de discussão) são agrupados e os pesos TFs são calculados para todos os termos em relação a um único documento. Uma vez feito isto, o cálculo do peso PDF é realizado, levando em consideração a frequência do termo em relação ao conjunto de documentos.

2.11 Considerações do Capítulo

Neste capítulo, apresentam-se a fundamentação teórica da pesquisa realizada. Sendo assim, foram abordados:

Os principais AVAs disponíveis no mercado, quais os locais de utilização desses AVAs, destacando principalmente a importância dos fóruns de discussões dentro dos AVAs. Dentre os AVAs, apresentados no capítulo, a ênfase principal foi para o *Moodle*, em virtude de que a maioria das instituições de ensino estão fazendo o uso desse AVA e, nesta pesquisa, auxiliando a sua aplicabilidade;

A fundamentação teórica sobre OA quanto a sua definição, taxonomia, como concebidos para ajudar os professores, onde são armazenados, descrição por meio de metadados e evolução tecnológica;

As principais funcionalidades do repositório digital, destacando: as definições, padrões internacionais, repositório institucionais, as vantagens de livre acesso e as suas características.

A revisão teórica sobre sistema de recomendação, explicando o contexto do comércio eletrônico, algumas empresas que utilizam sistema de recomendação para comercializar seus produtos, as principais técnicas utilizadas no sistema de recomendação;

Como fornecer aos usuários acesso à informação de seu interesse, através de sistemas de recuperação de informação, destacando: definições, apresentando uma visão lógica de um documento e os principais modelos utilizados pelo sistema de informação;

Os modelos clássicos utilizados no processo de RI: modelo booleano, modelo vetorial e modelo probabilístico; na implementação do sistema de

recomendação de OA nos fóruns de um AVA, o modelo escolhido foi modelo vetorial, em virtude da aplicação da técnica de atribuição de pesos para os documentos.

Revisão teórica sobre arquitetura orientada a serviços, sua definição, quais as suas vantagens, protocolo e padrões;

A Mineração de Texto, destacando: definição, bem como, sua importância foi justificada através de pesquisadores da área;

Revisão teórica sobre os algoritmos TF-IDF e TF*PDF, esses amplamente utilizados por aplicações para extração de palavras-chave de documentos. Mostrou-se o cálculo de peso de cada algoritmo e sua aplicabilidade.

Portanto, para o desenvolvimento do sistema de recomendação de OA, nesta pesquisa, utilizou-se algoritmo TF*PDF para mineração de texto como o objetivo de extrair palavras-chave, dentro dos fóruns do AVA. A decisão de utilizar o algoritmo TF*PDF se mostrou mais acertada, pois sua aplicação é mais indicada para atribuição de pesos, em fragmentos de textos, que tratam de uma mesma temática (extração de *Hot Topics*), diferente do TF-IDF, que trata textos sobre temas heterogêneos.

No próximo capítulo, foram apresentados os trabalhos relacionados, que compõe o estado da arte da presente pesquisa.

3 TRABALHOS RELACIONADOS

Neste capítulo, são descritas as pesquisas relacionadas com esta tese sobre modelo de recomendação, recomendação de OA e algoritmo de recomendação e integração ao AVA, levando-se em consideração alguns aspectos gerais que representam necessidades a serem supridas neste trabalho.

Considerando-se estas lacunas, observou-se que seria interessante propor um mecanismo que aproximasse os professores e alunos, que estivessem em ativa participação em um AVA, dos OAs disponíveis nos repositórios. Esta aproximação leva em consideração os indícios de seus interesses registrados em vários formatos, em diversas ferramentas disponíveis em AVAs. É notório que os usuários do AVAs anotam e fazem isto em fóruns, *chats*, diários de bordo, na descrição dos seus perfis, entre outros. Sendo assim, as pesquisas foram analisadas com o objetivo de justificar esta abordagem de recomendação de OA para alunos e professores, com base nos textos do fórum de um AVA.

Na pesquisa dos autores Manouselis, Vuorikari e Van Assche (2007), houve a proposta de um cálculo para recomendações automáticas para aluno, baseado em seu histórico de navegação recente, bem como a de explorar semelhanças e diferenças entre as preferências do utilizador e entre o conteúdo dos recursos de aprendizagem. A relevância deste trabalho consiste na recomendação automática de OA para alunos. Entretanto, como crítica, observa-se que esse trabalho não realiza a integração entre o AVA e o repositório digital de OAs.

Já a pesquisa de Betancur, Moreno e Ovalle (2009) apresenta um modelo de recomendação e recuperação de OAs, com objetivo de complementar um Sistema Tutorial Inteligente, a fim de aumentar a interação professor e aluno com os cursos em que estão cadastrados. Com o objetivo de realizar a recomendação dos OAs, o modelo utiliza-se de duas estratégias: para professores, o modelo proposto baseia-se em assuntos desenvolvidos pelo professor no curso; já para os alunos, nas suas características do perfil definidas pelos próprios discentes, através de interface. Positivamente, esta pesquisa realizou a criação de modelo de recomendação de OA. Mas, por outro lado, os autores não integraram o modelo de recomendação de OAs, levando-se em consideração os registros deixados pelos alunos dentro dos AVAs. Assim, neste sentido, o modelo não recomenda OA dentro dos fóruns de um AVA.

Os autores Ferreira e Raabe (2010) desenvolveram o LorSys (*Learning Object Recommender System*), um sistema de recomendação de OA, no formato *Sharable Content Object Reference Model (SCORM)*, voltado ao Sophia/Moodle. O LorSys utiliza a técnica de recomendação híbrida para localizar OA dentro do ambiente Sophia/Moodle. Sendo assim, o estudante visualiza os OAs recomendados por meio de um bloco adicionado na interface dos cursos. Consiste em ponto relevante desta pesquisa, a recomendação de OA dentro do AVA. Entretanto, a proposta do LorSys restringe-se a recomendar OA, do repositório disponível, somente dentro do ambiente Sophia/Moodle, ou seja, não recomenda OA disponibilizados em outros repositórios e não apresenta nenhum algoritmo para recomendação dos OAs no AVA em relação às postagens efetuadas pelo usuário nos fóruns.

Nesse sentido, os autores Vieira e Nunes (2011) desenvolveram o DICA, um sistema de recomendação de OA, que busca informações no Banco Internacional de Objetos Educacionais (BIOE) para recomendar objetos de aprendizagem, de acordo com as semelhanças existentes entre eles, através da utilização dos algoritmos *K-Means* e algoritmo *Term Frequency – Inverse Document Frequency (TF-IDF)*. Destaca-se como relevante, neste trabalho, a recomendação do OA dentro de um repositório, utilizando-se de algoritmo de recomendação. No entanto, a proposta do DICA tem característica de recomendação de OA, localizado somente dentro do repositório BIOE, não realizando nenhuma integração desse BIOE com um AVA.

Na pesquisa de Casali (2011), foi desenvolvido um protótipo de um sistema de recomendação com o objetivo de ajudar o usuário a localizar OA repositório Ariadne, de acordo com perfil do usuário, considerando-se suas características e preferências. Esta investigação primou pelo desenvolvimento de uma arquitetura *multiagente* para recomendação de OA. Mas, infelizmente, a pesquisa não realiza nenhuma integração entre o repositório Ariadne e um AVA. Além disso, foi elaborado um formulário na web onde o usuário tem que acessar para localizar os OAs de acordo com suas características e preferências descritas pelo próprio usuário, por meio de uma interface gráfica.

A tese de Rodrigues (2012) propõe um modelo para integração de repositórios de objeto de aprendizagem com Coletânea de Entidades de Suporte ao

uso de Tecnologia na Aprendizagem (CESTA2¹) a um AVA. A tese visa criar condições para que os OAs possam ser pesquisados e utilizados pelos professores e alunos através de *plugin*, disponível no próprio ambiente virtual. O modelo proposto demonstrou como esses dois sistemas podem ser integrados para permitir aos seus usuários acessar a interface do ambiente, realizar a busca no repositório. Contudo, nesta pesquisa, o usuário não recebe nenhuma recomendação dos objetos, dentro dos fóruns, porém observa-se que não é um problema da pesquisa, pois o objetivo desta foi apresentar a integração do ambiente virtual com repositório e como essa integração pode facilitar a busca por objetos.

Baseado no aporte teórico descrito, nesta pesquisa de tese, a tabela 2 apresenta o sumário dos trabalhos relacionados descritos e a referente pesquisa acerca das técnicas, mecanismos, integração ao AVA e recomendação nos fóruns utilizado ou “X” quando representa aspectos relevantes e “-” quando o trabalho analisado não representa ênfase aos aspectos.

Tabela 2 - Sumário dos trabalhos relacionados

Trabalhos relacionados Vs. Referente pesquisa de tese							
Aspectos relevantes	Manouselis; Vuorikari e Van Assche. (2007)	Betancur; Moreno e Ovalle (2009)	Ferreira e Raabe (2010)	Vieira e Nunes (2011)	Casali (2011)	Rodrigues (2012)	Referente pesquisa de tese
Algoritmo	X	-	-	X	-	-	X
Técnicas de Recomendação	X	X	X	X	X		X
Integração com AVA	-	X	X	X	-	X	X
Método para recomendação automática de OA	X						X
Recomendação de dentro fóruns de um AVA	-	-	-	-	-	-	X

Fonte: o autor (2015)

A Tabela 2 ilustrou um sumário dos trabalhos relacionados. Assim, a utilização de um sistema de recomendação integrado ao AVA, ainda é um obstáculo a ser superado, pois os alunos e colaboradores não sabem localizar esse

1 Disponível em: <http://cesta2.cinted.ufrgs.br/xmlui>

objeto de dentro dos fóruns do AVA. Dessa forma, a tese propõe um sistema computacional para recomendação de OA, baseado em postagens extraídas de dentro dos fóruns de um ambiente, tendo como contribuições: Método para recomendação automática de OA para professores e aluno; Técnica de mineração de textos; Técnicas de mineração de texto utilizando algoritmo TF*PDF; e, Integrar AVA com repositório digital, que serão assuntos abordados na próxima seção.

3.1 Contribuições

As principais contribuições nesta pesquisa transdisciplinar no desenvolvimento da tese são:

Método para recomendação automática de OA para aluno: já existente, como a pesquisa de Manouselis, Vuorikari e Van Assche (2007), mas não propõem recomendação automática com base nas postagens dos alunos e professores nos fóruns de um AVA.

Técnica de mineração de textos: também já existente com as pesquisas de Klemann (2009), Azevedo (2010), Macedo (2011), mas, não propõem recomendação de OA, com base nas palavras-chave extraídas de dentro dos fóruns de discussões de um AVA;

Técnicas de mineração de texto utilizando algoritmo TF*PDF para extração de palavras-chave dentro de um documento já existente como as pesquisa de Jahnvi e Radhika (2012); Zhe; Dong; Qi, Jianyi; Xinxin (2012); MA (2011), tais autores não propõe uma recomendação de OA baseada em palavras-chave, extraídas das postagens feitas por alunos e professores dentro de um AVA;

Integrar AVA com repositório digital para oferecer aos alunos e professores OA já existente, como a pesquisa de Rodrigues (2012), mas não oferecer OA com base nas postagens de alunos e professores.

Especificamente, esta pesquisa integra método para recomendação automática de OA para aluno e professores, aplicação de mineração de dados para sistema gestão educacional, técnica de mineração de textos, utilizando TF*PDF e integrar AVA com repositório digital, isso tudo dentro da área de mineração de textos e recomendação de informação. A contribuição, para estas áreas, é relacionada à

proposta para utilização de um sistema de recomendação de OA dentro dos fóruns de um AVA. A seguir tem-se as considerações propostas para o capítulo.

3.2 Considerações do Capítulo

Neste capítulo, foram apresentados os trabalhos relacionados com a tese. Dentre estes selecionados para esta pesquisa, distinguiram-se, como fundamentais, as pesquisas relacionadas: técnicas de recomendação; mecanismo de recomendação do OA; integração com ambiente virtual e recomendação de dentro dos fóruns de um AVA, visto que consideram-se os trabalhos selecionados para o estudo e com base na questão de pesquisa e nos objetivos desta tese; sendo assim, as principais contribuições esperadas são:

Apresentar um mecanismo de recomendação de OA, utilizando um algoritmo TF*PDF;

Utilizar técnicas de mineração de texto para processar as postagens em fóruns de um AVA;

Integrar um AVA e repositório digital para recomendar OA, utilizando técnicas de mineração de texto, aproximando os OA's do que é discutido no fórum;

Facilitar a localização de OA por parte dos professores e alunos em repositórios digitais em que estes estejam armazenados;

Recomendar OA, de acordo com discussões no fórum de um AVA;

E, finalmente, criar uma ferramenta que fomente as discussões em fóruns.

Portanto, o desafio desta tese é o desenvolvimento de um sistema para recomendação de OA, baseadas em postagens extraídas de um AVA para fins educacionais na EAD. No próximo capítulo, será apresentado o projeto do sistema de recomendação de OA, baseado em postagens extraídas de dentro dos fóruns de uma AVA.

4 METODOLOGIA DA PESQUISA

A operacionalização desta pesquisa foi realizada na Universidade Estadual do Maranhão - UEMA, por meio do Centro de Ciências Tecnológicas – CCT; para melhor compreender o universo da pesquisa, faz-se necessário um breve histórico. O CCT tem sobre a sua responsabilidade seis Cursos de Graduação: Engenharia Civil, Engenharia Mecânica, Engenharia de Computação, Arquitetura e Urbanismo, Curso de Formação de Oficial Bombeiro e Engenharia da Produção. Nesse contexto, foi selecionado o Curso de Engenharia de Computação em virtude de alguns docentes já estarem utilizando o AVA como ferramenta complementar para suas disciplinas presenciais.

Esse curso tem a finalidade de formar profissionais para atender às demandas dos diversos setores da sociedade, principalmente no que se refere à Engenharia de Software e Tecnologia da Informação, Automação e Controle, Telemática e Telecomunicações. Apresenta organização em semestre, com carga horária de 4.650 horas, distribuídas em dez períodos, envolvendo disciplinas de núcleo comum, específicas e estágio supervisionado.

Esta pesquisa tem como o objetivo geral propor um sistema computacional para recomendação de objeto de aprendizagem, baseado em postagens feitas de dentro dos fóruns de um ambiente virtual de aprendizagem. Para atingir esse objetivo, parte-se para dos objetivos específicos que são os seguintes: 1) realizar o levantamento bibliográfico dos principais modelos e técnicas de sistema de recomendação de informação; 2) identificar e analisar os fundamentos da área de recuperação da informação, mais especificamente mineração textual e sistemas de recomendação; 3) apresentar arquitetura de sistema de recomendação, baseada em postagens extraídas do fórum de um AVA, por mineração textual; 4) demonstrar um sistema de recomendação de OA, baseado em uma arquitetura *Web Services*; 5) apresentar a integração do sistema de recomendação ao AVA; 6) disponibilizar ao usuário de um AVA, através dos fóruns, acesso ao sistema de recomendação; 7) verificar se as palavras-chave, extraídas pelo sistema de recomendação, estão de acordo com as discussões realizadas no fórum de um AVA; 8) investigar se o sistema de recomendação disponibilizou OA de acordo com discussões no fórum de um AVA; e 9) avaliar se os OAs, recomendados pelo sistema, estão de acordo com

as postagens realizadas nos fóruns. Os resultados da pesquisa, quanto aos dois primeiros objetivos específicos, são relatados no capítulo 2; quanto ao terceiro, quarto, quinto e sexto, são descritos nos capítulos 3, 4, 5 e 6; O sétimo, oitavo, nono e o décimo são descritos nos capítulos 6 e 7 desta tese. Por se tratar de um número limitado de professores e alunos pesquisados, as análises desses resultados foram apenas qualitativas.

4.1 Sujeitos da Pesquisa

Especificamente, no curso de Engenharia de Computação, a pretensão é trabalhar com alunos e professores. Foram cadastrados, no AVA/Moodle, cinco professores para interagir e avaliar as atividades desenvolvidas nos fóruns e 25 alunos. As disciplinas avaliadas, para esta amostra, foram: Modelagem Banco de Dados e Interface de usuários e sistemas computacionais, com carga horária de 60 horas, que foram ministradas no primeiro semestre de 2016.

A escolha foi devido ao envolvimento dos alunos e professores com a proposta apresentada para o desenvolvimento da tese nesta pesquisa. A prática adotou ainda algumas atividades que clarificam e ajudam a alcançar a compreensão da totalidade, como:

a) Elaboração de instrumentos para coletar dados – foi elaborado no ambiente do AVA *Moodle* para cada sujeito da pesquisa, como também elaboram-se, roteiros dispostos na aplicação da recomendação dos OAs.

b) Tabulação dos dados coletados – a utilização do método matemático e estatístico é de extrema relevância para a compreensão do fenômeno em questão.

Foram analisadas as linhas e as entrelinhas dos objetivos desta pesquisa para o CCT e para o curso de Engenharia de Computação.

4.2 Instrumento de Coleta de Dados

O instrumento de coleta de dados utilizado foi a escala de *Likert* que consiste em um conjunto de itens apresentados, em forma de afirmações, perante os quais se solicita uma reação dos indivíduos. São feitas afirmações e, para cada uma, pede-se ao indivíduo que manifeste sua reação, escolhendo pontos da escala

(SAMPIERI, 2006). A escala *Likert* pode ser utilizada de maneira autoadministrada, em que o indivíduo marca sua resposta. O escalonamento tipo *Likert* consiste em um conjunto de afirmações, no qual ao respondente é solicitada a sua opinião em termos do grau de concordância ou discordância.

Nesta pesquisa, utiliza-se a escala de *Likert* com cinco graus para registrar a manifestação dos alunos e professores quanto ao grau de concordância, relacionado às assertivas propostas nos questionários. Dessa maneira, tem-se como objetivo qualificar o sistema de recomendação. Por isso, a escala de *Likert* foi associada a um valor numérico da seguinte maneira: (1) Discorda muito, (2) Discorda, (3) Nem concorda nem discorda, (4) Concorda e (5) Concorda muito.

Nesse sentido, realizou-se a elaboração dos questionários com questões abertas e fechadas, conforme Anexos 1 e 2. As questões fechadas são as que contêm alternativas de resposta previamente delimitadas, tais como várias alternativas, hierarquização de opções e múltiplas respostas. Esse tipo de questão apresenta algumas vantagens: facilidade de codificação; requer menor esforço e requer menor tempo de resposta; já as desvantagens são: limitação das respostas; não existe o 'porquê' das intenções e não se aprofunda no contexto.

Enquanto as questões abertas não se delimitam, a priori, às alternativas de resposta, porque o número de categorias de resposta é muito elevado, em teoria é infinito. São características vantajosas das questões abertas: aprofundamento de uma opinião, existência do "porquê" das intenções e serem explicativas; e suas desvantagens são: dificuldades de codificação, apresentam distorções e requerem um maior tempo de resposta.

4.3 Procedimentos para Coleta de Dados

No início do primeiro semestre de 2016, professores e alunos foram informados sobre os objetivos específicos da pesquisa. Após preencher o formulário, conforme Anexo A, iniciou-se o cadastramento no AVA/Moodle dos participantes e das disciplinas que fizeram parte da pesquisa. Para avaliar a recomendação do sistema, os participantes foram orientados a interagir nos fóruns de discussões. Assim, nos primeiros 20 (vinte) dias, foi trabalhado o incentivo às postagens no fórum do AVA pelos participantes e nos outros 40 dias, o uso do sistema de

recomendação, totalizando 60 dias, ou seja, dois meses para a análise do sistema de recomendação. Sendo assim, depois do acesso ao ambiente, os alunos e professores avaliaram cada OA recomendado.

Uma vez superada a fase de utilização do sistema de recomendação, a etapa que se segue é a de coleta de dados, que implica a obtenção de informações pertinentes sobre variáveis, acontecimentos, contextos ou objetos envolvidos na pesquisa. Fazem parte do processo de coleta as seguintes atividades, profundamente correlacionadas: a) selecionar um instrumento ou método de coleta dos dados ou desenvolver um; b) aplicar esse instrumento ou método para coletar os dados; c) preparar observações, registros e medições obtidas.

O processo de coleta de dados dispõe de uma série de instrumentos ou técnicas, sejam qualitativas ou quantitativas, e em um mesmo estudo pode se valer de ambos os tipos. Para a validação da medição, foi realizado um estudo para amostra, ou seja, 25 alunos e 5 professores. Nesta pesquisa, os dados foram coletados a partir dos questionários aplicados aos professores, conforme Apêndice A; e aos alunos, conforme Apêndice B, por meio, da *Web* através do *Google Form*. O *Google Forms* é uma plataforma do *Google* que permite aos seus usuários criarem formulários dinâmicos com perguntas e respostas personalizadas, podendo ser utilizadas as opções de múltipla escolha com escolha única ou múltipla, texto livre, datas, respostas longas ou curtas etc, bem como a inserção de figuras e vídeos nos formulários. A coleta das repostas é feita de maneira automática pela ferramenta com a organização em gráficos e a possibilidade de exportação dos resultados para planilhas.

4.3 Procedimentos para Análise de Dados

Para análise dos dados, utilizou-se a técnica não paramétrica de provas de hipóteses. Uma das vantagens dessa técnica é a sua aplicabilidade em pequenas amostras. A sua aplicação não exige suposição quanto à distribuição da população da qual se tenha retirado amostras para análises. Atualmente, existem várias provas estatísticas que utilizam técnicas não paramétricas. Para o desenvolvimento desta análise, buscou-se uma técnica que utilizasse nível de mensuração ordinal, com

duas amostras independentes (professores e alunos), extraídas da mesma população. Nesta pesquisa, aplicou-se o teste U, de Mann-Whitney (FONSECA, 1996), que não exige nenhuma consideração sobre as distribuições populacionais e suas variâncias. A seguir será apresentada a seção sobre recursos tecnológicos.

4.4 Recursos Tecnológicos

Os recursos tecnológicos são bastante dinâmicos, também exigindo que as tecnologias utilizadas, em seu desenvolvimento, sejam capazes de se executar em diversos ambientes e, acima de tudo, possibilitar uma fácil manutenção. Como objetivo de desenvolver um sistema de recomendação heterogêneo, é importante que haja algum mecanismo que facilite a integração desse sistema. Portanto, o uso da tecnologia de *Web Services*, para integração do AVA como o sistema de recomendação, possibilitou o seu desenvolvimento por meio dos seguintes recursos tecnológicos: Servidor Apache, Linguagem PHP, Banco de Dados (MySQL) e *microframework slim*.

4.5 Considerações do Capítulo

Neste capítulo, apresentou-se a metodologia para coleta, análise e recursos tecnológicos do sistema recomendação de OA, baseado em postagens extraídas de dentro dos fóruns de um AVA. Inicialmente, foi realizada a escolha do local onde foi desenvolvido a pesquisa, ou seja, foram identificados: o universo da pesquisa, o sujeito da pesquisa, os procedimentos para coleta e análise de dados e os recursos tecnológicos.

O objetivo do projeto do sistema de recomendação foi facilitar a compreensão das problemáticas correlacionadas ao desenvolvimento desta pesquisa, além de servir de piloto para a construção de uma versão mais refinada, com características de produto, no futuro. No capítulo a seguir, será apresentada a prototipagem do sistema de recomendação.

5 PROTOTIPAGEM DO SISTEMA DE RECOMENDAÇÃO DE OA DENTRO DO FÓRUM DE UM AVA

O uso das Tecnologias de Informação e Comunicação (TICs) trouxe um novo formato para que professores e alunos pudessem interagir. Tais mudanças permitiram uma nova dinâmica, que provocou uma alteração de paradigma no sentido que a individualização cedeu lugar à colaboração e a aprendizagem independente passou a ser sustentada por experiências colaborativas (PIVA; PUPO; GAMEZ; OLIVEIRA, 2011).

Uma das principais ferramentas dos AVAs, voltadas à construção do conhecimento de forma colaborativa, são os fóruns de discussões. Estes espaços fomentam discussões e troca de ideias em torno de temas propostos por seus participantes. É também um instrumento que permite a cada participante submeter sua colaboração referente ao tema proposto, buscando, assim, o entendimento mútuo.

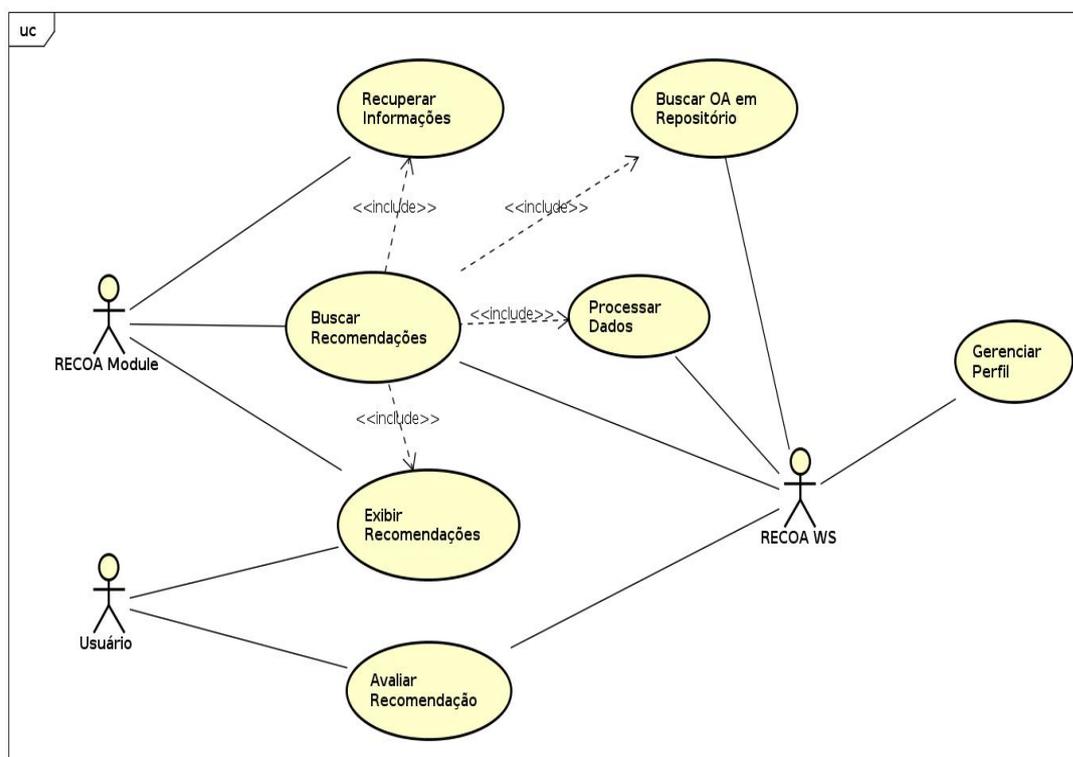
Segundo Silva (2011), o fórum é uma ferramenta assíncrona que representa um espaço para debates, no qual pode ocorrer o entrelaçamento de muitas vozes para a construção e desconstrução de pensamentos, para questionar e responder dúvidas, trilhando novos caminhos para a aprendizagem. Com relação à participação e ao envolvimento nos fóruns de discussões, Oliveira (2005) salienta que a participação, no espaço criado pelo fórum, pede preparo, geralmente provido por leituras adequadas, pesquisas, resgates ao *background* próprio de cada participante, entre outras formas de busca. Trata-se de organizar o pensamento, enriquecendo-o com pertinentes referências, permitindo o uso do espaço de discussões e de reflexões, proporcionado pelo fórum, para gerar colaborações e agregar ideias.

Partindo deste problema, esta pesquisa propõe um sistema de recomendação de OA, baseado em postagens extraídas dos fóruns de um AVA. Este sistema de recomendação foi denominado de RECOAWS. A modelagem do sistema RECOAWS será mostrada na seção a seguir.

5.1 Modelagem do Sistema RECOAWS

O diagrama de caso de uso representa a interação entre o sistema e os *web services* externos que foram utilizados. A tarefa principal desta tese é recomendar OA diretamente dentro dos fóruns de um AVA. A Figura 4 apresenta o relacionamento entre os casos de uso necessários para atingir este objetivo.

Figura 4 - Diagrama de Casos de Uso



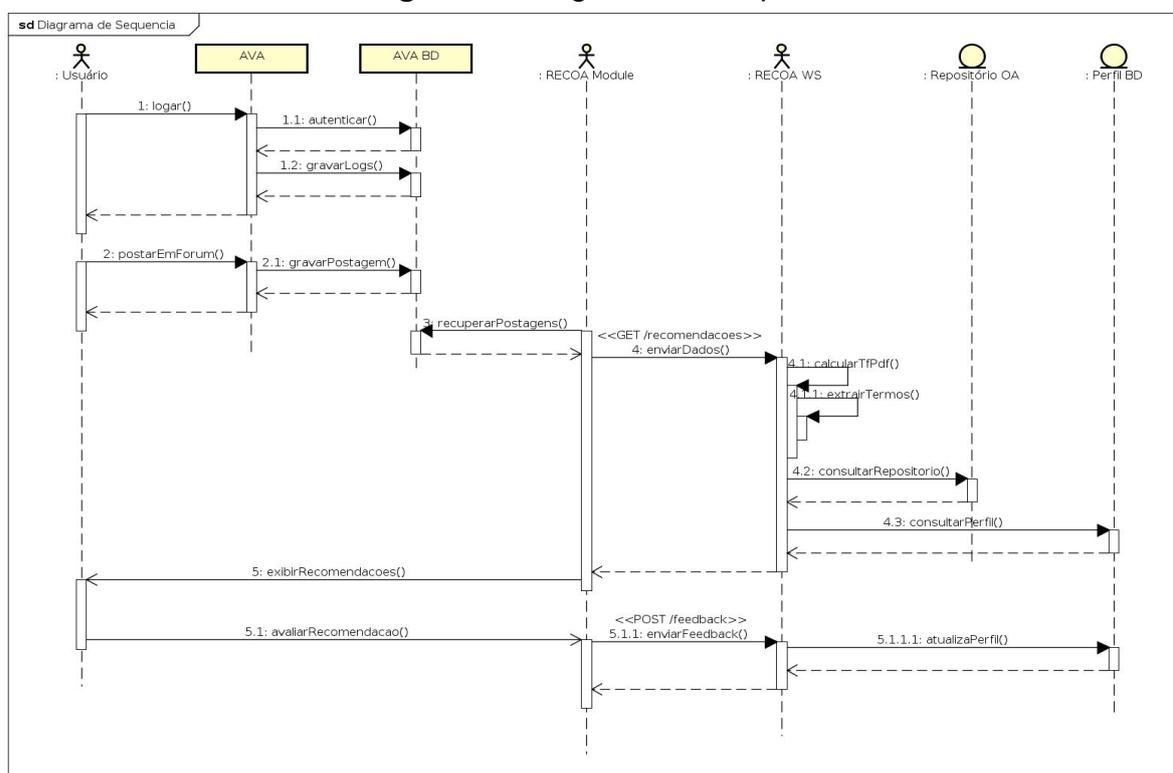
Fonte: o autor (2015)

Neste projeto, tem-se definidos os papéis de dois serviços, o de recomendação e o de avaliação da recomendação. No primeiro, o módulo específico do AVA realizou a chamada do serviço, passando, como parâmetro, o conteúdo das postagens dos fóruns e recebeu uma lista de recomendações de OAs. Já o segundo recebeu o *feedback* do usuário, que informou uma nota de 1 a 5, avaliando a pertinência do objeto que lhe foi recomendado.

Ainda no serviço de recomendação, é importante destacar que este agrega internamente as funções de pré-processamento de texto. Nesta etapa, o analisador de conteúdo atua no texto passado para o serviço, removendo as *stopwords*, ou seja,

tirando palavras que não acrescentam representatividade ao documento. Em seguida, é realizada a etapa de extração de *hot-topics*, que, como já foi dito, utiliza a técnica TF*PDF para seleção das palavras-chave, só então, depois destas etapas, é realizada a consulta ao repositório de OA para busca das recomendações. No protótipo proposto, foi utilizado a API do motor de busca *Bing* que oferece 5000 consultas gratuitas por mês, e permite que os *links* sejam retornados tanto em XML quanto em *JSON* (FONSECA; FARIAS; SILVA, 2014). A Figura 5 apresenta o diagrama de sequência e as interações entre os serviços de recomendação e avaliação do OA recomendados.

Figura 5 - Diagrama de sequência



Fonte: o autor (2015)

Considerando a característica de interoperabilidade, proposta para o sistema, o modelo arquitetural, utilizado foi o REST para o desenvolvimento dos serviços *Web*. Tal modelo foi proposto por Fielding (2000), de tal forma que ignora os detalhes de sintaxe de protocolo e de implementação, focando as regras de negócio, na comunicação entre os componentes e suas restrições. Suas características básicas são a sua essência cliente-servidor, a comunicação sem estado, utilização

de *cache*, interface uniforme e a utilização de camadas, além de ser independente de plataforma e de linguagem.

Conforme dito por Richardson (2008), para o desenvolvimento de web services REST, deve-se utilizar o protocolo HTTP, bem como seus verbos e mensagens de *status* para o tratamento das chamadas e retorno dos nossos serviços. Sendo assim, utilizou-se os *end-points* GET /recomendacao, para o serviço de recomendações, e POST/avaliacao, para o serviço de *feedback*. As especificações destes *web services* podem ser observadas no quadro 1. É importante destacar que foi utilizado o *JavaScript Object Notation* (JSON) para estruturação dos dados, será retornado um código de status que indicará o sucesso ou não da ação.

Quadro 1 - Especificações deste Web Services

Serviço	Entrada	Saída	Mensagens
GET /recomendacao	{ "texto": "..." }	[{ "titulo": "Reco 01", "descricao": "Desc Recome", "link": "http://www.site (...)", "tipo": "imagem" }, ...]	200: Sucesso 404: Nenhuma recomendação encontrada
POST /avaliacao	{ "ava": "...", "usuario": "...", "link": "...", "nota": 5 }		201: Sucesso

Fonte: o autor (2016)

Para se trabalhar de forma satisfatória com aspectos de segurança, e para garantir a confidencialidade dos dados, é imprescindível que os serviços estejam hospedados em um servidor com suporte a *TLS (Transport Layer Security)*, que garante um mecanismo de criptografia das mensagens trocadas com os clientes. Caso haja a necessidade da restrição dos serviços, somente para clientes

previamente cadastrados, é possível implementar um mecanismo baseado em *tokens*, em que cada cliente utiliza uma chave própria, através do cabeçalho *Authorization*, para acessar os *web services*.

5.2 Arquitetura do Sistema RECOAWS

A pesquisa apresenta uma arquitetura do sistema de recomendação, baseado em postagens extraídas do fórum de um ambiente virtual de aprendizagem. Um caminho possível para desenvolver a arquitetura foi fazer uso de uma ferramenta de mineração de texto, com objetivo de verificar se as palavras-chave extraídas pelo sistema RECOAWS, estão de acordo com as discussões realizadas no fórum. Uma vez identificadas, as palavras-chave são submetidas a um motor de busca de um repositório que retornará os OAs, que irão oferecer novos conteúdos ao usuário, para se aprofundar no contexto da discussão. Esses OAs foram avaliados pelos usuários e alimentarão um perfil que, no futuro, fará novas sugestões, baseadas em gostos anteriores dos usuários, fazendo uso agora de um mecanismo de filtragem colaborativa. A partir daí, propõe-se um sistema que coleta duas filas de prioridades: uma com prioridade zero, para itens de recomendados por filtragem colaborativa e outra para as recomendações oriundas de filtragem, baseada em conteúdo. O quadro 2 mostra a lista de prioridades (SILVA *et al*, 2015).

Quadro 2 - Lista de Prioridades

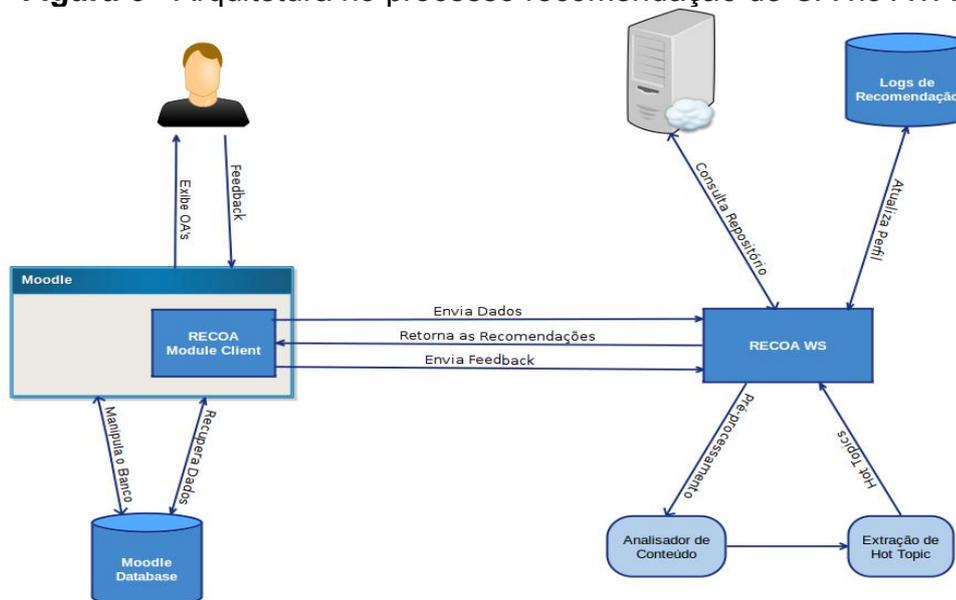
Prioridades	Técnicas
0	Filtragem Colaborativa
1	Filtragem Baseada em Conteúdo

Fonte: o autor (2015)

A pesquisa apresenta um sistema RECOAWS, que disponibiliza OAs aos usuários de um AVA, através dos fóruns de acesso ao sistema. Esse sistema visa recomendar também OAs de acordo com discussões no fórum de um AVA; por isso, foi concebido usando a tecnologia de recuperação de informação, mineração de textos e serviços *Web*. A Figura 6 apresenta a arquitetura proposta para este estudo e composta por dois subsistemas, um AVA, conforme Apêndice F, com um módulo específico para ele, e um serviço de recomendação genérico, chamado serviço WEB

RECOAWS, conforme Apêndice G. Para fins de experimento, um AVA escolhido foi o *Moodle*, sendo desenvolvido um módulo específico para a exibição das recomendações.

Figura 6 - Arquitetura no processo recomendação de OA no AVA



Fonte: o autor (2014)

Com a utilização de uma arquitetura orientada a serviços, garante-se que o acesso às recomendações possa se dar por qualquer AVA, bastando que estes façam chamadas aos *web services*, conforme apêndice G, de acordo com a interface acordada. A seguir, os principais componentes desse sistema serão detalhados.

5.3 Serviço de Recomendação

Para garantir que o sistema de recomendação esteja acessível a qualquer AVA, foi utilizada a tecnologia dos serviços *Web*. Essa tecnologia foi criada para garantir a interoperabilidade entre sistemas e é um meio simples, de baixo custo e multiplataforma para compartilhamento de funções e dados entre aplicações, conforme Sommerville(2011) e Zimmermann (2012).

Esse serviço recebeu os *logs* dos usuários e as atividades desenvolvidas em cada tópico dos fóruns de um AVA para extrair informações dos resultados, cujo

conteúdo foi representado, no sistema, através da classe documento, composta pelos atributos: nome, quantidade de palavras e uma lista de termos relevantes.

Para representar cada termo contido no documento, foi criada a classe termos. Esta, por sua vez, contém os atributos: termo e frequência do termo *tf* (SALTON; BUCKLEY 1988) e a frequência ponderada do termo *pdf* (BUN; ISHIZUKA 2001), em que o peso do termo é calculado através da multiplicação da frequência pela sua frequência ponderada ($TF * PDF$) e o número de ocorrências do termo. O processo do serviço de recomendação, como um todo, pode ser dividido nas seguintes etapas:

Analizador de Conteúdo: consiste no pré-processamento do texto do fórum. Como o conteúdo do fórum se encontra em linguagem natural, antes de se aplicar qualquer método estatístico de cálculo de relevância de termos, é feita uma triagem para separar os termos não representativos (*stopwords*), como advérbios, adjetivos, artigos e preposições, ou seja, classes que, em geral, não acrescentam representatividade ao documento e, muitas vezes, estão presentes somente para conectar frases.

Extração de *Hot Topics*: nesta etapa, será realizada a ponderação de termos (*Term Weighting* - *TF*) e é importante para determinar a relevância de uma palavra. Para se calcular a frequência do termo (*TF*), foi adotada uma abordagem diferente. Como não é feito o cálculo de similaridade entre os documentos a coleção de documentos, isto é, os diferentes tópicos do fórum de discussão são agrupados e os pesos *TFs* são calculados para todos os termos em relação a um único documento. Uma vez feito isso, o cálculo do peso *PDF* é realizado, levando em consideração a frequência do termo em relação ao conjunto de documentos.

No algoritmo $TF * PDF$, os termos dos documentos são ponderados e é realizada a ordenação deles para identificar quais são os principais termos referenciados nos fóruns de discussão. Portanto, feita a identificação dos principais tópicos discutidos nos fóruns, o sistema seleciona os cinco primeiros termos e os submetem ao motor de busca. Nesta pesquisa, foi utilizado-se a API do motor de busca *Bing*, que oferece cinco mil consultas gratuitas por mês e permite que os links sejam retornados tanto em XML quanto em JSON.

A *API Bing Search* permite usar os dados que foram recolhidos pelo motor de busca *Bing* em seu aplicativo de desktop, web, ou outro componente de software. Pode-se usar o *Bing Search API* com qualquer componente de software capaz de

emitir uma solicitação do *Hypertext Transfer Protocol Secure (HTTPS)*, podendo obter resultados de pesquisa em seis categorias, descritas no quadro 3 a seguir.

Quadro 3 - Categorias dos resultados da API Bing Search

Categoria	Descrição
<i>Web General</i>	Páginas web que estão relacionadas a <i>query</i> .
<i>Image</i>	Imagens e ícones que estão relacionadas a <i>query</i> .
<i>Video</i>	Vídeos que estão relacionadas a <i>query</i> .
<i>News</i>	Notícias de artigos que estão relacionadas a <i>query</i> .
<i>Related Search</i>	Pesquisa, expressões e seus correspondentes <i>Bing Uniform Resource Identificadores (URIs)</i> que estão relacionadas a <i>query</i> .
<i>Spelling Suggestions</i>	Sugestões de ortografia para quaisquer palavras com erros ortográficos na sua consulta.

Fonte: o autor (2015)

Cada item recuperado contém informações sobre o recurso relevante, tal como o seu título, descrição e URI. O número e tipo de campos, para cada resultado, variam por categoria. No caso de resultados de imagem e vídeo, cada um deles contém informações em miniatura que podem ser usadas para exibir as miniaturas na sua aplicação.

Além de usar as categorias de personalizar os resultados, pode-se dizer ao *Bing Search API* para filtrá-los em um conjunto. Por exemplo, podem-se limitar os resultados da pesquisa para um tipo de documento particular (por exemplo, Microsoft Word ou PDF), ou de uma região geográfica específica. É possível, ainda, especificar que as imagens devem ser de um determinado tamanho ou de uma determinada proporção, bem como solicitar apenas vídeos de certo período. Na categoria Notícias, elas podem ser filtradas por entretenimento, política ou esportes.

A *API Bing Search* permite usar os parâmetros, no quadro a seguir, para especificar o número de resultados que se deseja obter; por onde começar no conjunto de resultados (ou seja, a paginação); e o formato resultado. O quadro 4 apresenta os parâmetros da *API* do *Bing*.

Quadro 4 - Parâmetros da API Bing Search

\$top	Especifica o número de resultados retornados. Por padrão são retornados 50 para WEB, Imagens e Vídeos, e 15 para notícias.
\$skip	Especifica o ponto de partida dos resultados. O padrão é começar do zero.
\$format	Especifica o formato da resposta. As opções atuais são Atom (para XML) ou JSON. Padrão: Atom

Fonte: o autor (2015)

A API de *Bing Search* permite que os resultados da busca sejam recebidos em dois diferentes formatos: JSON e XML. O quadro 5 o mostra formato retornado pela API *Bing Search* e quadro 6 apresenta os links retornados contendo os seguintes elementos: *id*, *title*, *description*, *displayurl* e *url*.

Quadro 5 - Formato retornados pela API Bing Search

Formato	Descrição
<i>JavaScript Object Notation</i> (JSON)	Um formato de intercâmbio de dados leve. JSON é apoiado por muitas linguagens de programação e ambientes, como o Microsoft Visual Basic, Java, Perl, PHP, Ruby, C++ e C#.
<i>Extensible Markup Language</i> (XML)	Uma linguagem de marcação amplamente suportada que define um conjunto de regras para a codificação de documentos em um formato que seja legível e legível por máquina.

Fonte: o autor (2015)

Quadro 6 - Elementos que compõem os arquivos retornados pela API Bing Search

ID	Identificados
Title	Texto especificado na tag HTML <title> da página retornada
Description	Texto de descrição da página
DisplayUrl	URL Web que será apresentada ao usuário.
Url	URL completa

Fonte: o autor (2015)

Após feita a consulta, os *links* (no caso desta pesquisa, somente os cinco primeiros) são apresentados ao usuário como sugestão para leitura e aprofundamento no estudo. Um mecanismo de avaliação é proposto com a finalidade de verificar a relevância do *link* para o fórum o qual foi proposto.

O usuário poderá avaliar de acordo com o enquadramento do documento em meio à discussão proposta no fórum; caso o usuário goste de um *link* recomendado, é feito um processo de indexação do campo descrição, que apresenta um resumo do documento. Nesse contexto, o sistema armazena os termos de maior peso no

perfil do usuário, para que, em um momento posterior, possam ser feitas sugestões de outras matérias, com base em seu interesse.

A Recomendação trabalha em duas etapas, que são: na primeira etapa do serviço, utiliza a lista de *Hot Topics* e faz a consulta aos repositórios. Este, por sua vez, recomenda, utilizando a filtragem, baseada em conteúdo, que é oferecida a todos os participantes de um fórum.

Na segunda etapa, após o recebimento da recomendação, utiliza-se uma avaliação que será solicitada por parte do sistema. Sendo assim, quando alguns itens forem avaliados, será possível a oferta de OA por intermédio da filtragem colaborativa que, nesta proposta, terá maior prioridade em relação à filtragem baseada em conteúdo (FARIAS; FONSECA; SILVA, 2014).

5.4 Integração do Sistema RECOAWS com AVA

Para a integração do AVA, com o nosso serviço, foi desenvolvido um módulo específico, conforme apêndice F. Este módulo foi responsável por fazer as chamadas do serviço *web* RECOAWS, conforme Apêndice G, mostrando os OAs recomendados de dentro dos fóruns do ambiente. Além disso, este módulo tem o papel de receber o *feedback* do usuário e enviar para o serviço, a fim de que seja atualizado o perfil de recomendação do usuário.

Para se realizar a integração do AVA, com o sistema RECOAWS, foi utilizado o *Block Moodle*, que é um *plugin* desenvolvido para uso na página principal do site e páginas de cursos, sendo muito interessante por sua praticidade. Cada instância de um tipo de bloco pode ter suas próprias configurações e atribuições de papel. Um bloco é criado, basicamente, a partir de quatro arquivos (MOODLE ORG, 2016):

1) block_newblock.php => arquivo responsável pela definição da classe do bloco e pelo seu gerenciamento e geração da página de conteúdo a ser apresentado ao usuário. O termo 'newblock' deve ser substituído pelo nome do bloco. Exemplo: **block_recoa.php**.

2) db/acecces.php => arquivo que define as capacidades criadas pelo bloco. Ao menos duas capacidades são necessárias em um bloco: "myaddinstance"

e “addinstance”.

3) lang/en/block_recoa.php e lang/pt_br/block_recoa.php => arquivos que guardam os idiomas do bloco. Para cada idioma adicionado ao *plugin*, um diretório e o arquivo **block_recoa.php**, para este idioma, devem ser adicionados no diretório **lang**.

4) version.php => este arquivo guarda a versão do *plugin*, a versão mínima exigida do *moodle*, quais *plugins* são necessários para o funcionamento do *plugin*, dentre outras configurações.

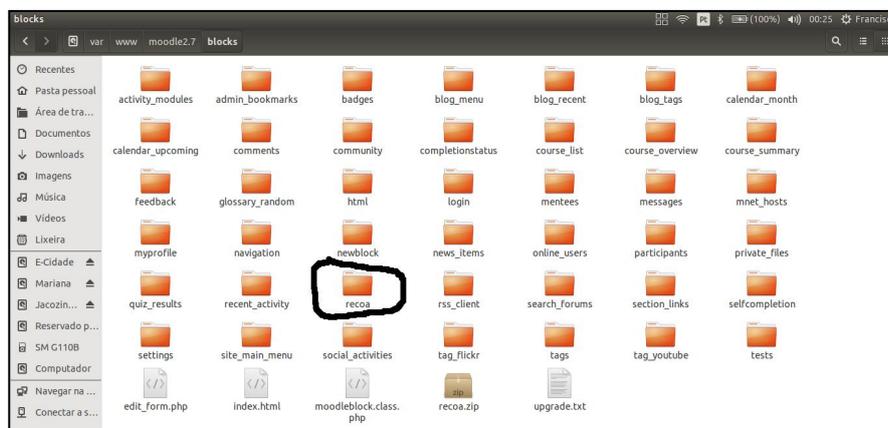
Outro arquivo importante em um bloco é o **lib.php**, pois, através dele, pode-se fazer a formatação do conteúdo de saída do bloco. Normalmente, há também o arquivo **styles.css**, que faz a estilização do conteúdo html.

Algumas variáveis especiais são utilizadas em um bloco, conforme sua necessidade; são variáveis globais, também chamadas objetos, definidas pela *API* do *moodle*. No bloco desenvolvido neste trabalho, foram utilizadas as seguintes variáveis globais: **\$DB**, **\$COURSE** e **\$PAGE**. O objeto **\$DB** é uma instância da classe *moodle_database.php*, responsável pela manipulação do banco de dados. O objeto **\$COURSE** fornece recursos para interação com um curso; o objeto **\$PAGE** disponibiliza o acesso a informações referentes à página.

No bloco deste trabalho, foram utilizados alguns recursos do objeto **\$DB**, tais como os métodos **\$DB->get_records_sql** e **\$DB->get_records**. Em relação ao objeto **\$COURSE**, utilizou-se tal objeto para recuperar o **id do curso**, a fim de que ele mesmo fosse passado via parâmetro na *query* a ser executada pelo método **\$DB->get_records_sql**. Por fim, o objeto **\$PAGE** é utilizado no arquivo **lib.php** em conjunto com o *javascript*, definido a partir do arquivo **module.js**; através do objeto **\$PAGE**, é montada uma estrutura html que compara o conteúdo de saída do bloco, que será mostrado para o usuário.

A figura 7 apresenta a instalação do bloco recoa. A instalação do bloco recoa, no *moodle*, consiste em algumas etapas simples que, normalmente, são executadas por um usuário administrador do ambiente. Para a instalação, é necessário copiar o diretório **recoa** para dentro do diretório **blocks** do *Moodle*.

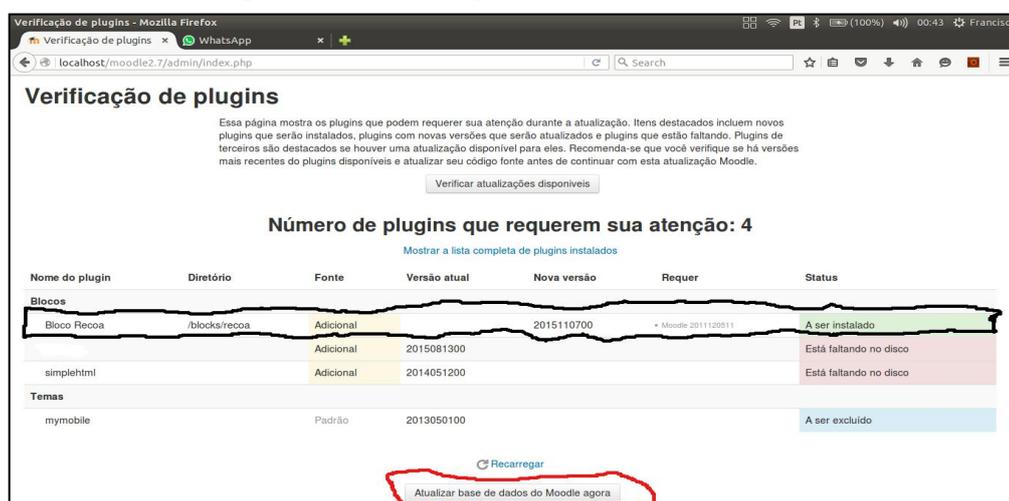
Figura 7 - Bloco recoa copiado para o diretório *blocks* do Moodle



Fonte: o autor (2015)

Após copiar o bloco, para o diretório indicado no *moodle*, e quando atualizar a página do ambiente, ou acessar a página inicial, será solicitada a realização de atualizações no ambiente, referentes ao novo bloco adicionado, conforme a figura 8. Deve-se clicar no botão “Atualizar base de dados do Moodle agora” para que as alterações sejam realizadas.

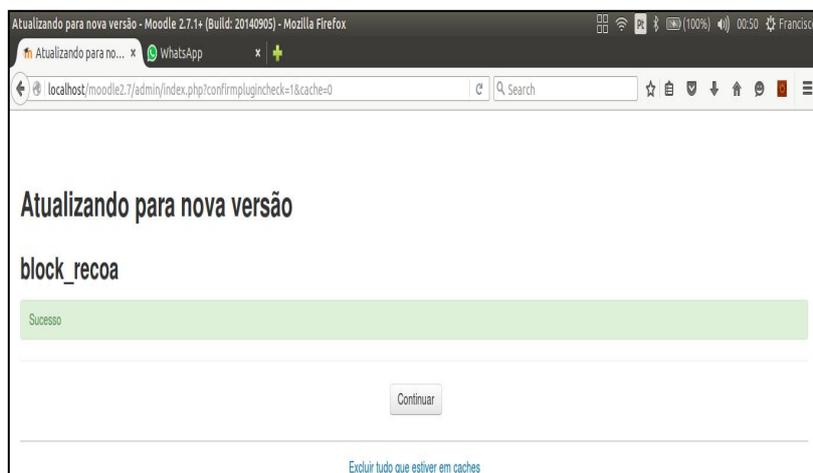
Figura 8 - Verificação de *plugins* do Moodle



Fonte: o autor (2015)

Finalizada a atualização, uma mensagem de sucesso é emitida pelo *moodle* e o usuário deve clicar no botão “Continuar” para retornar à navegação no ambiente, conforme a figura 9.

Figura 9 - Bloco recoa instalado

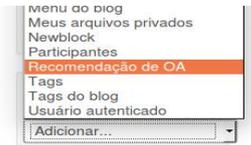


Fonte: o autor (2015)

Após a instalação do bloco, é possível realizar a habilitação do bloco no *moodle*; um bloco, no *moodle*, é um *plugin* que pode ser adicionado à páginas específicas no ambiente e, no caso do bloco recoa, é adicionado em uma página de discussão de um fórum pertencente a alguma disciplina de um curso.

Para adicionar um bloco na página, basicamente, há três etapas a serem seguidas. O quadro 7 apresenta as etapas para adicionar o bloco do *moodle* no curso.

Quadro 7 - Etapas para adicionar o bloco do *moodle* no curso

Etapas	Procedimentos
1. Clicar na caixa de seleção para adicionar um bloco	
2. Selecionar a opção Recomendação de AO	
3. Visualizar o bloco na página	

Fonte: o autor (2015)

5.5 Usuários

No sistema, o usuário faz sua avaliação, para o serviço de recomendação, por meio de módulo específico em seu AVA. Nessa avaliação, o usuário informa se o OA foi útil ou não para a discussão no fórum. Esta informação é incorporada, no seu perfil, para futuras recomendações.

Para a implementação do sistema RECOAWS, foi utilizada, a linguagem de programação PHP, tanto para o desenvolvimento dos *web services* REST quanto para o desenvolvimento do módulo para o *Moodle*. No desenvolvimento dos serviços, também foi utilizado o *microframework Slim* e, para a exibição das recomendações, a biblioteca *JQuery*. A figura 10 mostra a interface do sistema RECOAWS.

Figura 10 - Interface do RECOAWS



Fonte: o autor (2015)

O objetivo desta pesquisa foi verificar a eficiência na extração automática de palavras-chave dos tópicos de discussão do AVA e averiguar se o uso destas proporcionou a recuperação de *links* relevantes ao tema em discussão.

5.6 Conclusão do Capítulo

Neste capítulo, foi apresentada a prototipagem do sistema RECOAWS para a recomendação de OA dentro dos fóruns de um AVA. O modelo arquitetural utilizado, no desenvolvimento dos serviços *Web*, foi o *REST*. O sistema RECOAWS faz uso de uma ferramenta de mineração de texto para a identificação das palavras-chave que representam os tópicos de um fórum. O algoritmo utilizado para a extração de palavras-chave foi o TF*PDF. Neste algoritmo, os termos dos documentos são ponderados, ordenados, possibilitando a identificação dos principais termos referenciados nos fóruns de discussão.

O algoritmo TF*PDF permite a busca por palavras-chave, através da *API* do motor de busca *Bing*. Neste contexto, os OAs são disponibilizados aos usuários, possibilitando a recuperação de links pela *API* do motor de busca do *Bing*. Esses links são retornados tanto em XML quanto em JSON. Após a disponibilização dos OAs pelo sistema RECOAWS, o usuário faz a avaliação verifica a relevância dos objetos para o fórum proposto. Para a integração do AVA, com o sistema RECOAWS foi utilizado o *plugin* do *Block Moodle*. No próximo capítulo, serão apresentados, os resultados e as discussões do sistema RECOAWS.

6 RESULTADOS E DISCUSSÕES

Neste capítulo, são apresentados os resultados obtidos com a utilização do sistema RECOAWS no curso de Engenharia de Computação da Universidade Estadual do Maranhão. As estratégias pedagógicas desenvolvidas, na pesquisa, foram organizadas com os professores das disciplinas presenciais, os quais definiram as perguntas a serem abordadas em cada fórum. Nesse contexto, o sistema RECOAWS recomenda OA, levando-se em consideração as postagens feitas de dentro dos fóruns de um AVA.

Ao apropriar-se desses OAs, espera-se que os alunos e professores sejam capazes de utilizá-los e reutilizá-los de forma crítica; dessa forma, apresentam-se as estratégias que possibilitem a apropriação e a utilização desses recursos, impulsionando um fazer pedagógico capaz de envolver os educandos, incentivando-os a serem autores e produtores do seu próprio conhecimento.

Para o desenvolvimento do sistema RECOAWS, foram realizados alguns experimentos. Nesta etapa, são apresentados e descritos passos, como: a seleção das amostras de estudo, a preparação da sala virtual, a análise de dados coletados e os resultados encontrados, análises e discussões e contribuições da tese. As referidas etapas supracitadas são descritas a seguir.

a) **Definição da amostra:** utilizou-se as amostras das turmas do primeiro semestre de 2016, do curso de Engenharia de Computação, da Universidade Estadual do Maranhão (UEMA), mas em períodos diferentes nas disciplinas Modelagem de Banco de Dados e Interface do Usuário e Sistemas Computacionais da Graduação. A amostra foi realizada com 25 (vinte e cinco) alunos da graduação e 5 (cinco) professores.

b) **Preparação da sala virtual:** a preparação da sala virtual do sistema foi realizada por cinco professores, inseridos nas disciplinas ofertadas, com o objetivo de avaliar sistema de recomendação, verificar-se as palavras-chave estavam de acordo com as discussões nos fóruns. Os professores postaram, no fórum do AVA, o tema a ser debatido com os alunos. O quadro 8 apresenta os tópicos a serem discutidos no fórum do AVA.

Quadro 8 - Tópicos a serem discutidos nos fóruns do AVA

Disciplinas	Tópicos no fórum do AVA
Interface entre usuário e sistemas Computacionais	A interface para um sistema interativo são todas as peças do sistema com as quais as pessoas têm contato, física, perceptiva ou conceitualmente. Utilizando os conceitos de engenharia cognitiva, engenharia de usabilidade, engenharia ergonômica e engenharia semiótica, descreva um produto ou sistemas interativos que você utiliza no seu dia a dia.
Modelagem de Banco de Dados	Os bancos de dados são coleções de informações que se relacionam para criar um significado dentro de um contexto computacional. Operados por Sistemas Gerenciadores de Banco de Dados (SGBDs) como Oracle, MySQL, SQL e PostgreSQL, os BDs são a base para a maioria das aplicações utilizadas pelas empresas como os CRM, ERP, SCM, Service Desk. Descreva quais os principais SGBDs disponíveis, destacando seus pontos positivos e negativos.

Fonte: o autor (2016)

O quadro 8 apresentou os temas abordados pelo professores para discussões nos fóruns do AVA. O sistema RECOAWS recomendou OA de acordo com as postagens nos fóruns. Para a coleta e análise dos dados, foram utilizados os seguintes procedimentos:

a) **Coleta de dados:** para coleta de dados foram selecionadas as disciplinas de Graduação Interface do Usuário e Sistemas Computacionais e Modelagem de Banco de Dados, com seus respectivos alunos e professores, no total foram selecionados 25 (vinte e cinco) alunos e 5 (cinco) professores.

b) **Análise dos dados coletados:** A análise dos dados utilizados, nesta pesquisa, foi baseada na estatística não paramétrica U de *Man-Winter*, conforme anexo B, que apresenta todas as tabelas realizadas pelo software livre *BioEstat*², a fim de entender e interpretar comportamentos e atitudes dos alunos e professores quanto a utilização do sistema e a avaliação dos OAs recomendados.

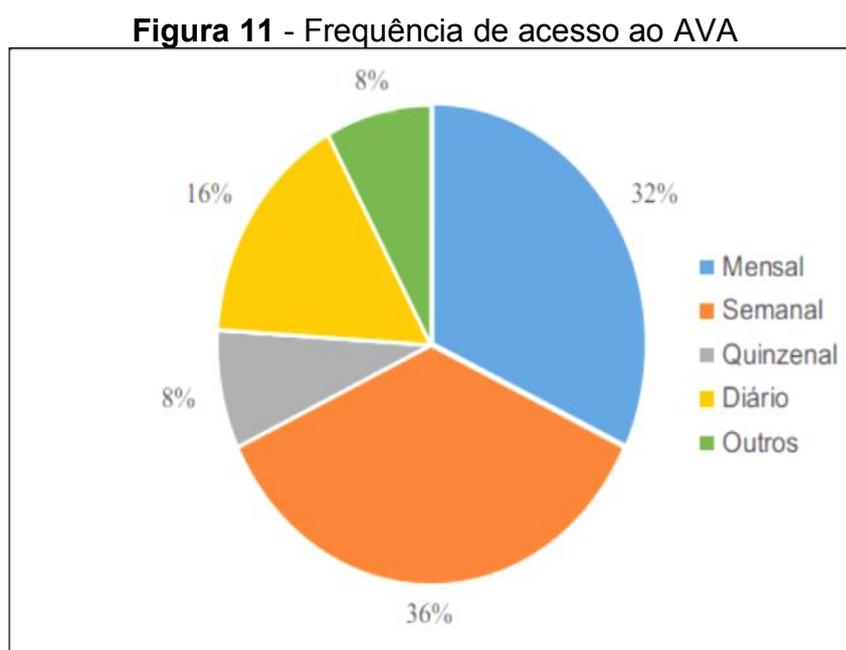
² Disponível em: <http://www.mamiraua.org.br/pt-br/downloads/programas/bioestat-versao-53/> acesso em: 20 maio 2016.

6.1 Resultados

Com o intuito de subsidiar as análises e discussões dos resultados, metodologicamente, utilizou-se a *Google Form* para a coleta dados, bem como manipulá-las e, assim, auxiliar na formulação das análises, discussões e conclusões baseadas nas tabelas e gráficos gerados. Para identificar os resultados das duas amostras (professores e alunos), foi solicitado aos professores, conforme apêndice A, e aos alunos, conforme apêndice B, que respondessem um questionário. Os dados obtidos foram determinantes para a elaboração dos percentuais para cada questão realizada no questionário.

6.1.1 Alunos

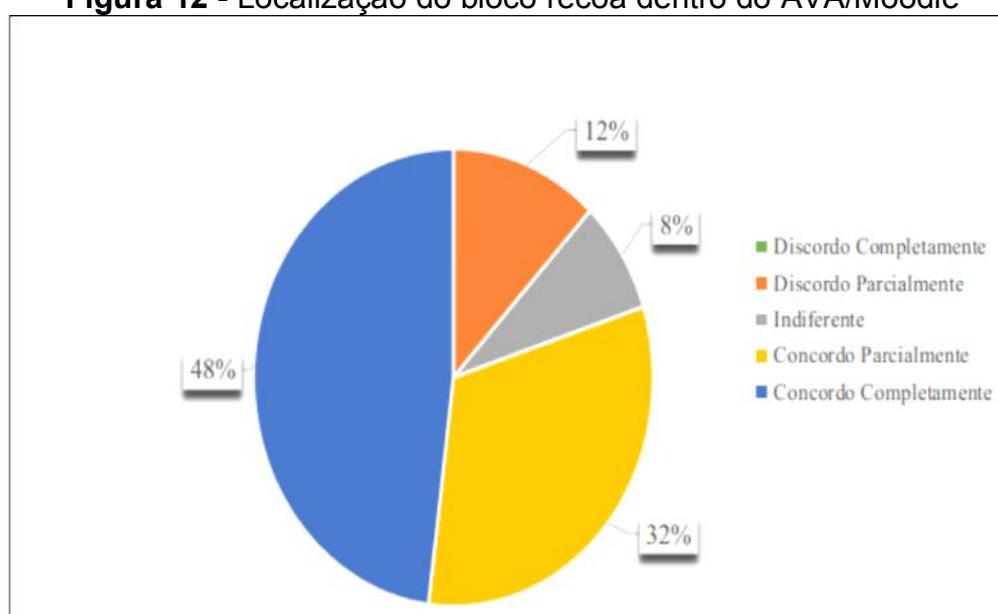
Neste momento, começa-se a questionar os alunos para avaliar a frequência de acesso. Assim, foi elaborada a seguinte pergunta: “Qual a sua frequência de acesso ao AVA?”. Nessa pergunta, tem-se a pretensão de verificar a tendência da variável “frequência de acesso” por parte dos alunos. A figura 11 mostra os percentuais de frequência de acesso ao AVA.



Fonte: o autor (2016)

A figura 11 apresentou o seguinte resultado: 36% dos alunos acessam o AVA semanalmente; 32%, mensalmente; 16%, diariamente; 8%, quinzenalmente; 8% não respondem com que frequência acessam o AVA. Depois de analisar o resultado da frequência de acesso, os alunos responderam a seguinte pergunta: “Foi fácil localizar o bloco recoa dentro do Moodle, através dos fóruns?”. Passou-se, então, a investigar sobre variáveis de “Localização do bloco recoa no Moodle”. A Figura 12 apresentou os percentuais de localização do bloco recoa no fórum.

Figura 12 - Localização do bloco recoa dentro do AVA/Moodle

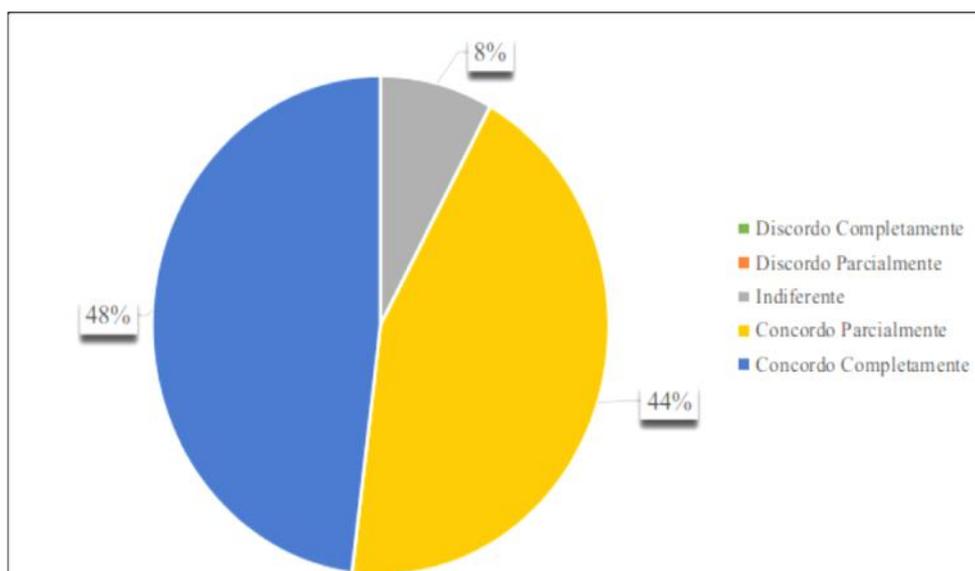


Fonte: o autor (2016)

A figura 12 apresentou o seguinte resultado: 48% dos alunos não tiveram nenhuma dificuldade em localizar o bloco recoa dentro do fórum; 32% tiveram algumas dificuldades, mas conseguiram localizar o sistema dentro do fórum e 20% dos alunos tiveram ajuda externa para localizar o sistema. Resumindo 80% dos alunos, que responderam a pesquisa, disseram que foi fácil localizar o bloco recoa dentro do fórum do Moodle. Após investigar a localização do bloco recoa, os alunos responderam à seguinte pergunta: “As palavras-chave extraídas pelo sistema possuem representatividade em relação aos textos dos fóruns?”. Neste momento, começou-se então a investigar sobre variáveis de “Palavras-chave extraídas dos fóruns”. A figura 13 ilustra o resultado que representa as opiniões dos alunos em

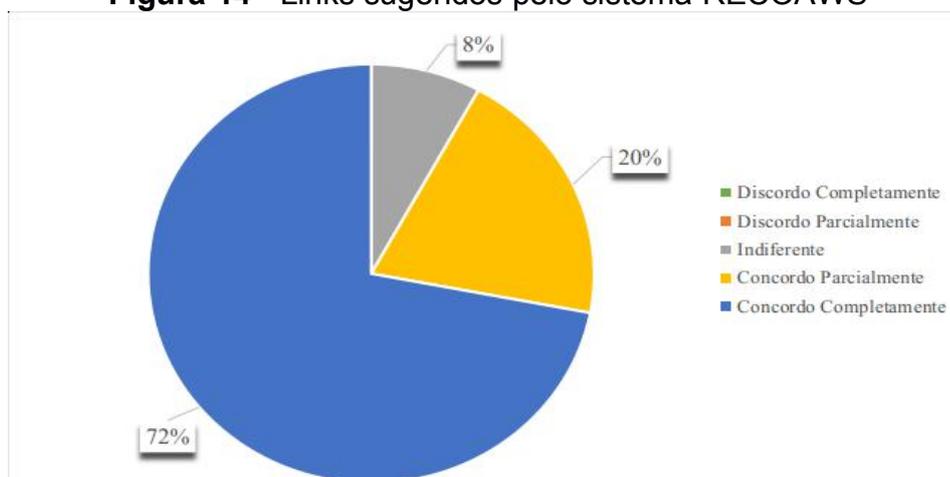
relação as palavras-chave extraídas pelo algoritmo TF*PDF, através do sistema RECOAWS.

Figura 13 – Palavras-chave extraídas pelo sistema RECOAWS, através do Algoritmo TF*PDF



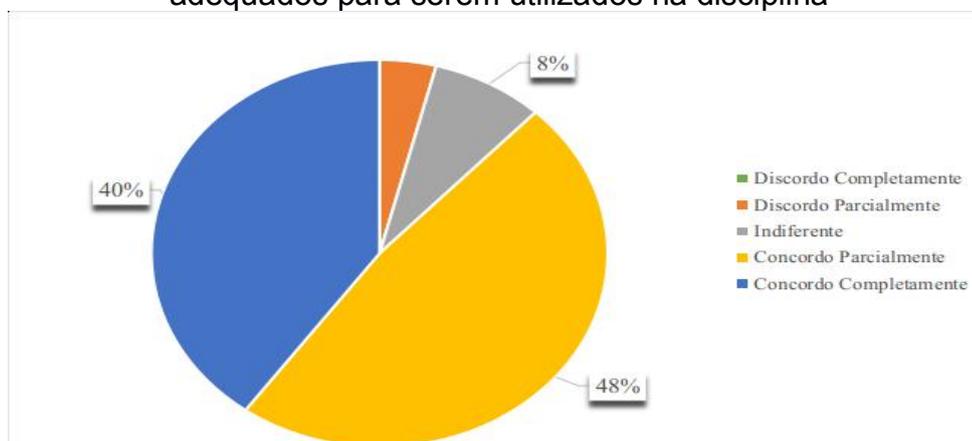
Fonte: o autor (2016)

O resultado ilustrado pela figura 13 apresenta o seguinte resultado: 92% dos alunos responderam que as palavras-chave, extraídas pelo sistema RECOAWS, por meio do algoritmo TF*PDF, estavam de acordo com as discussões dentro dos fóruns; e 8% dos alunos disseram que as palavras-chave não estavam de acordo as discussões dentro dos fóruns. Sendo assim, observou-se que, em sua maioria, as palavras-chave estavam de acordo com as postagens enviadas para os fóruns do Moodle. Após a análise da variável “palavras-chave extraídas dos fóruns”, os alunos responderam a seguinte pergunta: “Os links sugeridos pelo sistema RECOAWS estão de acordo com os temas em discussões dentro dos fóruns?”, começou-se, então, a investigar a variável sobre os “links sugeridos pelo sistema RECOAWS aos alunos nos fóruns”. Nesta variável, tem-se a pretensão de analisar se os *links* estavam de acordo com o tema em discussão no fórum do Moodle. A figura 14 apresentou os seguintes resultados em relação aos links sugeridos para os alunos de acordo com tema em discussão nos fóruns do AVA.

Figura 14 - Links sugeridos pelo sistema RECOAWS

Fonte: o autor (2016)

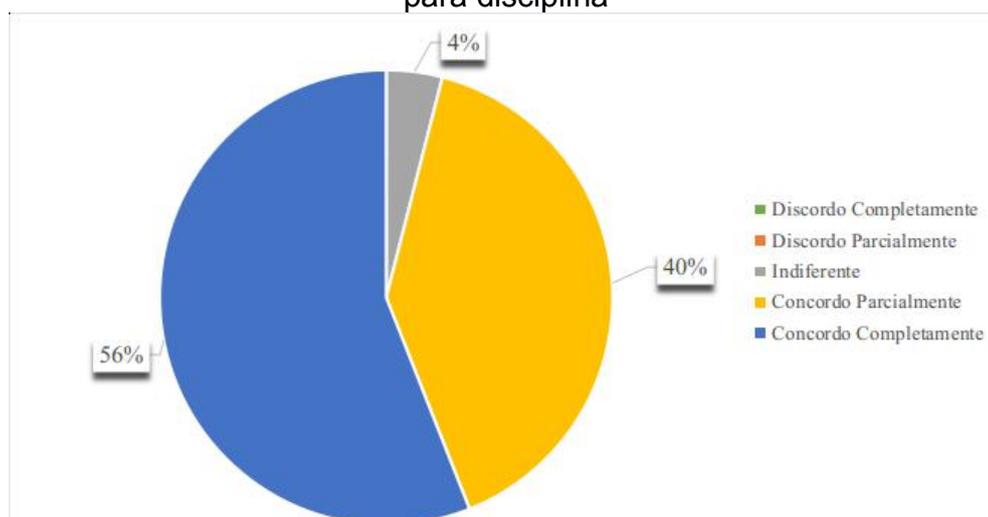
O resultado ilustrado, pela figura 14, apresentou o seguinte resultado: 72% dos alunos responderam que os *links* sugeridos pelo sistema de recomendação estão de acordo com tema em discussão dentro do fórum; 20% dos alunos concordam parcialmente que os *links* sugeridos estão de acordo com as discussões dentro do fórum do Moodle; e 8% dos alunos responderam que os *links* foram indiferentes ao tema da discussão. Depois, os alunos responderam: “Os objetos de aprendizagem disponibilizados pelo sistema RECOAWS foram adequados para serem utilizados na disciplina?”. A partir daí começou-se investigar a variável quanto à “disponibilização de OA”. A figura 15 apresentou a disponibilização dos OAs nas disciplinas ofertadas no AVA/Moodle.

Figura 15 - OAs recomendado pelo sistema RECOAWS que se mostraram adequados para serem utilizados na disciplina

Fonte: o autor (2016)

A figura 15 ilustrou o seguinte resultado: 48% dos alunos concordaram que os OAs recomendados estavam de acordo com as discussões dentro do fórum do AVA; 40% dos alunos concordaram parcialmente com as recomendações dos OA, apresentados pelo sistema; 8% dos alunos ficaram indiferentes as recomendações do OA pelo sistema; e 4% dos alunos discordaram parcialmente das recomendações do OA. Depois de investigar a variável “disponibilização dos OAs”, os alunos responderam a seguinte pergunta: “A integração do sistema RECOAWS ao AVA/Moodle facilitou a busca de OA para estudo do conteúdo da disciplina?”. Assim, começou-se a pesquisar sobre a variável “integração do sistema de RECOAWS com o AVA/Moodle”. A figura 16 apresentou a integração do sistema com o AVA/Moodle.

Figura 16 - Integração do sistema RECOAWS com AVA facilitou a busca de OA para disciplina

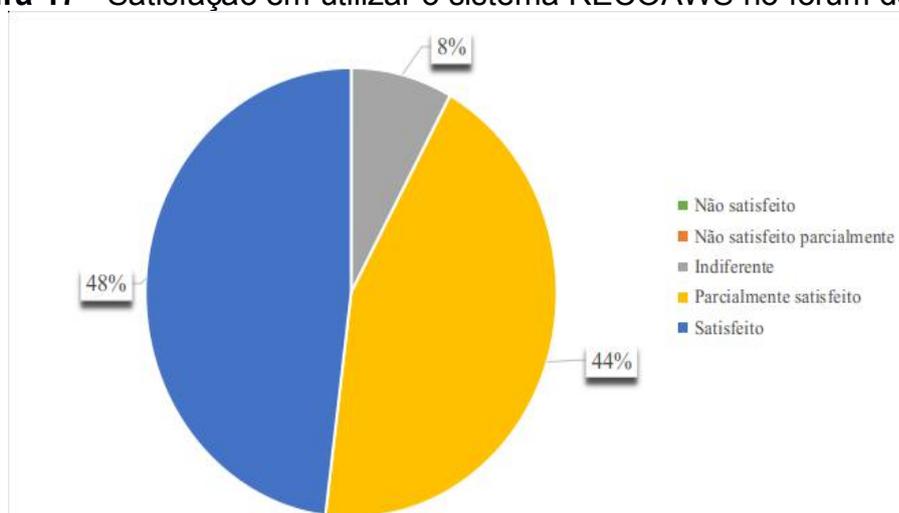


Fonte: o autor (2016)

Buscou-se, nesta abordagem, descobrir sobre a variável de “integração do sistema RECOAWS com o AVA/Moodle”. A figura 16 ilustrou o seguinte resultado: 56% dos alunos responderam que a integração do sistema com o AVA facilitou a busca por OAs; 40% dos alunos concordaram parcialmente que integração do sistema possibilitou a busca por OAs; e 4% disseram que foram indiferentes à integração do sistema com o AVA. Após investigar sobre a integração do sistema RECOAWS com um AVA, os alunos responderam a seguinte pergunta: “Fiquei satisfeito em utilizar o sistema RECOAWS no fórum do AVA/Moodle”. Neste momento, tem-se o objetivo de pesquisar sobre a variável “satisfação dos alunos em

utilizar o sistema RECOAWS no AVA/Moodle”. A figura 17 apresentou a satisfação dos alunos ao usar o sistema RECOAWS no AVA/Moodle.

Figura 17 - Satisfação em utilizar o sistema RECOAWS no fórum do AVA



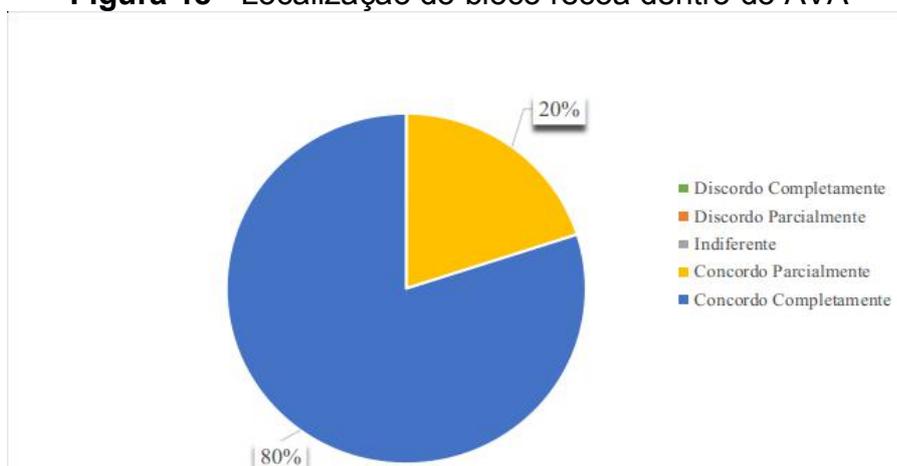
Fonte: o autor (2016)

Como pode-se observar, a figura 17 mostrou o seguinte grau de satisfação por parte dos alunos: 48% ficaram satisfeitos quanto à utilização do sistema RECOAWS no AVA/Moodle; 44% ficaram parcialmente satisfeitos com utilização do sistema no AVA/Moodle; e 8% dos alunos foram indiferentes quanto ao uso do sistema RECOAWS no AVA. Na próxima seção, foram abordadas as repostas dos professores.

6.1.2 Professores

Para analisar as respostas dos professores elaborou-se um questionário (Apêndice A) com perguntas abertas e fechadas, em que eles responderam, através do *Google Forms*, sobre o sistema RECOAWS. Os professores responderam a seguinte pergunta: “Foi fácil localizar o bloco recoa no AVA, através dos fóruns?”. O objetivo dessa pergunta foi a investigar a variável “localizar bloco recoa no AVA/Moodle”. A figura 18 mostrou o resultado sobre a localização do bloco no AVA.

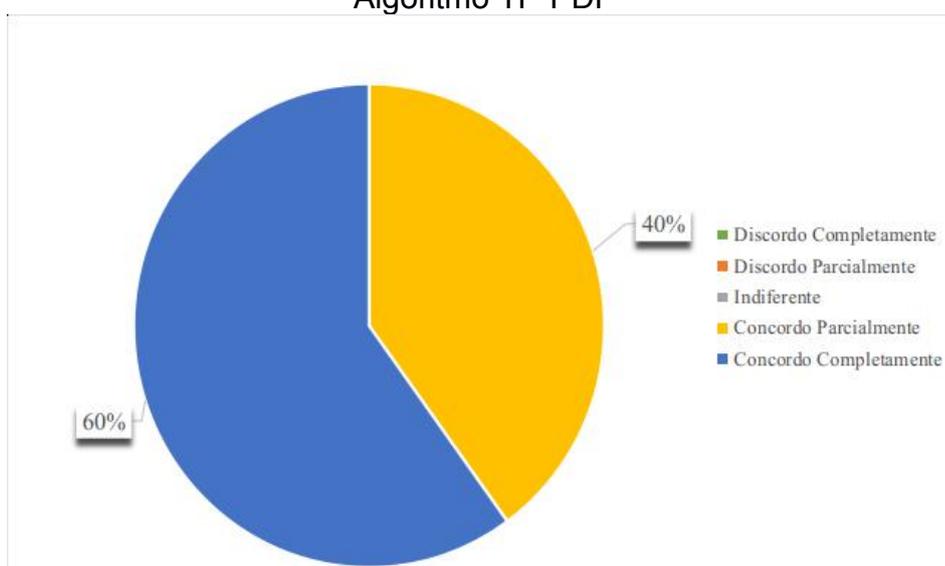
Figura 18- Localização do bloco recoa dentro do AVA



Fonte: o autor (2016)

A figura 18 obteve o seguinte resultado: 80% dos professores responderam que foi fácil a localização do sistema de recomendação no AVA; e 20% dos professores tiveram alguma dificuldade em encontrar o sistema de recomendação. Depois de investigar a localização do sistema de recomendação, os professores responderam: “As palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns?”. Começou-se a pesquisar sobre a variável “palavras-chave extraídas dos fóruns”. A figura 19 apresentou o resultado sobre palavras-chave extraídas do AVA.

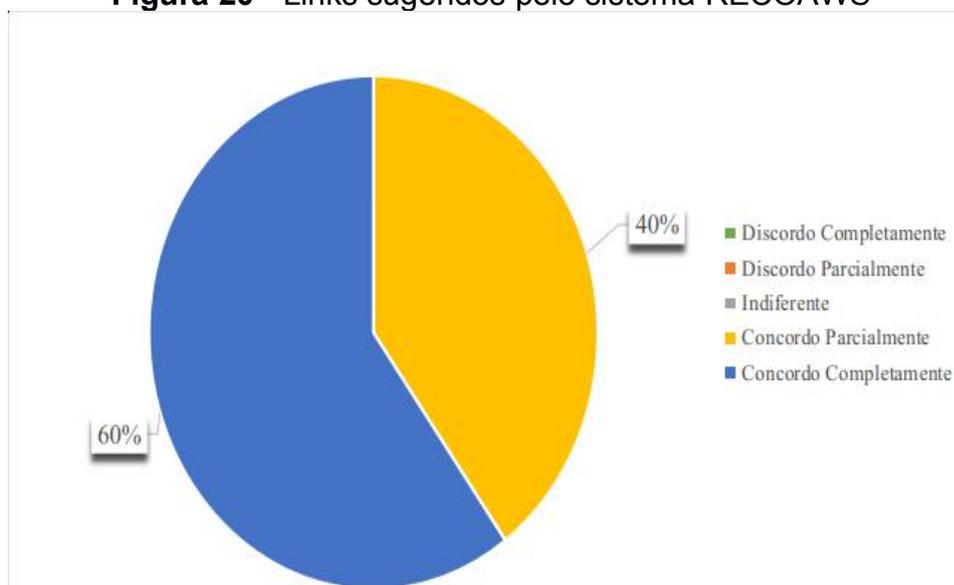
Figura 19 – Palavras-chave extraídas pelo sistema RECOAWS, através do Algoritmo TF*PDF



Fonte: o autor (2016)

A figura 19 apresentou o seguinte resultado: 60% dos professores responderam que as palavras-chave estavam de acordo com as discussões dentro dos fóruns do AVA/Moodle; e 40% dos professores concordaram parcialmente que elas estavam de acordo com as discussões dentro dos fóruns do AVA. Após investigar as “palavras-chave extraídas dos fóruns de um AVA/Moodle”, os professores responderam a seguinte pergunta: “As palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns?”. Começou a se pesquisar sobre a variável “*links* sugeridos pelo sistema RECOAWS”. A figura 20 ilustrou o resultado sobre links sugeridos pelo sistema RECOAWS.

Figura 20 - Links sugeridos pelo sistema RECOAWS

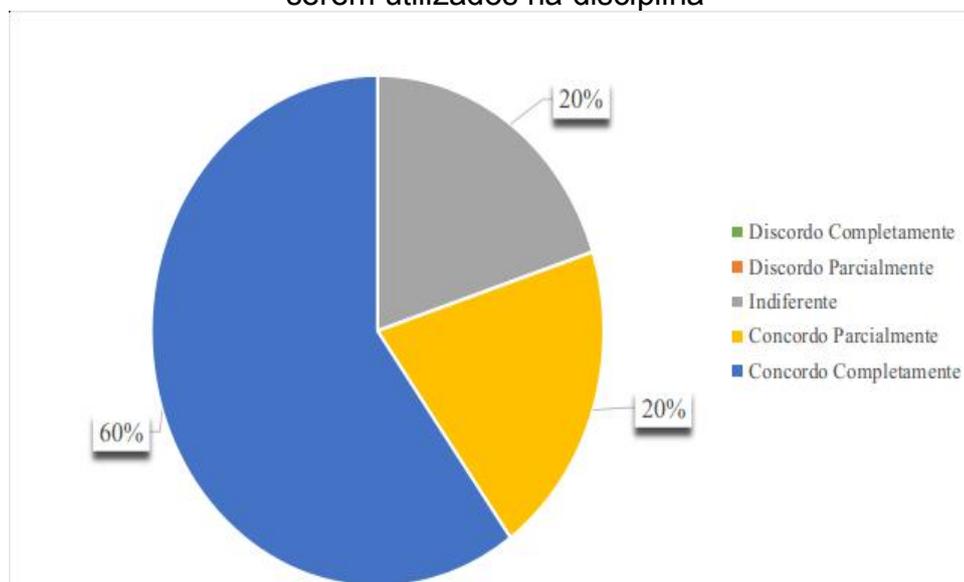


Fonte: o autor (2016)

Os resultados do questionário respondido pelo professores, em relação aos *links* sugeridos pelo sistema RECOAWS, foram os seguintes: 60% dos professores responderam que concordam que os *links* estão de acordo com as discussões dentro dos fóruns de um AVA; 40% dos professores concordaram parcialmente que os links sugeridos estão de acordo com as discussões dentro dos fóruns. Após investigar a variável “*links* sugeridos pelo sistema RECOAWS”, os professores responderam a seguinte pergunta: “Os OAs disponibilizados pelo sistema RECOAWS foram adequados para serem utilizados pela disciplina?”, começou-se a pesquisar sobre a variável “OAs disponibilizados pelo sistema”. A figura 21

apresentou o resultado sobre os OAs disponibilizados pelo sistema RECOAWS para os professores.

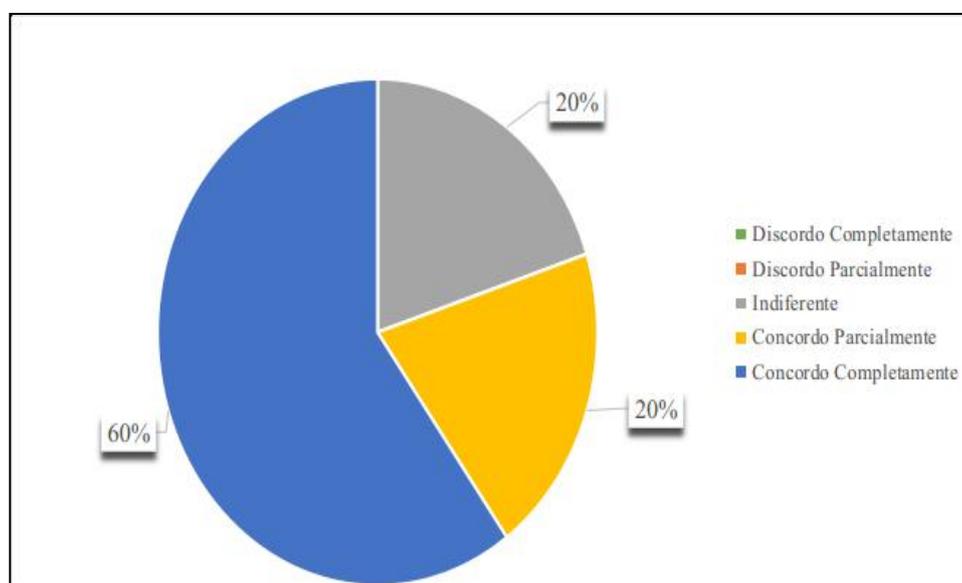
Figura 21 - OAs disponibilizados pelo sistema RECOAWS foram adequados para serem utilizados na disciplina



Fonte: o autor (2016)

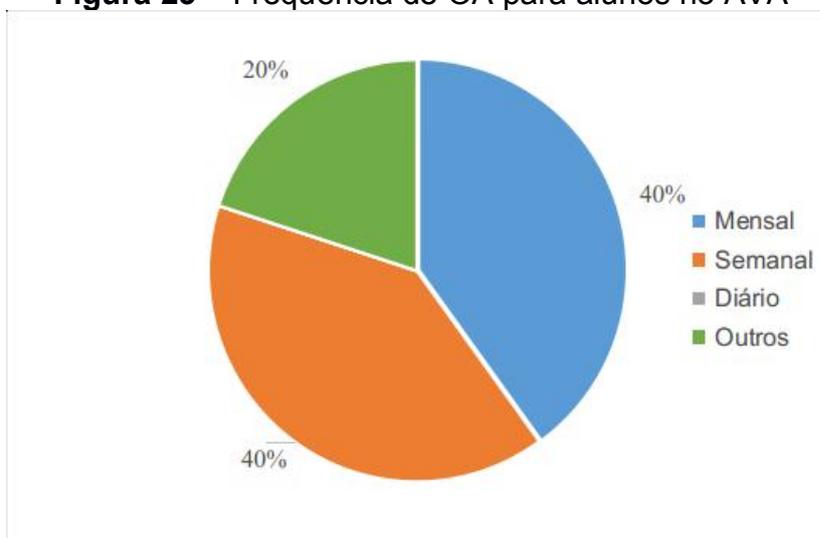
A figura 21 apresentou o seguinte resultado: 60% dos professores responderam que os OAs, disponibilizados pelo sistema, estavam de acordo com as discussões dentro dos fóruns; 20% dos professores concordaram parcialmente que os OAs estão de acordo de com as discussões dentro do fóruns e 20% dos professores não concordam e nem discordam que os OAs disponibilizados, pelo RECOAWS, foram adequados para disciplinas. Depois de investigar a disponibilização dos OAs no sistema, os professores responderam a seguinte pergunta: “A integração do sistema RECOAWS ao AVA/Moodle facilitou a busca de OA para disciplina?”. Começou-se a pesquisar sobre a variável “integração do sistema e a busca do OA para disciplina”. A figura 22 mostrou o resultado sobre a integração e a busca do OA.

Figura 22 - Integração do sistema RECOAWS com AVA facilitou a busca de OA para disciplina



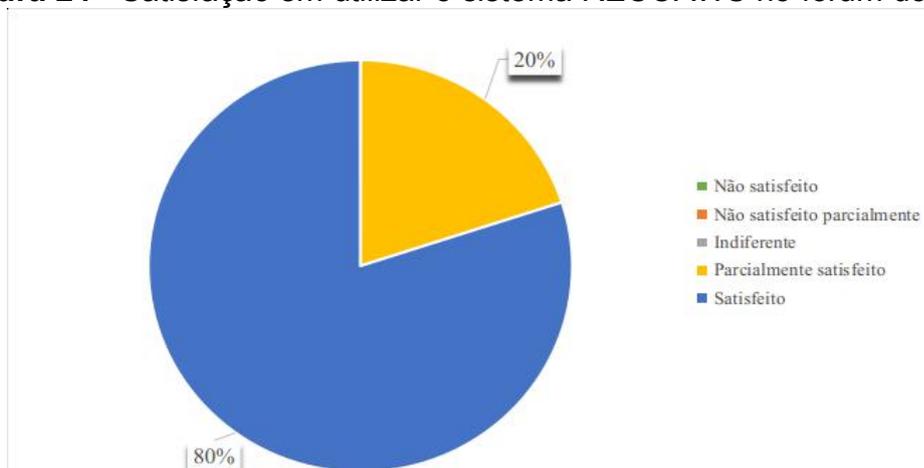
Fonte: o autor (2016)

A figura 22 apresentou o seguinte resultado: 60% dos professores responderam que a integração do sistema RECOAWS com o AVA e a busca de OAs facilitaram o estudo da disciplina; 20% dos professores concordaram parcialmente que a integração do sistema com o AVA e a busca de OAs facilitaram o estudo da disciplina; e 20% dos professores não concordam e nem discordam que a integração do sistema RECOAWS com o AVA e a busca de OAs facilitaram o estudo da disciplina. Depois de investigar a integração e a busca de OAs pelo sistema, os professores responderam a seguinte pergunta: “Com que frequência faço recomendação de OA aos alunos”, começou-se a pesquisar sobre a variável “frequência de recomendação de OA”. A figura 23 mostrou o resultado sobre a frequência de recomendação de OA.

Figura 23 – Frequência de OA para alunos no AVA

Fonte: o autor (2016)

A figura 23 mostrou o seguinte resultado: 40% dos professores responderam que recomendam OAs semanalmente; 40% dos professores recomendam OAs mensalmente; e 20% dos professores não informaram quando recomendam OAs para seus alunos. Depois de investigar sobre a frequência de recomendação de OAs, os professores responderam a seguinte pergunta: “Fiquei satisfeito em utilizar o sistema de recomendação de OAs nos fóruns do AVA/Moodle?”. Começou a se pesquisar sobre a variável “satisfação em utilizar o sistema RECOAWS”. A figura 24 apresenta o resultado sobre a satisfação em utilizar o sistema.

Figura 24 - Satisfação em utilizar o sistema RECOAWS no fórum do AVA

Fonte: o autor (2016)

A figura 24 apresentou o seguinte resultado: 80% dos professores ficaram satisfeitos em utilizar o sistema RECOAWS no fórum de um AVA; 20% ficaram parcialmente satisfeitos com a utilização do sistema.

Portanto, após análise de resultados obtidos pelos questionários de alunos e professores, quanto à localização, à extração de palavras-chave, aos links sugeridos e à satisfação dos usuários, foram analisadas as variáveis da pesquisas sobre o sistema RECOAWS. Na próxima seção, a análise dos resultados será abordada.

6.2 Análise de discussões dos resultados

Para se analisar os resultados obtidos com desenvolvimento do sistema RECOAWS, com base nas amostras, foi utilizado o teste *U* de *Mann-Whitney*, que é um teste não-paramétrico, aplicado para verificar se existe diferença significativa entre duas amostras independentes. Para análise destas amostras, adota-se o nível de significância 5% ($\alpha = 0,05$). Para realizar o teste, formulou-se a hipótese nula (H_0) informando de que “não há diferença de opiniões entre as amostras de professores e alunos, ou seja, para contrariar (H_0) formulou-se H_1 (Hipóteses alternativa) que “há diferença de opiniões entre professores e alunos”. O método do teste de hipótese aplicado será o valor de p (bilateral) e a média por meio do software *BioEstat*³. Por fim, o resultado geral, quando das amostra será destacado no Apêndice E. A tabela 3 – apresenta distribuição de frequência da localização do bloco recoa no AVA, através dos fóruns.

Tabela 3 - Localização do bloco recoa no AVA, através dos fóruns

Categorias	Professores		Alunos	
	Fi	Fr%	fi	Fr%
1- Discordo Completamente	0	0%	0	0%
2- Discordo Parcialmente	0	0%	3	12%
3- Indiferente	0	0%	2	8%
4-Concordo Parcialmente	1	20%	8	32%
5- Concordo Completamente	4	80%	12	48%
Total	5	100%	25	100%
Média dos Professores	4,8			
Média dos Alunos	4,16			
Valor de p	0,2105			

³ <http://www.mamiraua.org.br/pt-br/downloads/programas/bioestat-versao-53/> acessado em: 04/03/2016.

A partir de dados observados, na tabela 3, tem-se o resultado do valor $p=0,2105$, ou seja, maior que o valor de $\alpha = 0,05$. Nesse sentido, não se pode rejeitar a hipótese nula (H_0), ou seja, de que não há diferenças entre as opiniões de professores e alunos com relação à localização do bloco recoa no AVA/Moodle, através dos fóruns. Sendo assim, pode-se concluir que mais de 80% dos entrevistados responderam que foi simples a localização do bloco recoa dentro do fórum do AVA. A tabela 4 mostra que a distribuição de frequência de palavras-chave, extraídas possuem representatividade em relação ao texto dos fóruns.

Tabela 4 - Palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns

Categorias	Professores		Alunos	
	fi	Fr%	fi	Fr%
1- Discordo Completamente	0	0%	0	0%
2- Discordo Parcialmente	0	0%	0	0%
3- Indiferente	0	0%	2	8%
4-Concordo Parcialmente	2	40%	11	44%
5- Concordo Completamente	3	60%	12	48%
Total	5	100%	25	100%
Média dos Professores	4,6			
Média dos Alunos	4,4			
Valor de p (bilateral)	0,5970			

Na tabela 4, mostrou-se o resultado do valor $p=0,5970$; nesta amostra, observa-se que o valor p (bilateral) é maior que o valor de $\alpha = 0,05$. Nesse sentido, não se pode rejeitar a hipótese nula (H_0), pois não há diferenças entre as amostras para a pergunta “palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns”. Sendo assim, pode-se deduzir que mais de 90% dos entrevistados responderam que as palavras-chave possuem representatividade em relação ao texto dentro fóruns. A tabela 5 apresenta distribuição de frequência dos *links* estão de acordo com as discussões dentro dos fóruns de um AVA.

Tabela 5 - Os *links* estão de acordo com as discussões dentro dos fóruns de um AVA

Categorias	Professores		Alunos	
	fi	Fr%	fi	Fr%
1- Discordo Completamente	0	0%	0	0%
2- Discordo Parcialmente	0	0%	0	0%
3- Indiferente	0	0%	2	8%
4-Concordo Parcialmente	2	40%	5	20%
5- Concordo Completamente	3	60%	18	72%
Total	5	100%	25	100%
Média Professores	4,6			
Média Alunos	4,64			
Valor de p (bilateral)	0.7596			

A tabela 5 mostrou o seguinte resultado para o valor $p=0,7596$, significa dizer que o valor p (bilateral) é maior que o valor de $\alpha = 0,05$. Nesse sentido, não se pode rejeitar a hipótese nula H_0 pois não há diferenças de opiniões entre professores e alunos para a pergunta sobre os “*links* recuperados pelo sistema RECOAWS estão de acordo com as discussões dentro dos fóruns de um AVA”. Sendo assim, pode-se concluir que mais de 92% dos entrevistados responderam que os *links* recuperados, pelo sistema RECOAWS estão de acordo com as discussões dentro dos fóruns de um AVA. A tabela 6 apresentou que a distribuição de frequência da disponibilização de OAs pelo sistema RECOAWS foram adequados para serem utilizados pela disciplina.

Tabela 6 – Disponibilização de OA pelo sistema RECOAWS foram adequados para serem utilizados pela disciplina

Categorias	Professores		Alunos	
	fi	Fr%	fi	Fr%
1- Discordo Completamente	0	0%	0	0%
2- Discordo Parcialmente	0	0%	1	4%
3- Indiferente	1	20%	2	8%
4-Concordo Parcialmente	1	20%	10	40%
5- Concordo Completamente	3	60%	12	48%
Total	5	100%	25	100%
Média dos Professores	4,4			
Média dos Alunos	4,32			
Valor de p (bilateral)	0.8674			

Na tabela 6, o resultado para o valor $p=0,8674$ é maior que o valor de $\alpha = 0,05$. Sendo assim, não se pode rejeitar a hipótese nula H_0 , pois não há diferenças entre

as amostras para a disponibilização de OAs pelo sistema RECOAWS, foram utilizados pela disciplina. Logo, pode-se concluir que mais de 88% dos entrevistados responderam que disponibilização de OAs pelo sistema foram adequados para serem utilizados pela disciplina. A tabela 7 mostra a distribuição de frequência de integração do sistema RECOAWS com o AVA facilita a busca de OAs.

Tabela 7 – Integração do sistema RECOAWS com o AVA facilitou a busca de OA

Categorias	Professores		Alunos	
	fi	Fr%	fi	Fr%
1- Discordo Completamente	0	0%	0	0%
2- Discordo Parcialmente	0	0%	0	0%
3- Indiferente	1	20%	1	4%
4-Concordo Parcialmente	1	20%	10	40%
5- Concordo Completamente	3	60%	14	56%
Total	5	100%	25	100%
Média Professores	4,4			
Média Alunos	4,52			
Valor de p	0.9114			

Na Tabela 7, como o valor $p=0.9114$ é maior que o valor de $\alpha = 0,05$, não se pode rejeitar a H_0 de que não há diferenças entre as amostras para a integração do sistema RECOAWS com o AVA, ele facilitou a busca de OAs. Dessa forma, pode-se concluir que mais de 80% dos entrevistados responderam que a integração do sistema RECOAWS com um AVA facilitou a busca de OAs. A tabela 8 apresentou a distribuição de frequência de satisfação dos usuários com sistema RECOAWS nos fóruns de uma AVA.

Tabela 8 – Satisfação em utilizar Sistema RECOAWS nos fóruns de um AVA

Categorias	Professores		Alunos	
	fi	Fr%	fi	Fr%
1- Discordo Completamente	0	0%	0	0%
2- Discordo Parcialmente	0	0%	0	0%
3- Indiferente	0	0%	2	8%
4-Concordo Parcialmente	1	20%	12	48%
5- Concordo Completamente	4	80%	11	44%
Total	5	100%	25	100%
Média dos Professores	4,8			
Média dos Alunos	4,36			
Valor de p (bilateral) =	0.1910			

Na Tabela 8, o valor de $p = 0,1910$, ou seja, é maior que o valor de $\alpha = 0,05$. Desse modo, não se pode rejeitar a hipótese nula H_0 , ou seja, que não há diferenças entre as opiniões de alunos e professores com relação a satisfação em utilizar sistema de recomendação de objetos de aprendizagem nos fóruns de um AVA. Portanto, pode-se concluir que mais de 92% dos entrevistados responderam que tiveram satisfação em utilizar Sistema RECOAWS nos fóruns de um AVA.

Baseado nas tabelas 3 a 8 verifica-se que a variável “localização do sistema RECOAWS no AVA/Moodle”, “palavras-chave extraídas dos fóruns”, “disponibilização de OA recomendados pelo sistema RECOAWS”, “integração do sistema de RECOAWS com o AVA/Moodle” e “satisfação em utilizar o sistema RECOAWS no AVA/Moodle” oferece resultados satisfatório, conforme ficou comprovado pelo teste *U* de *Mann-Whitney* para as amostras. Conclui-se que o protótipo do sistema RECOAWS atendeu aos objetivos propostos inicialmente. No próximo capítulo, apresentam-se as conclusões e trabalhos futuros da tese.

7 CONCLUSÕES E TRABALHOS FUTUROS

Esta pesquisa teve como objetivo o desenvolvimento de um sistema de recomendação integrado ao AVA/Moodle para oferecer aos seus usuários OAs, que fossem de acordo com as discussões dos fóruns. Nesse sentido, ao apropriar-se destes OAs, esperava-se que os alunos e professores fossem capazes de utilizá-los e reutilizá-los de forma crítica, aproveitando esse conteúdo como fundamentação para futuras discussões nos fóruns.

O objetivo geral, como descrito na introdução é: propor um sistema computacional para recomendação de OAs, levando-se em consideração as postagens feitas de dentro dos fóruns de um AVA. Para atingir o objetivo geral, a pesquisa atendeu todos os objetivos específicos que foram propostos inicialmente para o desenvolvido sistema RECOAWS, utilizando o algoritmo TF*PDF, no sentido de extrair palavras-chave, baseada em postagens realizadas nos fóruns de uma AVA/Moodle. Nesse contexto, desenvolveu-se o sistema RECOAWS a partir de uma arquitetura *Web Services*, possibilitando a recomendação de OAs para os usuários, baseados em postagens extraídas de dentro dos fóruns. Para avaliar a contribuição do sistema foi disponibilizado para professores e alunos dentro do AVA/Moodle, no curso de Engenharia de Computação, nas disciplinas Modelagem de Banco de Dados e Interface de usuário e sistemas computacionais, o sistema denominado RECOAWS, para recomendação de OA.

Nesse contexto, com o objetivo de saber se as palavras-chave extraídas de dentro dos fóruns estavam de acordo com as discussões, foi feita a avaliação dos usuários sobre o sistema RECOAWS, que apresentou, na tabela 4, o seguinte resultado: mais de 90% disseram que as palavras-chave possuem representatividade em relação ao texto dentro dos fóruns. Nessa abordagem, não se observou diferença de opiniões para os entrevistados em relação às palavras-chave.

Assim, com a pretensão de saber os *links* sugeridos, pelo sistema RECOAWS, estavam de acordo com as discussões dentro dos fóruns de um AVA, a tabela 5 apresentou os seguintes resultados mais de 92% dos *links* sugeridos, estavam de acordo com as discussões dentro do AVA. Sendo assim, não se observou diferença de opiniões para os entrevistados em relação aos *links* sugeridos pelo sistema. Dessa forma, investigou-se também que objetivos específicos, sobre a

disponibilização de OAs pelo sistema RECOAWS, demonstraram resultados satisfatórios, por parte das amostras analisadas. Conforme a tabela 6, mais de 80% dos entrevistados avaliaram que o sistema atendeu à expectativa em relação à disponibilização de OAs dentro dos fóruns do AVA.

Sendo assim, verificou-se que não houve divergência de opiniões nas amostras analisadas. Na avaliação sobre a integração do sistema do RECOAWS com os fóruns do AVA, pode-se constatar, a partir da análise do resultado da tabela 7, que mais de 80% da amostra pesquisada respondeu que é simples a integração do sistema com o AVA. Como se pode observar, não houve divergência de opiniões entre os entrevistados.

Por fim, aborda-se o grau de satisfação do usuário sobre o sistema RECOAWS. Na tabela 8, essa amostra apresentou os seguintes resultados acima de 92% dos entrevistados relatam que ficaram satisfeito, com o sistema de recomendação. Neste sentido, pode-se concluir que o objetivo satisfação, quanto ao uso da ferramenta, também foi alcançado no desenvolvimento do sistema.

As perspectivas de trabalhos futuros nesta pesquisa apontam adicionar novas funcionalidades a ela e, para isso, seguem alguns desafios:

Recomendação das postagens baseadas em número de citações;

Recomendação das postagens com base nos artigos relacionados;

Mostrar a situação individual de cada aluno de forma mais detalhada em relação às suas postagens;

Ampliação do sistema para ser utilizado em outras estruturas, tais como, redes sociais, *Twitter*, *Facebook*, *Google+*.

Portanto, após análises e discussões conclui-se que o sistema RECOAWS atendeu os objetivos propostos quanto à localização, disponibilização, palavras-chave, *links* sugeridos e satisfação dos usuários para recomendação de OAs. Desse modo, acredita-se que o sistema se apresenta como uma relevante contribuição para a área de Informática na Educação e em muitos outros contextos. A pesquisa desenvolveu um sistema que, potencialmente, diminui o trabalho de busca por informações por parte do usuário, uma vez que o sistema recomenda automaticamente informações correlacionadas ao seu trabalho. Como todo trabalho de pesquisa, esta desencadeou várias problemáticas passíveis de serem

investigadas em outras oportunidades. A seguir, elencam-se alguns aprimoramentos que foram identificados:

Estender o mecanismo de recomendação, incorporando a possibilidade de fazer filtragem colaborativa;

Investigar outros mecanismos de recomendação híbrida;

Aprimorar a arquitetura do sistema de forma que ela se torne genérica e possa ser utilizada em outros domínios do conhecimento;

Intensificar os testes do protótipo para fazer ajustes finos nas etapas de processamento de texto;

Fazer ajustes na arquitetura e no protótipo para melhorar o consumo de recursos computacionais.

REFERÊNCIAS

- ABAWAJY, J. **Analysis of Asynchronous Online Discussion Forums for Collaborative Learning**. International Journal of Education and Learning, v. 1, n. 2, p. 2012.
- ALMEIDA, M. E. B. de. **Educação a distância na Internet: abordagens e contribuições dos ambientes digitais de aprendizagem**. **Educação e Pesquisa**, São Paulo, v. 29, n. 2, p. 327-340, jul./dez. 2003.
- AZEVEDO, B. F. T.; BEHAR, P. A.; REATEGUI, E. B. **Aplicação da mineração de textos em fóruns de discussão**. **RENOTE**, v. 8, n. 3, 2010.
- BAEZA-YATES, R. et al. **Modern information retrieval**. New York: ACM press, 1999.
- BAEZA-YATES, R.; RIBEIRO-NETO, Berthier. **Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition)**. Pearson Education Ltd., Harlow, England 2011, 944 p.
- BETANCUR, D.; MORENO, J.; OVALLE, D. A. **Modelo para la recomendación y recuperación de objetos de aprendizaje en entornos virtuales de enseñanza/aprendizaje Recommendation and retrieval model of learning object in virtual environments for teaching/learning**. Revista Avances en Sistemas e Informática, v. 6, n. 1, 2009.
- BRASIL, FEDERAÇÃO EDUCA. **FEB–Federação de Repositórios Educa Brasil**. 2010. Disponível em: < <http://feb.ufrgs.br> >. Acesso em: 04 jan. 2014.
- BUN, Khoo Khyou.; ISHIZUKA, Mitsuru. **Information Area Tracking and Changes Summarizing in WWW** In: Proc. of WebNet 2001, International Conf. on WWW and Internet. Orlando, Florida, 2001. p. 680-685.
- BUN, K. K. ISHIZUKA, M. **Topic Extraction from News Archive Using TF*PDF Algorithm**. 3rd International Conference on Web Information Systems Engineering, 2002
- BURKE, R. **Hybrid Web Recommender Systems**. In Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.). **The Adaptive Web: Methods and Strategies of Web Personalization.**, Lecture Notes in Computer Science, v. 4321, p. 377-408, Springer, Berlin-Heidelberg, 2007.
- CASALI, A et al. **Sistema inteligente para la recomendación de objetos de aprendizaje**. LACLO 2010 Best Papers. Generación Digital, n. 16, 2011.
- CAZELLA, S.C; REATEGUI, E. B. **Sistemas de Recomendação**. São Leopoldo: XXV Congresso da Sociedade Brasileira de Computação, 2005.
- COLE, J.; FOSTER, H. **Using Moodle: Teaching with the popular open source course management system**. O'Reilly Media, Inc., 2007.
- DCMI **Metadata Terms**, Disponível em: <http://dublincore.org/documents/2012/06/14/dcmi-terms/>. Acesso em maio de 2014.
- FABRE, M. C. J. M; TAROUÇO, L. M. R; TAMUSIUNAS, F. R. **Reusabilidade de objetos de educacionais**. In: RENOTE(Revista Eletrônica de Novas Tecnologias na

Educação). Porto Alegre: s.ed.,v.1, n.1, fevereiro de 2003. Disponível em:<http://www.cinted.ufrgs.br/CESTA/objetosdeaprendizagem_sucesu.pdf>. Acesso em 20 de novembro de 2015.

FERREIRA, V. H; RAABE, A. L. **LORSys–Um Sistema de Recomendação de Objetos de Aprendizagem SCORM**. RENOTE, v. 8, n. 2., jul 2010

FIELDING, R. T. **Architectural styles and the design of network-based software architectures**. Tese de Doutorado. University of California, Irvine. 2000.

FONSECA, L. C. C.; FARIAS, M. P.; SILVA, R. J. **Sistema de Recomendação de Links para o fomento de discussões em fóruns de um Ambiente Virtual de Aprendizagem**. RENOTE, v. 12, n. 2.

FONSECA, J. S. **Curso de Estatística**. São Paulo: Atlas, 1996.

GARIN, R. S.; LICHTNOW, D.; PALAZZO, L. A. M.; LOH, S.; KAMPPFF, A. J. C.; PRIMO, T.; OLIVEIRA, J. P. M.; LIMA, J. V; **O uso de técnicas de recomendação em um sistema para apoio à aprendizagem colaborativa**. Revista brasileira de informática na educação. V. 14, n. 3 (set./dez. 2006), p. 49-59.

HAN, Jiawei; PEI, Jian; KAMBER, Micheline. **Data mining: concepts and techniques**. Elsevier, 2011.

HAUGHEY, Margaret; MUIRHEAD, Bill. **Evaluating Learning Objects for Schools**. E-Journal of Instructional Science and Technology, v. 8, n. 1, p. 1, 2005.

HEERY, R; ANDERSON, S. **Digital repositories review. 2005**. Joint Information Systems Committee: UKOLN.

IEEE Learning Technology Standards Committee (LTSC). **Draft Standard for Learning Object Metadata (IEEE 1484.12.1-2002)**. Julho de 2002. Disponível em: http://ltsc.ieee.org/doc/wg12/LOM_1484_12_1_v1_Final_Draft.pdf. Acesso em: 02 maio 2014.

IMS. **IMS Global Learning Consortium**. 2002. Disponível em: <http://www.imsglobal.org/>. Acesso em maio de 2014. Acesso em: 05 maio 2014.

JAHNAVI, Y. RADHIKA, Y. **A Cogitate Study on Text Mining**. **International Journal of Engineering and Advanced Technology (IJEAT)** ISSN: 2249 – 8958, v.1, Issue-6, aug. 2012.

JONES, K. S. **A statistical interpretation of term specificity and its application in retrieval**. **Journal of Documentation**. v.28 1972. p. 11-21

KAUR, K.; GUPTA, V. **A Survey of Topic Tracking Techniques**. **International Journal of Advanced Research in Computer Science and Software Engineering**. v.2, Issue 5, may. 2012

KLEMMANN, M.; REATEGUI, E. B.; LORENZATTI, A.. **O Emprego da Ferramenta de Mineração de Textos SOBEK como Apoio à Produção Textual**. In: Anais do Simpósio Brasileiro de Informática na Educação. 2009.

KONCHADY, M. **Text Mining Application Programming**. Thomson. 2006.

KUECHLER, W. L. **Business applications of unstructured text**. Communications of ACM, v.50, n.10, p. 86–93, 2007.

LEITE, F.; AMARO, B.; BATISTA, T.; COSTA, M. **Repositórios Institucionais: Boas Práticas para Construção de Repositórios Institucionais da Produção Científica**. Brasília: IBICT, 2012.

LIU, Yi. et al. **Characterizing RESTful Web Services Usage on Smartphones: A Tale of Native Apps and Web Apps**. Web Services (ICWS), 2015 IEEE International Conference on. IEEE, 2015.

LOPES, M. C. S. **Mineração de Dados Textuais Utilizando Técnicas de Clustering para o Idioma Português**. Tese, Universidade Federal do Rio de Janeiro, 2004.

MA, H. **Hot Topic Extraction Using Time Window**. International Conference on Machine Learning and Cybernetics, Guilin, 10-13 July 2011.

MACEDO, A. L.; BEHAR, P. A.; REATEGUI, E. B. **Rede de Conceitos: tecnologia de mineração de texto para apoiar práticas pedagógicas no acompanhamento da escrita coletiva**. Revista Brasileira de Informática na Educação, v. 19, n. 01, p. 04, 2011.

MANOUSELIS, N.; VUORIKARI, R.; VAN ASSCHE, F. **Simulated analysis of MAUT collaborative filtering for learning object recommendation**. In: **Proceedings of the 1st Workshop on Social Information Retrieval for Technology Enhanced Learning**. 2007. p. 27-35.

MOODLE ORG. Disponível em: <<https://moodle.org/>>. Acesso em: 13 março 2015.

OLIVEIRA, G. P. **O Fórum em um Ambiente Virtual de Aprendizado Colaborativo**. **Revista Digital de Tecnologia Educacional e Educação a Distância**.v. 2. 2005.

OPENDOAR (2015). **Directory of open access repositories**. <http://www.opendoar.org/>. acesso em 20 de nov. 2015.

PENTEADO, F.; GLUZ, J. C.; GALAFASSI, C. **Suporte Técnico/Pedagógico aos OAs pelos AVAs: uma Análise Comparativa das Pesquisas Recentes**. In: Anais do Simpósio Brasileiro de Informática na Educação. 2011.

PIVA, D.; PUPO, R.; GAMEZ, L.; OLIVEIRA, S. **EaD na Prática. Planejamento, métodos e ambientes de educação online**. Rio de Janeiro: Elsevier, 2011.

REN, Y.; DU, Y.; HUANG, X.; XU, Y. **Topic Detection of News Stories with Formal Concept Analysis**. **Journal of Information & Computational Science** v.8, n. 9, p. 1675-1682, 2011.

REZENDE, S. O.; MARCACINI, R.; M. MOURA, M. F. **O uso da Mineração de Textos para Extração e Organização Não Supervisionada de Conhecimento**. Revista de Sistemas de Informação da FSMA n. 7, p. 7-21, 2011

RICCI, F.; ROKACH, L.; SHAPIRA, B. **Introduction to recommender systems handbook**. In: **Recommender systems handbook**. Springer US, 2011. p. 1-35.

RICHARDSON, L.; RUBY, S. **RESTful web services**. O'Reilly Media, Inc., 2008

RODRIGUES, A. P. **Integração de ambiente virtual de aprendizagem com repositório digital**. Tese de Doutorado, Porto Alegre, 2012.

- ROMERO, C.; VENTURA, S.; PECHENIZKIY, M.; BAKER, R. S. **Handbook of educational data mining**. CRC Press, 2010.
- ROMERO, C.; VENTURA, S.; GARCÍA, E. **Data mining in course management systems: moodle case study and tutorial**. Computers & Education, v. 51, n. 1, p. 368-384, 2008.
- SALTON, G.; BUCKLEY, C. 1988. **Term-weighting approaches in automatic text retrieval**. *Information Processing & Management* ,v.24, n.5, p. 513-523, 1988.
- SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, P. B. **Metodologia de la investigación**. 2006.
- SCHERER, S. **Educação bimodal: habitantes, visitantes ou transeuntes?** In: VALENTE, J. A. e BUSTAMANTE, S. B. V. Educação a Distância: prática e formação do profissional reflexivo. São Paulo: Avercamp, 2009. 259 p.
- SHINTAKU, M.; MEIRELLES, R. **Manual do Dspace: administração de repositórios**. Salvador: EDUFUBA, 2010.
- SILVA, R. S. **Moodle para autores e tutores**. São Paulo: Novatec, 2011.
- SILVA, REINALDO; GOMES, M. A. N. ; VICARI, R. M. ; FONSECA, L. C. C. ; FARIAS, M. P. . **Modelo Computacional para Recomendação de Objeto de Aprendizagem baseado em anotações extraídas do Ambiente Virtual de Aprendizagem**. In: IX Congresso Latino-Americano Interdisciplinar do Adolescente (CLIOA) 2015, 2015, Porto Alegre. Anais do IX Congresso Latino-Americano Interdisciplinar do Adolescente (CLIOA) 2015, 2015.
- SILVA, FRANCISCO ; SILVA, JOSENILDO DA ; SILVA, REINALDO ; FONSECA, LUÍS CARLOS. **Um modelo preditivo para diagnóstico de evasão baseado nas interações de alunos em fóruns de discussão**. In: XXVI Simpósio Brasileiro de Informática na Educação, 2015, Maceió. p. 1187-1196.
- SILVA, F. C.; SILVA, R. J. ; FONSECA, L. C. C.; SILVA, J. C. **Uma ferramenta para visualização das tendências de evasão geradas por mineração de dados a partir das interações em fóruns de discussão**. In: X Conferência Latino-Americana de Objetos e Tecnologias de Aprendizagem - LACLO 2015, 2015, Maceió. Uma ferramenta para visualização das tendências de evasão geradas por mineração de dados a partir das interações em fóruns de discussão. Porto Alegre: Sociedade Brasileira de Computação - SBC, 2015, 2015. p. 237-246.
- SOMMERVILLE, I. **Engenharia de Software**. 9. ed. São Paulo: Pearson Prentice Hall, 2011.
- VICARI, R. M.; BEZ, M.; SILVA, J. M. C.; RIBEIRO, A.; GLUZ, J. C.; PASSERINO, L.; SANTOS, E.; PRIMO, T.; ROSSI, L.; BORDIGNON, A.; BEHAR, P. FILHO, R.; ROESLER, Valter. **Proposta Brasileira de Metadados para Objetos de Aprendizagem Baseados em Agentes (OBAA)**. Renote: Revista Novas Tecnologias na Educação [recursos eletrônico]. Porto Alegre, 2010.
- VICARI, R. et al. **Padrão para Metadados de Objetos de Aprendizagem Multiplataforma**. Porto Alegre: UFRGS, 2009.

- VICARI, R. M. ; GLUZ, J. C.; SANTOS, E. R.; PRIMO, T. T. ; ROSSI, L. H. L. ; BORDIGNON, A. ; PASSERINO, L. M. Relatório Técnico RT-OBAA-01 **Proposta de Padrão para Metadados de Objetos de Aprendizagem Multiplataforma**. 2009.
- VIEIRA, F.J.R.; NUNES, M.A.S.N. Dica: **Sistema de Recomendação de Objetos de Aprendizagem Baseado em Conteúdo**. Scientia Plena, v. 8,n. 5, 2012.
- WAGH, K; THOOL, R. **A comparative study of soap vs rest web services provisioning techniques for mobile host**. Journal of Information Engineering and Applications, v. 2, n. 5, p. 12-16, 2012.
- W3C WORLD WIDE WEB CONSORTIUM. **Simple Object Access Protocol(SOAP)** Disponível em: <<http://www.w3.org/TR/2000/NOTE-SOAP-20000508>>. Acesso em: 11de jul 2015.
- W3C, 2015. **“Sobre a W3C”**. Disponível em: < <http://www.w3c.br>>. Acesso em 10 de jun. de 2015.
- WILEY, D. **Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy, in The instructional use of learning objects**: 2000 Disponível em: <<http://reusability.org/read/chapters/wiley.doc>> Acesso em 15 de junho de 2014.
- WIVES, Leandro K. **Tecnologias de descoberta de conhecimento em textos aplicadas à inteligência competitiva**. Exame de Qualificação. Universidade Federal do Rio Grande do Sul, , Rio Grande do Sul, 2002. Programa de Pós-Graduação em Computação.
- YANG, Yao Jung; WU, Chuni. **An attribute-based ant colony system for adaptive learning object recommendation**. Expert Systems with Applications, v. 36, n. 2, p. 3034-3047, 2009.
- XML EXTENSIBLE MARKUP LANGUAGE. **Extensible Markup Language (XML) 1.0 (Fifth Edition)**. Disponível em: <<http://www.w3.org/TR/xml/>>.Acesso em: 25 out. 2015
- ZIMMERMANN, O.; TOMLINSON, M.; PEUSER, S. **Perspectives on Web Services: Applying SOAP, WSDL and UDDI to Real-World Projects**.Springer Science & Business Media, 2012.
- ZHE, G.; DONG, L.; QI, L.; JIANYI, Z.; YANG, X.; XINXIN, N.; **An Online Hot Topics Detection Approach Using the Improved Ant Colony Text Clustering Algorithm**. Journal of Convergence Information Technology (JCIT),v. 7, n. 2, February 2012.

APÊNDICE A – QUESTIONÁRIO PROFESSORES

Sistema de Recomendação de Objeto de Aprendizagem - RECOAWS - Professores

1. Foi fácil localizar o sistema de recomendação no AVA, através do Fóruns.
 Completamente Difícil Difícil mais ou menos fácil
 Fácil Completamente Fácil

2. As palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns?
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

3. Os links sugeridos estão de acordo com o tema da discussão?
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

4. Os OAs, disponibilizados pelo sistema RECOAWS, foram adequados para serem utilizados pela disciplina.
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

5. A integração do sistema de RECOAWS ao AVA/MOODLE facilita a busca de OA.
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

6. A integração do sistema de recomendação ao AVA/MOODLE despertou interesse do aluno pela disciplina.
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

7. Qual o grau de satisfação em utilizar o sistema RECOAWS no AVA/Moodle.
 completamente insatisfeito Insatisfeito Mais ou menos satisfeito
 Satisfeito Completamente Satisfeito

8. Costumo recomendar/sugerir e disponibilizar OAs aos meus alunos através do AVA.
 Não Completamente Não Parcialmente Indiferente
 Sim Parcialmente Sim Completamente

9. Com que frequência faço recomendação de OA aos alunos no AVA.
 Diariamente Semanalmente Quinzenalmente
 Mensalmente Outros

10. Faça crítica e sugestões sobre o sistema RECOAWS.

APÊNDICE B – QUESTIONÁRIO ALUNOS

Sistema de Recomendação de Objeto de Aprendizagem - RECOAWS – Alunos

1. Qual é a sua frequência de acesso ao AVA?
 Diariamente Semanalmente Quinzenalmente
 Mensalmente Outros

2. Foi fácil localizar o sistema de recomendação no AVA, através dos Fóruns.
 Completamente Difícil Difícil Mais ou menos fácil
 Fácil Completamente Fácil

3. As palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns?
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

4. Os links sugeridos estão de acordo com o tema da discussão?
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

5. Os OAs, disponibilizados pelo sistema RECOAWS, foram adequados para serem utilizados pela disciplina.
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

6. A integração do sistema de RECOAWS ao AVA/MOODLE facilita a busca de OA.
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

7. A integração do sistema de recomendação ao AVA/MOODLE despertou meu interesse pela disciplina.
 Discordo Completamente Discordo Parcialmente Indiferente
 Concordo Parcialmente Concordo Completamente

8. Qual o grau de satisfação (a) em utilizar o sistema RECOAWS no AVA/Moodle.
 Completamente insatisfeito Insatisfeito Mais ou menos satisfeito
 Satisfeito Completamente Satisfeito

9. Faça crítica e sugestões sobre o sistema RECOAWS.

APÊNDICE C – DECLARAÇÃO DE LIBERAÇÃO DO CCT/UEMA

UNIVERSIDADE
ESTADUAL DO
MARANHÃO

**DECLARAÇÃO DE LIBERAÇÃO DA PESQUISA**

Eu, Jorge de Jesus Passinho e Silva, na qualidade de Diretor de Centro de Ciências Tecnológica- CCT/UEMA. Autorizo a realização da pesquisa intitulada Sistema de recomendação de objeto de aprendizagem baseado em postagens extraída do ambiente virtual de aprendizagem, a ser conduzido sob a responsabilidade do professor Reinaldo de Jesus da Silva, matrícula 2207777, e declaro que o CCT/UEMA disponibilizou infraestrutura necessária a realização da pesquisa.

São Luís, 07 de Março de 2016
Jorge de Jesus Passinho e Silva
CPF:238376593-72
Matrícula: 9571

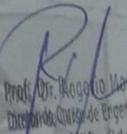
APÊNDICE D – DECLARAÇÃO DE LIBERAÇÃO DO CURSO DE ENGENHARIA DE COMPUTAÇÃO


Universidade Estadual do Maranhão
Curso Graduação em Engenharia da Computação

DECLARAÇÃO DE LIBERAÇÃO DA PESQUISA

Eu, Rogério Moreira Lima Silva, na qualidade de Diretor de Curso de Engenharia de Computação - CCT/UEMA. Autorizo a realização da pesquisa intitulada Sistema de recomendação de objeto de aprendizagem baseado em postagens extraída do ambiente virtual de aprendizagem, a ser conduzido sob a responsabilidade do pesquisador professor Reinaldo de Jesus da Silva, matrícula 2207777, e declaro que o Curso de Engenharia de Computação disponibilizou infraestrutura necessária à realização da referência pesquisa.

São Luis, 09 de Março de 2016.


Prof. Dr. Rogério Moreira Lima Silva
Diretor do Curso de Engenharia de Computação
Matr. 2448488
Rogério Moreira Lima Silva
Matricula: 2448488
CPF: 489630973-15

SCIENTIA AD VITAM

Cidade Universitária Paulo VI. C.P. 09, Tirirical – CEP. 65055-970 – São Luís/MA. Fones: (98) 3245-5461 Fax: (98) 3245-5882
C.N.P.J. 06.352.421/0001/68 - Criada nos termos da Lei nº. 4.400 de 30.12.1981

APÊNDICE F - CÓDIGO FONTE PRINCIPAL MÉTODO DO BLOCO RECOAWS

```

public function get_content(){
    if($this->content !== null) return $this->content;

    global $DB, $USER, $CFG, $COURSE, $PAGE;

    echo '<input type="hidden" id="usuario-logado" value="" . $USER->username . "" />';

    $PAGE->requires->js(new moodle_url($CFG->wwwroot .
'/blocks/RECOA/module.js'));

    $sqlPostagens = getSqlpostagens();
    $result = $DB->get_records_sql($sqlPostagens, array($COURSE->id, $_GET['d'] ));
    $lista_posts = array();
    foreach ($result as $record) {
        $temp = preg_replace('/(<V[^>]+?>)(<[^>V][^>]*?>)/', '$1 $2', $record->message);
        $temp = str_replace("&nbsp;", " ", $temp);
        $temp = strip_tags($temp);
        $lista_posts[] = $temp;
    }

    $payload = json_encode(array("textos" => $lista_posts));
    $service_url = URL_WSRECOA . '/recomendacoes';
    $curl = curl_init($service_url);
    curl_setopt($curl, CURLOPT_RETURNTRANSFER, true);
    curl_setopt($curl, CURLOPT_POST, true);
    curl_setopt($curl, CURLOPT_POSTFIELDS, $payload);
    curl_setopt($curl, CURLOPT_HTTPHEADER, array('Authorization:
xxxx_AuthorizationKey__xxxx'));
    $curl_response = curl_exec($curl);

    if ($curl_response === false) {
        $info = curl_getinfo($curl);
        curl_close($curl);
        die('Erro: ' . var_export($info));
    }
    curl_close($curl);

    $decoded = json_decode($curl_response);
    if (isset($decoded->response->status) && $decoded->response->status ==
'ERROR') {
        die('Erro:' . $decoded->response->errormessage);
    }

    $recomendacoes = $decoded->recomendacoes;
    $palavras = $decoded->palavraschave;

```

```
foreach ($palavras as $palavra) {
    $palavraschave .= $palavra . ", ";
}

$texts = $videos = $images = "";
foreach ($recomendacoes as $rec) {
    $texts .= showTexts($rec, $palavraschave);
    $videos .= showVideos($rec, $palavraschave);
    $images .= showImages($rec, $palavraschave);
}

$recoaTables = block_recoa_print_recomendations($texts, $videos, $images);
$this->content = new stdClass;
$this->content->text = $recoaTables;
$this->content->footer = 'RECOA';
return $this->content;
}
```

APÊNDICE G - SERVIÇO WEB RECOAWS

```

$app -> get('/recomendacoes', function() use($app)
{
$body = $app -> request() -> getBody();
$payload = json_decode($body);
$textos = $payload -> textos;
$tipoRecoa = $app -> request() -> post('tipoRecoa');
if ($tipoRecoa == null)
{
$tipoRecoa = "All";
}

$tipoRecoa = urldecode($tipoRecoa);
$response = array();
$objRecomendacaoLinks = new GerarRecomendacao();
$result = $objRecomendacaoLinks -> getRecomendacaoLinks($textos,
$tipoRecoa);

if ($result)
{
echoResponse(200, $result);
} else
{
echoResponse(404, null);
}
});

// Cálculo dos Pesos
Function calcularPesosDosTermos($documento_unificado,
$lista_documentos) {
$TF = new PesoTF1();

```

```

$PDF = new PesoPDF();
$TFPDF = new PesoTFPDF();
$listaTermos = $documento_unificado->getListaTermosProcessados();
$tamanhoListaTermos = count($listaTermos);
for ($i = 0; $i < $tamanhoListaTermos; $i++)
{
    $termo = $listaTermos[$i];
    $termfreq = $TF->calcularPeso($termo);

    $termo->setTf($termfreq);
    $inverseDF = $PDF->calcularPeso($termo->getTermo(),
$lista_documentos);
    $termo->setIdf($inverseDF);

    $tfidf = $TFPDF->calcularPeso($termfreq, $inverseDF);
    $termo->setTfidf($tfidf);
    $listaTermos[$i] = $termo;
}
return $listaTermos;
}

```

// Extração de palavras chave, cálculo de pesos e busca de recomendações

```

function getRecomendacaoLinks($lista_textos, $tipoRecoa)
{
    $query = "";
    $lista_documentos = $this->cria_documentos($lista_textos);
    $token = new Tokenization();
    $documento_unificado = $token->juntaDocumentos($lista_documentos);
    $calculoPesos = new CalculoPesos();
    $termos = $calculoPesos->calcularPesosDosTermos($documento_unificado, $lista_documentos);
    $termosOrdenados = $documento_unificado->ordenaListaTermos($termos);
}

```

```

for($i=0; $i<5; $i++)
    $query .= $termosOrdenados[$i]->getTermo() . " ";
$response = array();
$response['palavraschave'] = array();
$response['recomendacoes'] = array();
$token = strtok(trim($query), " ");
while ($token != false) {
    $response['palavraschave'][] = $token;
    $token = strtok(" ,:;()?![]{}\'\"");
}

$apiKey = 'xxx_chave_fictícia_xxx';
$bing = new BingSearch($apiKey);
if($tipoRecoa == "Web" || $tipoRecoa == "All")
{
    $jsonobj = $bing->queryWeb($query);
    foreach ($jsonobj->d->results as $value)
    {
        $reco = array();
        $reco['titulo'] = $value->Title;
        $reco['descricao'] = $value->Description;
        $reco['link'] = $value->Url;
        $reco['tipo'] = "web";
        $response['recomendacoes'][] = $reco;
    }
}
if($tipoRecoa == "Image" || $tipoRecoa == "All")
{
    $jsonobj = $bing->queryImage($query);
    foreach ($jsonobj->d->results as $value)
    {
        $reco = array();

```

```
        $reco['titulo'] = $value->Title;
        $reco['link'] = $value->MediaUrl;
        $reco['tipo'] = "image";
        $response['recomendacoes'][] = $reco;
    }
}

if($tipoRecoa == "Video" || $tipoRecoa == "All"){
    $jsonobj = $bing->queryVideo($query);
    foreach ($jsonobj->d->results as $value) {
        $reco = array();
        $reco['titulo'] = $value->Title;
        $reco['link'] = $value->MediaUrl;
        $reco['tipo'] = "video";
        $response['recomendacoes'][] = $reco;
    } }
return $response;
}
```

ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Estamos realizando um estudo de proposta de doutorado com o objetivo de investigar sistema de recomendação de objeto de aprendizagem, baseado em postagens extraídas dos fóruns de um ambiente virtual de aprendizagem. Assim, você está sendo convidado(a) para realizar as atividades desse processo, em um formato a distância. A sua colaboração poderá contribuir para a construção do conhecimento científico e beneficiar perspectivas de intervenções pedagógicas futuras. A participação, na pesquisa, é totalmente voluntária. Esta pesquisa é coordenada pela Professora Dra. Rosa Maria Vicari (PPGIE/UFRGS) e pelo Doutorando Reinaldo de Jesus da Silva, do Programa de Pós-Graduação em Informática na Educação da Universidade Federal do Rio Grande do Sul (UFRGS) com quem podem ser obtidas maiores informações (Av. Paulo Gama, 110 - prédio 12105 - 3º andar sala 332 CEP: 90040-060 - Porto Alegre – RS – Brasil; reinaldo@uema.br).

Se você tiver dúvidas em relação à pesquisa ou quiser comentar algum aspecto relacionado a ela, pode contatar os pesquisadores responsáveis. A participação, na pesquisa, é voluntária. Portanto, caso não queira participar, você não precisa assinar este termo nem participar da pesquisa. O fato de não querer participar da pesquisa não lhe trará nenhum prejuízo.

Após o encerramento do processo, você pode solicitar uma devolutiva individual. Os resultados globais da pesquisa serão publicados posteriormente em algum periódico ou evento científico da área de Informática na Educação, sem identificação da identidade dos participantes. Na apresentação dos resultados desse trabalho, sua identidade será mantida no mais rigoroso sigilo. Serão omitidas todas as informações que permitam identificá-lo(a).

Este documento foi revisado e aprovado pelo Comitê de Ética da Universidade Federal do Rio Grande do Sul (número CAAE 20469713.0.0000.5347).

Pelo presente Termo de Consentimento, eu, _____, declaro que sou maior de 18 anos e que fui informado dos objetivos e da justificativa da presente pesquisa, e estou de acordo em participar dela. Fui igualmente informado: a) da liberdade de participar ou não da pesquisa, bem como do meu direito de retirar meu consentimento, a qualquer momento, e deixar de participar do estudo, sem que isso me traga qualquer prejuízo; b) da garantia de receber resposta a qualquer dúvida acerca dos procedimentos e outros assuntos relacionados com a pesquisa; c) da segurança de que não serei identificado e de que se manterá o caráter confidencial das informações registradas; d) que as informações obtidas serão arquivadas sem identificação pessoal junto ao banco de dados do pesquisador responsável; e) que os dados da pesquisa serão arquivados sob a guarda do pesquisador responsável por cinco anos e depois destruídos.

Data ____/____/____ Assinatura do participante:

Assinatura do pesquisador responsável: _____

ANEXOS B - RESULTADOS DAS AMOSTRAS CALCULADAS ATRAVÉS DO SOFTWARE *BIOESTAT* ⁴

A seguir são apresentadas tabelas com as descrições de todos os resultados gerados pelo software estatístico BioEstat 5.0.

Tabela 9 - Resultados da análise descritiva da amostra dos professores.

Média	4,8
Erro padrão	0,2
Mediana	5
Modo	5
Desvio padrão	0,447213595
Variância da amostra	0,2
Curtose	5
Assimetria	-2,236067977
Intervalo	1
Mínimo	4
Máximo	5
Soma	24
Contagem	5

Fonte: Autor

Tabela 10 - Resultados da análise descritiva da amostra dos alunos.

Média	4,16
Erro padrão	0,205588586
Mediana	4
Modo	5
Desvio padrão	1,02794293
Variância da amostra	1,056666667
Curtose	0,175713882
Assimetria	-1,095569543
Intervalo	3
Mínimo	2
Máximo	5
Soma	104
Contagem	25

Fonte: Autor

⁴ Disponível em: <http://www.mamiraua.org.br/pt-br/downloads/programas/bioestat-versao-53/acesso> em: 20 maio 2016.

Tabela 11 - Resultados da localização do bloco recoa no AVA, através dos fóruns.

Resultados	Professores	Alunos
Tamanho da amostra	5	25
Soma dos Postos (Ri)	100	365
Mediana =	5	4
U =	40	
Z(U) =	1,2521	
p-valor (unilateral) =	0,1053	
p-valor (bilateral) =	0,2105	

Fonte: Autor

Tabela 12 - Resultados das palavras-chave extraídas possuem representatividade em relação ao texto dos fóruns.

Resultados	Professores	Alunos
Tamanho da amostra	5	25
Soma dos Postos (Ri)	87	378
Mediana =	5	4
U =	53	
Z(U) =	0,5287	
p-valor (unilateral) =	0,2985	
p-valor (bilateral) =	0,597	

Fonte: Autor

Tabela 13 – Resultados dos *links* sugeridos estão de acordo com tema da discussão

Resultados	Professores	Alunos
Tamanho da amostra	5	25
Soma dos Postos (Ri)	72	393
Mediana =	5	5
U =	57	
Z(U) =	0,3061	
p-valor (unilateral) =	0,3798	
p-valor (bilateral) =	0,7596	

Fonte: Autor

Tabela 14 - Resultados da integração do sistema RECOAWS ao AVA/MOODLE facilita a busca de OA.

Resultados	Professores	Alunos
Tamanho da amostra	5	25
Soma dos Postos (Ri)	75,5	389,5
Mediana =	5,00	5
U =	60,50	
Z(U) =	0,1113	
p-valor (unilateral) =	0,4557	
p-valor (bilateral) =	0,9114	

Fonte: Autor

Tabela 15 - Resultados do grau de satisfação em utilizar o sistema RECOAWS no AVA/MOODLE.

Resultados	Professores	Alunos
Tamanho da amostra	5	25
Soma dos Postos (Ri)	101	364
Mediana =	5	4
U =	39	
Z(U) =	1,3077	
p-valor (unilateral) =	0,0955	
p-valor (bilateral) =	0,1910	

Fonte: Autor