UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Crystiam Kelle Pereira e Silva

Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais

Juiz de Fora
2015

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Crystiam Kelle Pereira e Silva

Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação, do Instituto de Ciências Exatas da Universidade Federal de Juiz de Fora como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Orientadora: Dra. Fernanda Cláudia Alves Campos.

Juiz de Fora
2015

Ficha catalográfica elaborada através do Programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Pereira e Silva, Crystiam Kelle.
Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais / Crystiam Kelle Pereira e Silva. -- 2015.
140 f.

Orientadora: Fernanda Cláudia Alves Campos
Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação, 2015.

1. Sistemas de Recomendação. 2. redes sociais. 3. perfil do usuário. I. Alves Campos, Fernanda Cláudia, orient. II. Título.

Crystiam Kelle Pereira e Silva

"Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais"

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Computação da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre.

Aprovada em 27 de março de 2015.

BANCA EXAMINADORA

Profª. Dra. Fernanda Cláudia Alves Campos – Orientador
Universidade Federal de Juiz de Fora

Profª. Dra. Regina Maria Maciel Braga Villela
Universidade Federal de Juiz de Fora

Prof. Dr. Victor Stroele de Andrade Menezes
Universidade Federal de Juiz de Fora

Prof. Dr. Sean Wolfgand Matsui Siqueira
Universidade Federal do Estado do Rio de Janeiro

Dedico este trabalho à minha família e a todos que acreditaram que ele poderia ser concretizado. Em especial, à minha mãe e ao companheiro de sempre, Solano.

AGRADECIMENTOS

Meus agradecimentos iniciam-se para aqueles que desde os primeiros dias da minha vida escolar me incentivaram a buscar crescimento através da educação: meus pais. À minha mãe agradeço o carinho, todos os ensinamentos de vida, o incentivo e as palavras de ânimo em todos os momentos de desânimo. Ao meu pai agradeço os ensinamentos que pode me dar até o momento que esteve comigo e a vontade que ele me deixou de sempre vencer para orgulhá-lo, estivesse ele perto ou ali, um pouco mais longe, onde os meus olhos não podem alcançá-lo.

Meu agradecimento ao incentivador, amigo e companheiro de sempre, Solano, pelo apoio em todos os momentos, pelos dias de paz, pela compreensão pela minha ausência, pelo carinho que tranquiliza e por transformar esse desafio em algo mais leve.

Agradeço aos professores Eduardo Barrére e Tarcísio Lima e aos colegas Francisco Henrique, Thiago Nery e Paulo Alceu pelo incentivo no momento da decisão de iniciar essa jornada.

Meus agradecimentos a todos os professores do Programa de Pós-Graduação em Ciência de Computação, em especial aos professores José Maria, Regina Braga pelos muitos ensinamentos, parceria e incentivo.

Ao professor Victor Ströele sou extremamente grata pela paciência, eficiência e delicadeza nos ensinamentos.

À minha orientadora, Fernanda Campos, agradeço por tudo. Seus ensinamentos, sua experiência, sua paciência, sua objetividade, seus conselhos e o incentivo desde o primeiro momento que foram imprescindíveis para a conclusão deste trabalho.

A todos os colegas de mestrado, obrigada pela parceria. Aos participantes do estudo de caso agradeço ainda mais pela disponibilidade e paciência.

Aos amigos e colegas do CEAD, muito obrigada por compreenderem a minha dedicação ao mestrado e contribuírem com ela. Ao ex-diretor do CEAD, Flávio Takakura, agradeço a compreensão da importância na dedicação a esta pesquisa e ao incentivo em todos os momentos.

Agradeço ainda à Universidade Federal de Juiz de Fora pelo incentivo à minha qualificação.

"Quem tem um porquê enfrenta qualquer como."
Victor Frankl

RESUMO

Existem inúmeros recursos educacionais distribuídos em diferentes repositórios que abordam um conjunto amplo de assuntos e que possuem objetivos educacionais distintos. A escolha adequada desses recursos educacionais é um desafio para os usuários que desejam usá-los para a sua formação intelectual. Nesse contexto surgem os Sistemas de Recomendação para auxiliar os usuários nessa tarefa. Para que seja possível gerar recomendações personalizadas, torna-se importante identificar informações que ajudem a definir o perfil do usuário e auxiliem na identificação de suas necessidades e interesses. O uso constante e cada vez mais intenso de algumas ferramentas tecnológicas faz com que inúmeras informações a respeito do perfil, dos interesses, das preferências, da forma de interação e do comportamento do usuário possam ser identificadas em decorrência da interação espontânea que ocorre nesses sistemas. Esse é o caso, por exemplo, das redes socais. Neste trabalho é apresentada a proposta e o desenvolvimento de uma arquitetura capaz de extrair características do perfil e do contexto educacional dos usuários, através da rede social Facebook e realizar recomendações de recursos educacionais de forma individualizada e personalizada que sejam condizentes com essas características. A solução proposta é apoiada por técnicas de extração de informações e ontologias para a extração, definição e enriquecimento das características e interesses dos usuários. As técnicas de Extração de Informação foram aplicadas aos textos associados às páginas curtidas e compartilhadas por usuários nas suas redes sociais para extrair informação estruturada que possa ser usada no processo de recomendação de recursos educacionais. Já as ontologias foram usadas para buscar interesses relacionados aos temas extraídos. A recomendação é baseada em repositório de objetos de aprendizagem e em repositórios de dados ligados e é realizada dentro das redes sociais, aproveitando o tempo despendido pelos usuários nas mesmas. A avaliação da proposta foi feita a partir do desenvolvimento de um protótipo, três provas de conceito e um estudo de caso. Os resultados da avaliação mostraram a viabilidade e uma aceitação relevante por parte dos usuários no sentido de extrair informações sobre os seus interesses educacionais, geradas automaticamente da rede social Facebook, enriquecê-las, encontrar interesses implícitos e usar essas informações para recomendar recursos educacionais. Foi verificada também a possibilidade da recomendação de pessoas, permitindo a formação de uma rede de interesses em torno de um determinado tema, indicando aos usuários bons parceiros para estudo e pesquisa.

Palavras chave: Sistemas de Recomendação, redes sociais, perfil do usuário

ABSTRACT

There are several educational resources distributed in different repositories that address to a wide range of subjects and have different educational goals. The proper choice of these educational resources is a challenge for users who want to use them for their intellectual development. In this context, recommendation systems may help users in this task.In order to be able to generate personalized recommendations, it is important to identify information that will help to define user profile and assist in identifying his/her needs and interests. The constant and ever-increasing use of some technological tools allows the identification of different information about profile, interests, preferences, interaction style and user behavior from the spontaneous interaction that occurs in these systems, as, for example, the social networks. This paper presents the proposal and the development of one architecture able to extract users´ profile characteristics and educational context, from the Facebook social network and recommend educational resources in individualized and personalized manner, consistent with these characteristics. The proposed solution is supported by Information Extraction Techniques and ontologies for the extraction, enrichment and definition of user characteristics and interests. The Information Extraction techniques were applied to texts associated with "LIKE" and shared user´s pages on his social networks to extract structured information that can be used in the recommendation process of educational resources, the ontologies were used to search to interests related to extracted subjects. The recommendation process is based on learning objects repositories and linked data repositories and is carried out within social networks, taking advantage of user time spent at the web. The proposal evaluation was made from the development of a prototype, three proofs of concept and a case study. The evaluation results show the viability and relevant users´ acceptance in order to extract information about their educational interests, automatically generated from the Facebook social network, enrich these information, find implicit interests and use this information to recommend educational resources. It was also validated the possibility of people recommendation, enabling the establishment of interest network, based on a specific subject, showing good partners to study and research.

Keywords: *recommendation system, social network, user profile*

LISTA DE TABELAS

Tabela 1 - Legendas usadas na análise das propostas de recomendação.....44
Tabela 2 - Comparativo entre as propostas de recomendação.....44
Tabela 3 - Principais características da proposta de recomendação BROAD-RSI. 47
Tabela 4 - Elementos contextuais usados no BROAD-RSI. 53
Tabela 5 - Relação entre metadados e informações extraídas. 58
Tabela 6 - Comparação das fontes de recursos educacionais usadas na recomendação. 65
Tabela 7 – Parte da tabela usada no cálculo da aderência dos recursos educacionais. 78
Tabela 8 - Exemplo de ordenação dos recursos educacionais. 79
Tabela 9 – Respostas do Pesquisador A para as recomendações feitas (Pergunta 1).....106
Tabela 10 - Respostas do Pesquisador B para as recomendações feitas (Pergunta 1).....107
Tabela 11 - Respostas do Pesquisador A para as recomendações feitas (Pergunta 2). 108
Tabela 12 - Respostas do Pesquisador B para as recomendações feitas (Pergunta 2).....108
Tabela 13 - Respostas do Pesquisador A para as recomendações feitas (Pergunta 3). 109
Tabela 14 - Respostas do Pesquisador B para as recomendações feitas (Pergunta 3).....110
Tabela 15 - Tipos de entrevistas. 115
Tabela 16 –Possibilidades de resultados de recomendações. 117
Tabela 17 - Métricas para avaliar Sistemas de Recomendação. 117
Tabela 18 - Avaliação dos usuários sobre as características de perfil e de contexto extraídos. 118
Tabela 19 - Avaliação dos participantes sobre interesses extraídos. 119
Tabela 20 - Avaliação dos participantes sobre interesses descobertos. 120
Tabela 21 - Avaliação dos participantes sobre interesses descobertos através das duas abordagens usadas. 120
Tabela 22 - Resultados da avaliação das recomendações. 122
Tabela 23 - Resultado das avaliações para a questão 1 e questão 2. 122

LISTA DE FIGURAS

Figura 1 - Organização da dissertação. 27
Figura 2 - Visão geral da rede de dados do *The Linking Open Data*. Fonte: CYGANIAK; JENTZSCH (2014). 36
Figura 3 - Página da DBPEDIA relacionada ao tema "Sistema de Recomendação". 37
Figura 4- Fluxo simplificado de recomendação do BROAD-RSI. 50
Figura 5 – Representação da arquitetura do BROAD-RSI. 51
Figura 6 - Permissão necessária para a extração de dados do usuário no Facebook. 52
Figura 7 - Exemplos de textos descritivos de páginas no Facebook. 55
Figura 8 - Recursos disponíveis no conjunto de dados ligados da *Open University*. 59
Figura 9 - Recomendação de Links recuperados da DBPEDIA (Tema *Semantic Web*). 60
Figura 10 - Recomendação de Links recuperados da DBPEDIA (Tema *E-learing*). 60
Figura 11 - Recomendação de vídeos do Youtube (Tema *Data Mining*). 61
Figura 12 - Rede de pessoas com um determinado interesse comum. 62
Figura 13 - Recomendação enviada para a rede social do usuário sobre o tema "*Data Mining*". 64
Figura 14 - Recomendação enviada para a rede social do usuário sobre o tema "*Software Engineering*". 64
Figura 15 - Representação da arquitetura do BROAD-RSI (Tecnologias). 67
Figura 16 - Exemplo de informações disponíveis através da API do Facebook. 69
Figura 17 - Informações sobre os livros lidos por um determinado usuário. 70
Figura 18 - Esquema de uso da *AlchemyAPI*. 71
Figura 19 - Extração de entidades, conceitos e palavras-chave usando a AlchemyAPI. 72
Figura 20 - Ontologias a respeito de conceitos sobre o tema "*Pharmacology*". 73
Figura 21 - Ontologias a respeito de conceitos sobre o tema "*Software Engineering*". 74
Figura 22 - Exemplo de SPARQL feita na taxonomia da ACM. 74
Figura 23 - Visão geral do processo de recomendação do BROAD-RSI. 79
Figura 24 - SPARQL usada na recuperação de recursos na DBPEDIA. 80
Figura 25 - Página da DBPEDIA referente ao tema "*Data Mining*". 81
Figura 26 - Links recuperados através da DBPEDIA. 81
Figura 27 – SPARQL executada no *Dataset* da DBPEDIA. 82
Figura 28 - Recomendação de vídeos do Youtube. 83
Figura 29 - Trecho do código para envio da recomendação para o Facebook. 85

Figura 30 - Características de perfil do Pesquisador A. 90
Figura 31 - Características de perfil do Pesquisador B. 91
Figura 32 - Grandes áreas de interesse do Pesquisador A. 91
Figura 33 - Interesses especificamente específicos do Pesquisador A. 92
Figura 34 - Interesses diversos do Pesquisador A. 92
Figura 35 - Grandes áreas de interesse do Pesquisador B. 93
Figura 36 - Interesses especificamente educacionais do Pesquisador B. 93
Figura 37 - Interesses diversos do Pesquisador B. 93
Figura 38 - Termos extraídos do texto descritivo da área de "*Data Mining*" no Facebook. 96
Figura 39 - Termos encontrados através da consulta à uma ontologia usando o termo "*Data Mining*". 97
Figura 40 - Termos extraídos do texto da página descritiva da área de interesse "*Wireless Network*" no Facebook. 97
Figura 41 - Recomendações extraídas do repositório do BROAD-RSI para o tema "*Data Mining*" e seus temas relacionados. 102
Figura 42 - Recomendações de vídeos associados ao tema "*Data mining*" e seus temas relacionados. 103
Figura 43 - Recomendação de links sobre o tema "*Data Mining*" e seus temas relacionados. 103
Figura 44 - Recomendações de recursos educacionais armazenados no repositório do BROAD-RSI a respeito do tema "*Wireless Network*" e seus temas relacionados. 104
Figura 45 - Recomendação de links sobre o tema "*Wireless Network*" e seus temas relacionados. 104
Figura 46 - Recomendação de vídeos sobre o tema "*Wireless Network*" e seus temas relacionados. 105
Figura 47 - Envio de recomendação na página do Facebook do Pesquisador A. 105
Figura 48 - Avaliações dos usuários para as recomendações feitas a eles (Pergunta 1). 108
Figura 49 - Avaliação em relação ao compartilhamento das recomendações nas redes sociais. 109
Figura 50 - Avaliação em relação à aderência da recomendação ao perfil. 110
Figura 51 - Distribuição dos participantes do estudo de caso por atividade profissional. 115
Figura 52 – Tela para o usuário avaliar as recomendações dentro do BROAD-RSI. 116

LISTA DE GRÁFICOS

Gráfico 1 - Avaliação dos interesses extraídos do Pesquisador A. 94
Gráfico 2 - Avaliação dos interesses extraídos do Pesquisador B. 94
Gráfico 3 - Avaliação dos interesses descobertos dos Pesquisadores A e B. 98
Gráfico 4 - Comparativo entre o enriquecimento por ontologia (círculo externo) e o enriquecimento por técnicas de Extração de Informação (círculo interno). 98
Gráfico 5 - Avaliação dos usuários sobre as características de perfil extraídas. 119
Gráfico 6 – Avaliação dos interesses. 121
Gráfico 7 - Avaliação das recomendações. 122
Gráfico 8 – Resultados das avaliações para a questão 1 e questão 2. 123
Gráfico 9 - Uso de redes sociais. 125
Gráfico 10 – Redes sociais nas quais os usuários gostariam de receber recomendações. 126
Gráfico 11 - Tipo de mídia preferida pelo usuário. 128

LISTA DE ABREVIAÇÕES

ACM	*Association for Computing Machinery*
API	*Application Programming Interface*
AVA	Ambiente Virtual de Aprendizagem
FOAF	*Friend of Friend*
HTTP	*Hypertext Transfer Protocol*
LOD	*Linked Open Data*
LOM	*Learning Object Metadata*
MOOC	*Massive Open Online Course*
OA	Objeto de Aprendizagem
OWL	*Web Ontology Language*
PLE	*Personal Learning Environment*
POC	Prova de conceito
RDF	*Resource Description Framework*
RFID	*Radio Frequency Identification*
SIOC	*Semantically-Interlinked Online Communities*
SR	Sistema de Recomendação
URI	*Uniform Resource Identifier*
URL	*Uniform Resource Locator*
W3C	*World Wide Web Consortium*
XML	*eXtensible Markup Language*

SUMÁRIO

1 INTRODUÇÃO 18
1.1 MOTIVAÇÃO 19
1.2 PROBLEMA 22
1.3 HIPÓTESES 23
1.4 METODOLOGIA DE PESQUISA 23
1.5 OBJETIVOS 25
1.6 ORGANIZAÇÃO DA DISSERTAÇÃO 25
2 FUNDAMENTAÇÃO TEÓRICA 28
2.1 SISTEMAS DE RECOMENDAÇÃO 28
2.1.1 Técnicas de Recomendação 29
2.2 SISTEMAS SENSÍVEIS AO CONTEXTO 30
2.3 EXTRAÇÃO DE INFORMAÇÃO 31
2.4 WEB SEMÂNTICA 32
2.4.1 Ontologias 33
2.4.2 Dados Ligados 34
2.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO 38
3 PROPOSTAS DE SISTEMAS DE RECOMENDAÇÃO 40
3.1 CONSIDERAÇÕES FINAIS DO CAPÍTULO 46
4 BROAD-RSI 48
4.1 ARQUITETURA 50
4.2 EXTRAÇÃO DE INFORMAÇÕES 52
4.2.1 Definição do Perfil Educacional 53
4.3 ENRIQUECIMENTO DE PERFIL 54
4.4 REPRESENTAÇÃO SEMÂNTICA 56
4.5 RECOMENDAÇÃO 57
4.5.1 Abordagem baseada em Repositórios de Objetos de Aprendizagem 57
4.5.2 Abordagem baseada em Dados Ligados 58
4.5.3 Recomendação de vídeos 60
4.5.4 Recomendação de pessoas 61
4.5.5 Armazenamento e entrega da recomendação aos usuários 63
4.6 CONSIDERAÇÕES FINAIS DO CAPÍTULO 64

5 DESENVOLVIMENTO DO PROTÓTIPO 66
5.1 EXTRAÇÃO DE INFORMAÇÕES 67
5.2 ENRIQUECIMENTO DE PERFIL 70
5.3 REPRESENTAÇÃO SEMÂNTICA 75
5.4 RECOMENDAÇÃO 76
5.4.1 Abordagem baseada em repositórios de objetos de aprendizagem 76
5.4.2 Abordagem baseada em Dados Ligados 80
5.4.3 Abordagem baseada em vídeos 82
5.4.4 Recomendações de pessoas 83
5.4.5 Armazenamento e entrega da recomendação aos usuários 84
5.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO 85
6 AVALIAÇÃO 87
6.1 PROVAS DE CONCEITO 87
6.1.1 Escolha dos participantes 88
6.1.2 Coleta dos dados 88
6.1.3 Primeira PoC: características de perfil e de contexto do usuário 89
6.1.4 Segunda PoC: enriquecimento de interesses educacionais extraídos 95
6.1.5 Terceira PoC: recomendações aderentes ao perfil e ao contexto do usuário 100
6.2 ESTUDO DE CASO 112
6.2.1 Planejamento do estudo de caso 113
6.2.2 Seleção dos indivíduos 114
6.2.3 Preparação e coleta dos dados 115
6.2.4 Métricas 117
6.2.5 Análise quantitativa dos dados 118
6.2.6 Análise qualitativa dos dados 125
6.2.7 Ameaças a validade 128
7 CONSIDERAÇÕES FINAIS 130
7.1 CONCLUSÕES 130
7.2 CONTRIBUIÇÕES 131
7.3 LIMITAÇÕES 132
7.4 TRABALHOS FUTUROS 133
REFERÊNCIAS BIBLIOGRÁFICAS 135

1 INTRODUÇÃO

Nos últimos anos foi possível notar uma grande diversificação na maneira que a população busca informações a respeito de diferentes temas e interesses. A facilidade de acesso aos dispositivos computacionais, tais como computadores, *tablets*, *smartphones* e, principalmente, o aumento do acesso à Internet trouxeram contínuas e rápidas transformações no comportamento da população em relação ao acesso, à busca e à disseminação da informação. Segundo Manuel Castells (CASTELLS, 2004), estamos em um processo de transformação estrutural multidimensional associado ao surgimento de um novo paradigma tecnológico, com base nas tecnologias de informação e comunicação.

Toda essa transformação pode ser notada também na área da Educação. As ferramentas computacionais estão cada vez mais presentes no processo de acesso à informação, seja através da realização de cursos formais, nas modalidades a distância ou semipresencial, apoiados por ferramentas tecnológicas; seja no dia-a-dia da população que usa diversas ferramentas computacionais na procura informal por conteúdos de interesse, usando blogs, canais de vídeos, cursos livres, redes sociais, entre outros ambientes. Uma pesquisa realizada pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação em 2013 (COMITÊ GESTOR DA INTERNET NO BRASIL – CGI.BR, 2014) mostrou que 96% dos professores de escolas públicas brasileiras usam recursos educacionais disponíveis na Internet para preparar aulas ou atividades com os alunos. Os tipos de recursos mais utilizados são imagens, figuras, ilustrações ou fotos (84%), textos (83%), vídeos (74%) e questões de prova (73%). O uso de jogos chega a 42%, apresentações prontas, 41%, e programas e softwares educacionais, 39%.

Existem inúmeros recursos educacionais distribuídos em diferentes repositórios que abordam um conjunto amplo de assuntos e que possuem objetivos educacionais distintos, tornando a escolha adequada dos recursos educacionais um desafio para os usuários que desejam usá-los para a sua formação intelectual. Em virtude disso, surgiram diversas ferramentas tecnológicas que contribuem com a localização e seleção dos recursos educacionais, tais como repositórios de Objetos de Aprendizagem (HODGINS, 2004) (LEHMAN, 2007) e Sistemas de Recomendação de recursos educacionais.

No entanto, para que essas ferramentas tecnológicas possam localizar recursos educacionais que atendam às necessidades do usuário é importante que elas possuam informações a respeito do seu perfil, seus interesses, preferências e o contexto no qual ele está inserido. Nesse sentido, é apresentada nesta dissertação uma proposta de extração de

características de perfil e de contexto dos usuários, através das suas interações em redes sociais e de uso de tais características para gerar recomendações personalizadas e individualizadas de recursos educacionais que sejam aderentes ao perfil e ao contexto de cada usuário.

Este capítulo fornece uma visão geral da dissertação, a motivação para o desenvolvimento da proposta apresentada, bem como a sua relação com o problema exposto. São apresentadas ainda as hipóteses desta pesquisa e a metodologia usada para testá-las.

1.1 MOTIVAÇÃO

Diante da diversidade de recursos educacionais disponíveis em diferentes fontes e da dificuldade de seleção desses recursos, de tal forma que eles atendam às necessidades específicas de cada usuário, um sistema de recomendação (SR) tem um papel importante para auxiliar os usuários a encontrarem recursos educacionais relevantes e pertinentes aos seus perfis e ao contexto em que estão inseridos. Segundo Burke (2002), um SR pode ser "qualquer sistema que produza recomendações individualizadas ou que tenha o efeito de guiar o usuário de forma personalizada para objetos do seu interesse ou que lhes sejam úteis dentre diversas opções possíveis".

Para que se seja possível gerar recomendações personalizadas, torna-se importante identificar informações que ajudem a definir o perfil do usuário e auxiliem na identificação de suas necessidades e interesses. O uso constante e cada vez mais intenso de algumas ferramentas tecnológicas faz com que inúmeras informações a respeito do perfil, dos interesses, das preferências, da forma de interação e do comportamento do usuário possam ser identificadas em decorrência da interação espontânea que ocorre nesses sistemas. No ambiente educacional isso ocorre com frequência em Ambientes Virtuais de Aprendizagem - AVAs (MCWHORTER; LINDHJEM, 2013), em sistema acadêmicos e em sistemas que apoiam cursos massivos – MOOCS (PAPPANO, 2012). Já em um contexto mais amplo o mesmo pode ser notado, por exemplo, em sites de redes sociais (BOYD; ELLISON, 2007), tais como Twitter[1], Linkedin[2], Facebook[3], Instagram[4], entre outros.

---

[1] https://twitter.com/?lang=pt

[2] https://br.linkedin.com/

[3] https://www.facebook.com/

[4] https://instagram.com/

Todos esses sistemas educacionais poderiam contribuir para a extração de características de perfil e de contexto dos usuários, no entanto esses sistemas são geralmente restritos a uma instituição de ensino ou são iniciativas específicas para um grupo de usuários. O grande alcance das redes sociais torna-se uma forte motivação para explorar o grande volume de dados ali gerados, as possibilidades de interação entre os usuários e as facilidades de disseminação da informação, que são próprias desses sites.

Recomendações feitas dentro de ambientes educacionais tradicionais, como os AVAs e MOOCs, possuem suas vantagens, como por exemplo, um foco maior no plano de estudo, maior precisão nos temas abordados, entre outros. Porém, geralmente, possuem alcance restrito a pessoas que buscaram formação intelectual através desses ambientes. A inclusão de ambientes com maior alcance, como as redes sociais, no processo de orientar e instruir pessoas, traz a possibilidade de atingir pessoas que não estejam inseridas em um ambiente formal de aprendizagem, mas que desejam buscar conhecimento através da Web.

Segundo Junco (2011), o site de rede social mais popular entre estudantes universitários é o Facebook. De acordo com uma pesquisa realizada pela *QS Quacquarelli Symonds Ltd* (VANOOZZI; BRIDGESTORK, 2013) entre 2012 e 2013, com 918 entrevistados em 26 países distribuídos pela Europa, Ásia, América Latina, América do Norte e África, o uso do Facebook entre estudantes varia entre 78 e 96% nessas regiões, sendo que a América Latina é a região que mais utiliza o Facebook. Segundo o Ibope (2011), no Brasil, 72% das pessoas que acessam a internet já incorporaram às suas rotinas o hábito de navegar em pelo menos uma rede social online.

O tempo dispendido pelos usuários nas interações em redes sociais é outro fator que motiva o uso desses ambientes associados a Sistemas de Recomendação. A mesma pesquisa realizada pela *QS Quacquarelli Symonds Ltd* mostrou que, cerca de 30% dos entrevistados, na América Latina e África, passam 5 ou mais horas nas redes sociais. Esse número se somado ao percentual de pessoas que passam entre 3 e 5 horas chega a 78% na América Latica e 54% na África. Esse número é um pouco menor em regiões como América do Norte, Europa e Ásia. No entanto o percentual de pessoas que usam as redes sociais por mais de 5 horas por dia permanece entre 24 e 26%. Se somados ao percentual dos que utilizam as redes sociais entre 3 e 5 horas os números chegam a 64% na América do Norte, 50% na Europa e Ásia.

Por tudo isso, torna-se muito motivador explorar o uso das informações geradas espontaneamente nas redes sociais para auxiliar o processo de busca e seleção de recursos educacionais de forma individualizada e personalizada, levando em consideração as

preferências, os interesses e outras características do perfil educacional do usuário. É motivador ainda, a possibilidade de usar as redes sociais para apresentar aos usuários instruções educacionais que contribuam com a sua formação intelectual, explorando o tempo que eles permanecem conectados a esses ambientes e também o poder de disseminação da informação que as redes sociais possuem, que pode contribuir para a ampla distribuição de recursos educacionais entre os seus usuários.

Uma motivação adicional para a realização deste trabalho é avançar as pesquisas relacionadas ao projeto BROAD (CAMPOS *et al*., 2012) (REZENDE *et al*., 2013), (PEREIRA *et al*., 2014a) (PEREIRA *et al*., 2014b) (PEREIRA *et al*., 2014c) acrescentando à sua proposta inicial, a definição do perfil e do contexto educacional dos usuários, através da aquisição de informações disponibilizadas por eles nas redes sociais e a recomendação de objetos de aprendizagem considerando tais características do usuário. A principal ideia do projeto BROAD é que cada nova versão da arquitetura represente um passo à frente em relação aos objetivos da pesquisa considerando a adoção de novas tecnologias.

O Projeto BROAD do Núcleo de Pesquisa em Engenharia do Conhecimento – NEnC, integra professores da Pós graduação em Ciência da Computação, alunos de mestrado e de iniciação científica e analistas de sistemas. Além disso, os professores integrantes do Projeto atuam no curso de Licenciatura em Computação, modalidade a distância, e tem nesse curso um ambiente interessante para a disseminação dos produtos gerados. A convergência dos resultados obtidos tem contribuído para a elaboração e composição de disciplinas a distância e ampliação do alcance da oferta da Educação.

No projeto BROAD, realizar recomendações de Objetos de Aprendizagem através de informações de contexto, metadados e planos didáticos é uma evolução natural da pesquisa. Considerando esse desafio na arquitetura apresentada(REZENDE *et al*., 2013), a recomendação qualificada dos OA é feita pelo professor à partir do seu plano didático e de um sistema multi-agentes, considerando as características do contexto do estudante.

A arquitetura do BROAD-RS está estruturada em quatro níveis:

- Multi-agent layer: esta camada descreve todos os agentes, seus respectivos papeis e suas inter-relações e interações com as aplicações. Os agentes relacionam os dados de contexto e perfil do aluno e plano didático do professor encontrados no banco de dados com a rede de ontologias.
- Middle layer: é a camada de serviços, utilizada para acessar as aplicações da arquitetura. É responsável pela interação dos agentes com a base de conhecimento.

- Backend layer: é a camada de conhecimento, onde se encontra a rede de ontologias e onde os OAs são armazenados. É nesta camada que a aplicação armazena os planos didáticos e as recomendações já realizadas em um banco de dados (BROAD-RS DB), interagindo diretamente com a base de conhecimento, ampliando a capacidade de expressão semântica das informações de contexto e perfil do aluno.
- Client layer: implementa o cadastro de objetos de aprendizagem, o cadastro de alunos e professores e a interface e telas dos usuários que pode ser desenvolvida em qualquer linguagem ou framework.

Este projeto pretende evoluir essa proposta, através do BROAD-RSI, acrescentando a definição do perfil e do contexto educacional e social dos alunos, através da aquisição semiautomática de informações disponibilizadas em redes sociais e ambientes virtuais de aprendizagem. A proposta prevê ainda a recomendação no contexto de redes sociais, levando orientações e recursos educacionais para os alunos fora de um ambiente acadêmico.

1.2 PROBLEMA

Conforme explicado anteriormente, para gerar recomendações de recursos educacionais de forma personalizada e individualizada, faz-se necessário que o sistema de recomendação possua informações acerca dos usuários, como por exemplo informações pessoais, interesses e preferências.

Muitos Sistemas de Recomendação utilizam formulários, nos quais o usuário informa alguns itens de sua preferência para, então, iniciar o processo de recomendação. A aplicação de formulário mostra-se como uma solução para esse problema. No entanto exige que o usuário execute uma ação manual e mantenha esse perfil sempre atualizado para sustentar o processo de recomendação.

No caso de Sistemas de Recomendação de recursos educacionais, outras alternativas usadas para obter informações a respeito do perfil educacional do usuário são a análise de logs do próprio sistema de recomendação, de ambientes virtuais de aprendizagem e de sistemas de controle acadêmico, o uso de interações passadas para buscar indícios de interesses e preferências atuais, mineração de dados educacionais, entre outros. Essas alternativas são eficazes, no entanto exigem que exista um grande volume de dados a serem analisados e, geralmente, tais dados somente são obtidos a partir do constante uso desses sistemas. Além

disso, algumas dessas alternativas exigem que os usuários estejam inseridos em um ambiente acadêmico, restringindo assim o alcance das recomendações educacionais.

Nesse contexto, surge o problema tratado neste trabalho: a necessidade de obter informações a respeito do perfil e do contexto de usuários, que não necessariamente estão inseridos em um ambiente acadêmico, sem que eles tenham que informá-las manualmente e atualizá-las com frequência, tornando possível oferecer-lhes recomendações de recursos educacionais de forma individualizada e personalizada.

1.3 HIPÓTESES

As hipóteses desta pesquisa estão relacionadas à possibilidade de extrair automaticamente características de perfil e de contexto do usuário e gerar recomendações individualizadas e personalizadas, aderentes a essas características:

- H1: Se extrairmos dados, gerados espontaneamente, a partir da interação dos usuários em redes sociais, então será possível identificar características de perfil e de contexto do usuário.
- H2: Se usarmos as características de perfil e de contexto extraídas da interação do usuário em redes sociais, então será possível realizar recomendações educacionais relevantes a esses usuários.

1.4 METODOLOGIA DE PESQUISA

A metodologia de pesquisa usada neste trabalho inclui: (1) a revisão bibliográfica; (2) o mapeamento sistemático (KITCHENHAM, 2007); (3) a proposta e o desenvolvimento de uma arquitetura capaz de extrair características do perfil e do contexto dos usuários, através de redes sociais e realizar recomendações de recursos educacionais de forma individualizada e personalizada que sejam condizentes com essas características; e (4) as avaliações realizadas da proposta, feitas através do desenvolvimento de um protótipo, de provas de conceito (DE MELLO *et al.*, 2012) e de um estudo de caso (WOHLIN *et al.*, 2012).

Na revisão bibliográfica foi possível obter o conhecimento atualizado ("estado da arte") nas áreas de Extração de Informação, Web Semântica e Sistemas de Recomendação. Todos os trabalhos relacionados e citados nesta pesquisa foram catalogados usando o software

Mendeley[5]. O uso desse software permitiu buscas mais eficientes por palavras-chave, autores e *tags*, além de auxiliar na construção da bibliografia presente no final desta pesquisa.

A realização do mapeamento sistemático (PEREIRA; *et al.*, 2014) teve como objetivo buscar Sistemas de Recomendação sensíveis ao contexto na área educacional e, então, identificar os elementos de contexto usados nesses sistemas. O resultado do mapeamento sistemático ajudou na escolha dos elementos contextuais a serem considerados na proposta de recomendação de recursos educacionais apresentada nesta pesquisa.

Após a realização da revisão bibliográfica e do mapeamento sistemático, foi feita a proposta de uma arquitetura capaz de extrair características do perfil e do contexto dos usuários através da rede social Facebook e realizar recomendações de recursos educacionais de forma individualizada e personalizada que sejam condizentes com essas características. A solução proposta é apoiada por técnicas de Extração de Informações e ontologias para a extração e definição de características e interesses dos usuários. Já a recomendação é feita através de três abordagens: baseada em repositórios de Objetos de Aprendizagem (HODGINS, 2004), baseada em repositórios de Dados Ligados (BIZER *et al.*, 2009) e baseada em um repositório de vídeos.

A validação da proposta foi feita através do desenvolvimento de um protótipo, cujos objetivos são: (1) extrair características de perfil e de contexto do usuário, através da rede social Facebook; (2) enriquecer interesses educacionais extraídos; (3) exibir aos usuários recomendações educacionais aderentes ao seu perfil e ao contexto; (4) apoiar o processo de avaliação da proposta por parte do usuário, ou seja, possibilitar que os usuários avaliem as características de perfil e de contexto extraídas e as recomendações feitas dentro do próprio protótipo. Após o desenvolvimento do protótipo, foram realizadas três provas de conceitos (*Proof of Concept* – PoC) (DE MELLO *et al.*, 2012) e, posteriormente, um estudo de caso (WOHLIN *et al.*, 2012) também foi conduzido.

Através da realização das três provas de conceito foi possível verificar a viabilidade técnica da proposta e ter os primeiros indícios acerca das hipóteses que estão sendo avaliadas nesta dissertação. Além disso, elas auxiliaram a formulação do questionário usado como instrumento de avaliação no estudo de caso. A partir do uso do protótipo, durante a realização das provas de conceito, foi possível ajustar alguns parâmetros para a avaliação, solucionar problemas técnicos e perceber algumas questões que poderiam ser melhor avaliadas em um estudo de caso.

[5] http://www.mendeley.com/

Assim sendo, os resultados observados nas três provas de conceito apresentadas motivaram o planejamento e execução de um estudo de caso visando aumentar a capacidade de observação e a confiança na viabilidade da proposta do BROAD-RSI.

1.5 OBJETIVOS

Esta pesquisa tem o objetivo principal de propor a arquitetura BROAD-RSI, capaz de extrair e explorar as informações disponíveis nas redes sociais para identificar características de perfil e de contexto dos usuários e, a partir dessas características, fazer recomendações de recursos educacionais individualizadas e personalizadas aos usuários.

O objetivo principal desta pesquisa pode ser decomposto nos seguintes objetivos específicos:

i. identificação de características do perfil e do contexto do usuário a partir de informações geradas, espontaneamente, nas redes sociais;
ii. uso do perfil do contexto do usuário para auxiliar no processo de seleção e recomendação de recursos educacionais adequados ao usuário;
iii. oferecimento de recomendações educacionais a usuários que não são alunos regulares de nenhum curso ou instituição, no entanto possuem interesse em se capacitar;
iv. oferecimento de recomendações educacionais dentro dos ambientes das redes sociais;
v. indicação de pessoas com interesses educacionais semelhantes, buscando aproximá-las na intenção de formarem uma rede de interesses em torno de um determinado assunto;
vi. avanço do projeto BROAD (REZENDE *et al.*, 2013) acrescentando a definição do perfil e do contexto dos usuários, através da aquisição de informações disponibilizadas por eles nas redes sociais.

1.6 ORGANIZAÇÃO DA DISSERTAÇÃO

Este capítulo apresentou uma introdução, a motivação que orienta a busca da solução do problema a ser tratado, as hipóteses consideradas, a metodologia, os objetivos e a organização da dissertação. No capítulo 2 são apresentados os pressupostos teóricos, que dão o embasamento teórico necessário para o entendimento da solução proposta. No capítulo 3 é

realizado um levantamento a respeito de diferentes Sistemas de Recomendação de recursos educacionais e é feito uma comparação entre eles. No capítulo 4 é proposta a arquitetura BROAD-RSI capaz de extrair e explorar as informações disponíveis na rede social Facebook para identificar características de perfil e de contexto dos usuários e, a partir dessas características, fazer recomendações de recursos educacionais individualizadas e personalizadas aos usuários. No capítulo 5 é mostrada a implementação da proposta do BROAD-RSI. No capítulo 6 é feita a avaliação da solução apresentada e, por fim, são feitas as considerações finais no capítulo 7.

A figura 1 é uma síntese da organização desta dissertação.

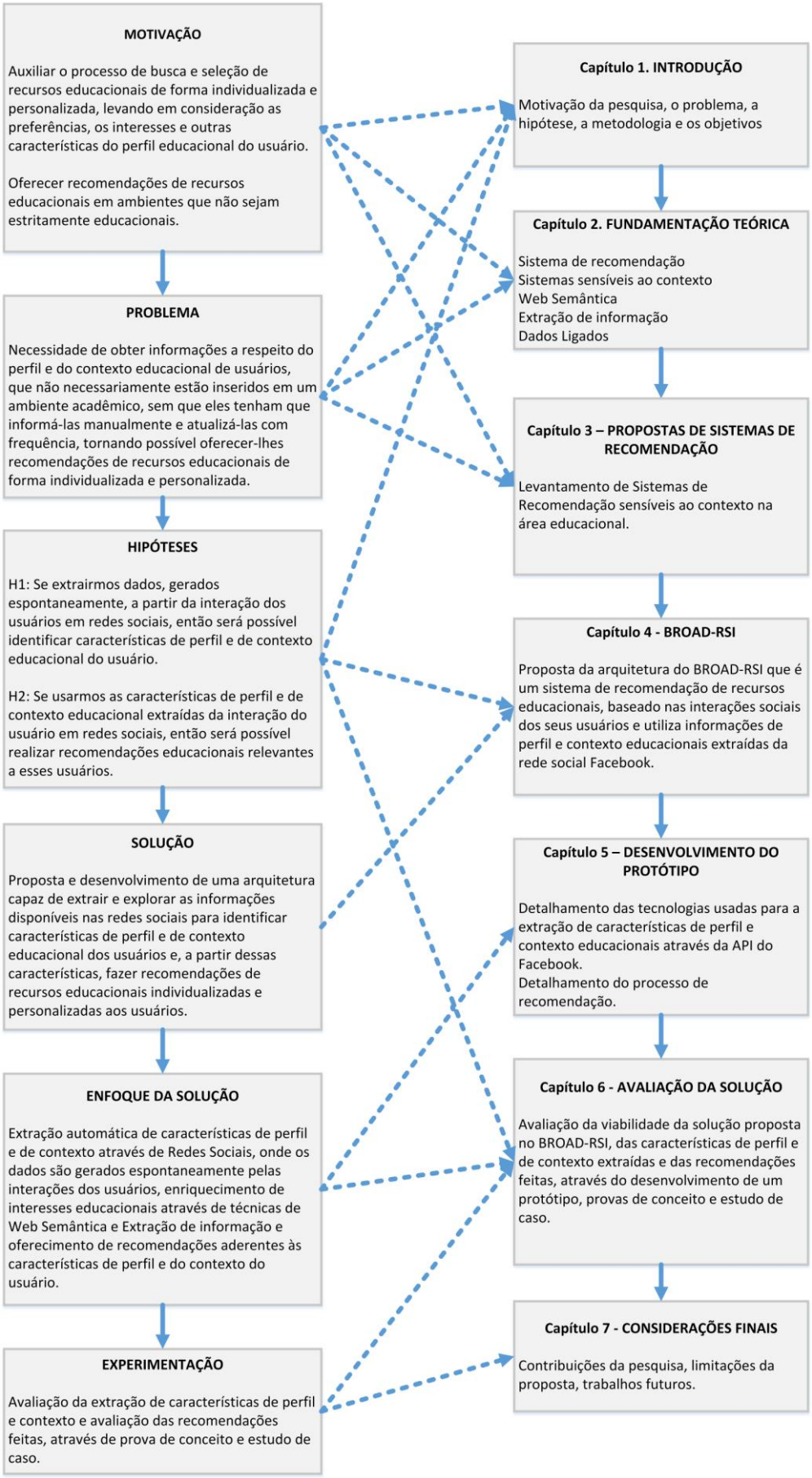


Figura 1 - Organização da dissertação.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os pressupostos teóricos que dão o embasamento teórico necessário para a solução proposta. Os conceitos relacionados à Sistemas de Recomendação e técnicas de recomendação são fundamentais para entender como são feitas as recomendações de recursos educacionais e quais foram as escolhas feitas na formulação da proposta. Os conceitos relacionados a sistemas sensíveis ao contexto são úteis para elucidar as características desses sistemas, bem como a definição de contexto, já que a proposta apresentada nesta dissertação considera alguns elementos de contexto para apoiar a escolha de recursos educacionais. Os conceitos de Extração de Informação são essenciais para entender as atividades aplicadas na solução proposta para extrair as entidades nomeadas, conceitos e palavras-chave dos textos presentes nas páginas compartilhadas pelos usuários. Já os conceitos relacionados a Web Semântica são necessários para a compreensão tanto da representação semântica do perfil do usuário, quanto para o enriquecimento de áreas de interesses, feitas através de ontologias. Já os conceitos relacionados a dados ligados são úteis para a melhor compreensão de um dos conjuntos de recursos educacionais usados na recomendação.

2.1 SISTEMAS DE RECOMENDAÇÃO

Segundo Burke (2002), um sistema de recomendação pode ser: “qualquer sistema que produza recomendações individualizadas ou que tenha o efeito de guiar o usuário de forma personalizada para objetos do seu interesse ou que lhes sejam úteis dentre diversas opções possíveis”. Esse tipo de sistema, basicamente, precisa de três componentes para fornecer recomendações: (1) os dados de contexto, ou seja, as informações que o sistema tem antes do processo de recomendação começar; (2) dados de entrada, que são as informações fornecidas sobre o usuário, a fim de fazer uma recomendação, e (3) o algoritmo usado na recomendação que opera com dados de contexto e de entrada, a fim de proporcionar recomendações para um usuário.

De acordo com COSTA, AGUIAR & MAGALHÃES (2013) os Sistemas de Recomendação servem na área educacional, por exemplo, para filtrar recursos educacionais contemplando os perfis individuais dos alunos ou até mesmo perfis de um grupo de alunos de uma disciplina. Dessa forma, esses sistemas podem auxiliar o professor na composição de recursos digitais para uma disciplina e no acompanhamento dos alunos, uma vez que a

preocupação em identificar e tratar perfis diferenciados é parcialmente ou até mesmo completamente resolvida com a personalização provida por esses sistemas.

Alguns outros exemplos de tarefas educacionais apoiadas por Sistemas de Recomendação são (MANOUSELIS *et al*., 2011): recomendar itens que ajudem o aluno durante a realização de uma tarefa (FRITZEN *et al*., 2012) (KLASNJA-MILICEVIC *et al*., 2011); descobrir e indicar novos itens de interesse (REZENDE *et al*., 2013); adaptar conteúdo e interface para o usuário (GASPARINI *et al*., 2011); recomendar itens em uma sequência relevante (KARAMPIPERIS; SAMPSON, 2005); apresentar recursos educacionais ao usuário, enquanto ele navega em outros sites (FRITZEN *et al*., 2012) (ALHARBI *et al*., 2012) (WANG; WU, 2011); encontrar bons pares para desenvolvimento de trabalhos (DWIVEDI; BHARADWAJ, 2012) (RECKER; WILEY, 2001) (GREER et al., 1998); recomendar pessoas com interesses, preferências e estilos de aprendizagem semelhantes (DWIVEDI; BHARADWAJ, 2012); entre outros.

2.1.1 **Técnicas de Recomendação**

Para realizar as recomendações, os Sistemas de Recomendação utilizam algumas técnicas que permitem a escolha do conteúdo mais apropriado a um determinado usuário. O modelo de BURKE (2002) categoriza as técnicas de recomendação em cinco conjuntos: filtragem colaborativa (FC); baseada em conteúdo (FBC); demográfico; baseada na utilidade e baseada no conhecimento. Segundo ADOMAVICIUS & TUZHILIN (2005), as técnicas de filtragem colaborativa e baseada em conteúdo podem ser usadas de forma combinada, gerando assim, a técnica de recomendação híbrida.

Na técnica de filtragem colaborativa, os itens são selecionados para um usuário tendo como base a avaliação feita por outros usuários com perfis (preferências, interesses) similares. Já a baseada em conteúdo, seleciona os itens para a recomendação comparando os seus conteúdos com o de outros itens dos quais o usuário tenha gostado no passado. Na técnica demográfica, o conteúdo é recomendado para um usuário baseado em outros usuários que possuem as mesmas informações demográficas. A técnica baseada na utilidade gera o conjunto de itens baseando-se na utilidade que estes terão aos usuários e, por fim, a técnica baseada em conhecimento considera o conhecimento de como esses itens podem atender às necessidades e preferências do usuário.

Todas as técnicas de recomendação possuem vantagens e desvantagens. A principal vantagem da filtragem baseada em conteúdo é que através dela é possível recomendar novos

itens a usuários, mesmo que esses itens não tenham recebido avaliação de nenhum outro usuário. No entanto, essa técnica impede recomendação a usuários que não tenham acessado nenhum outro material no passado, já que ela realiza comparações entre itens novos e itens que o usuário já tenha acessado no passado.

A filtragem colaborativa possui a vantagem de não precisar de histórico das recomendações feitas ao usuário para realizar as primeiras recomendações a eles. Porém, nesta técnica há o problema do *cold-start* (PARK; CHU, 2009), onde não há classificações históricas sobre um determinado item feita por outros usuários. De forma geral, a fase inicial do processo de recomendação, onde os dados disponíveis são extremamente escassos, é um desafio em sistemas que utilizam filtragem baseada em conteúdo ou filtragem colaborativa, já que ambas são baseadas em aprendizagem, ou seja, necessitam que o sistema esteja em uso para realizarem novas recomendações de qualidade.

No caso das técnicas de localização, conhecimento e utilidade há a vantagem de não ser necessário ter informações sobre as recomendações feitas no passado, nem das avaliações de outros usuários aos itens disponíveis para serem recomendados, como acontece com as técnicas de filtragem colaborativa e baseada em conhecimento. No entanto, elas requerem conhecimento prévio sobre o perfil e o contexto dos usuários, seus interesses e preferências para realizar as recomendações. PARK & CHU (2009) indicam que uma maneira de amenizar as dificuldades em determinar um perfil inicial para o usuário é solicitar informações básicas a ele, através de questionários ou de outros sistemas que ele já utilize, permitindo assim o envio das primeiras recomendações.

2.2 SISTEMAS SENSÍVEIS AO CONTEXTO

Um sistema é considerado sensível ao contexto se ele usa o contexto para fornecer informações e/ou serviços relevantes ao usuário, sendo que a relevância depende da tarefa do usuário (DEY & ABOWD, 1999). Os sistemas sensíveis ao contexto buscam adaptar suas funcionalidades e o seu comportamento de acordo com o contexto em que seus usuários estão inseridos.

Em uma pesquisa realizada por BAZIRE & BRÉZILLON (2005) foram detectadas mais de 150 definições de contexto, provenientes principalmente da Web, evidenciando que não há um consenso em relação a esse conceito. Uma das definições de contexto mais citadas e usadas é a definição de DEY & ABOWD (1999): “contexto é qualquer informação que pode ser usada para caracterizar a situação de uma entidade. Uma entidade é uma pessoa, lugar ou

objeto que é considerada relevante para a interação entre o usuário e uma aplicação, incluindo o próprio usuário e a aplicação". Em outras palavras, VIEIRA & TEDESCO (2009) explicam que contexto é o que está por trás da habilidade de discriminar o que é ou não importante em um dado momento.

Ainda a respeito da definição de contexto, VIEIRA & TEDESCO (2009) fazem uma distinção entre os conceitos de contexto e elementos contextuais: um elemento contextual é qualquer dado, informação ou conhecimento que permite caracterizar uma entidade ou domínio; o contexto da interação entre um agente e uma aplicação para executar alguma tarefa, é o conjunto de elementos contextuais instanciados que são necessários para apoiar a tarefa atual.

Apesar da diversidade de definições de contexto, duas conclusões consensuais puderam ser extraídas do trabalho de BAZIRE & BRÉZILLON (2005) (apud Vieira *et al*, 2006): (1) o contexto atua como um conjunto de restrições que influenciam o comportamento de um sistema embutido em uma dada tarefa; (2) a definição de contexto depende da área de conhecimento.

Na área educacional, existem inúmeras propostas (SUDHANA *et al*., 2013) (ZHENG *et al*., 2009) (GALLEGO *et al*, 2013) (VERBERT *et al*., 2012) que buscam considerar diferentes elementos contextuais para auxiliar a execução de tarefas tais como: apoiar os alunos a atingirem metas específicas de aprendizagem; recomendar recursos educacionais; adaptar sistemas de apoio à aprendizagem, determinar o melhor momento e local para envio de orientações educacionais, entre outras.

2.3 EXTRAÇÃO DE INFORMAÇÃO

Segundo PISKORSKI & YANGARBER (2013) o objetivo da Extração de Informação é identificar um conjunto predefinido de conceitos em um domínio específico, ignorando outras informações irrelevantes, ou seja, derivar informações estruturadas a partir de texto não estruturado. De acordo com JIANG (2012), a Extração de Informação é uma tarefa importante na área de mineração de texto e tem sido extensivamente estudada em várias outras comunidades de pesquisa, incluindo o Processamento de Linguagem Natural, Recuperação de Informações e Mineração Web. Dentre as informações estruturadas pode-se destacar a extração automática de entidades, relações entre entidades e atributos (SARAWAGI, 2008).

As principais características que podem ser encontradas em uma técnica de Extração

de Informação são (MOENS, 2006) (apud FERNANDES, 2013): (i) necessidade de informação a ser obtida de uma fonte de dados normalmente composta por conteúdo textual ou imagens; (ii) impossibilidade de obter a resposta por humanos devido ao grande volume de dados; (iii) inexistência de estrutura formal nos dados como a obtida em um banco de dados, por exemplo, o que não permite uma consulta direta feita por computadores.

Algumas tarefas comuns de Extração de Informação são (MOENS, 2006):

- Reconhecimento de entidades: reconhecimento e classificação de expressões próprias em texto, tais quais pessoas, empresas ou localidades;
- Correferência de frases: reconhecimento de frases que façam referência à mesma situação ou entidade em um texto;
- Reconhecimento de papel semântico: atribuição de papéis semânticos a elementos sintáticos que constituem uma sentença;
- Relação entre entidades: reconhecimento do relacionamento entre duas ou mais entidades, sendo atribuído a este relacionamento um papel semântico;
- Reconhecimento temporal: detecção e reconhecimento no texto de expressões temporais.

A Extração de Informação tem uma grande variedade de uso, em diferentes domínios. Entre elas podemos citar (JIANG, 2012): (1) área médica, onde as técnicas são aplicadas para descobrir relações entre genes específicos, proteínas e outras entidades médicas em grandes bases de artigos científicos; (2) profissionais da área financeira, que muitas vezes precisam buscar assuntos específicos a partir de notícias para auxiliar a sua tomada de decisão; (3) serviços de inteligência, que analisam grande quantidade de textos a procura de informações a respeito de terrorismo, por exemplo; (4) motores de busca na Web, que usam a Extração de Informação como uma das etapas no processo de recuperação de itens na Web.

2.4 WEB SEMÂNTICA

Os primeiros conceitos de Web Semântica foram apresentados por Tim Berners-Lee em 2001: “a Web Semântica não é uma Web separada, mas sim uma extensão da atual, na qual informações ganham um significado bem definido, permitindo que computadores e pessoas trabalhem melhor em cooperação” (BERNERS-LEE *et al*., 2001).

De acordo com ANTONIOU & HARMELEN (2008), existem duas interpretações a respeito dos objetivos da Web Semântica. Na primeira interpretação o objetivo da Web Semântica é permitir a integração de fontes de dados estruturados e semiestruturados através

da Web. Para isso, seria necessário expor conjuntos de dados na Web em formato RDF (*Resource Description Framework*) e usar RDF *Schema* para expressar a semântica deste conjunto de dados, a fim de permitir a integração e reutilização para fins específicos. Na segunda interpretação, o objetivo da Web Semântica é melhorar a Web atual, como, por exemplo, melhorando motores de busca, personalização dinâmica de sites e enriquecimento semântico das páginas da Web existentes. Nesse caso, estão envolvidos metadados semânticos que possibilitam, por exemplo a extração de conceito, reconhecimento da entidades, classificação automática e outros. ANTONIOU & HARMELEN (2008) enfatizam que, obviamente, há sobreposição das duas interpretações, apesar dos objetivos, tecnologias e fontes semânticas serem distintas.

O uso de técnicas e padrões da Web Semântica na área educacional tem trazido avanços em alguns aspectos, como por exemplo, personalização de conteúdo e recursos educacionais para os estudantes, recomendação de recursos educacionais, coleta de dados referentes a interação dos alunos com os ambientes educacionais (DEVEDZ, 2004; OHLER, 2008).

BITTENCOURT *et al*. (2008) apontam que, na área da educação, as pesquisas em Web Semântica estão voltadas para os seguintes tópicos: construção de ontologias para *e-learning*, uso de ontologias e padrões da Web Semântica para estruturação, representação, indexação e recuperação compartilhável e interoperabilidade de recursos educacionais, uso de ontologias e padrões de Web Semântica para apoio de sistemas *e-learning* inteligentes, uso de contexto baseado em Web Semântica para adaptação, personalização de aplicações *e-learning*.

2.4.1 **Ontologias**

Nos últimos anos, ontologias vêm sendo muito usadas na área da Ciência da Computação, no entanto não se restringe a essa área, sendo também muito explorada na área da medicina, biologia, filosofia e gerenciamento do conhecimento. Sendo assim, muitas definições distintas são encontradas, estando cada uma associada a áreas específicas. Algumas delas são exploradas em (GUARINO; GIARETTA, 1995).

Na área da computação, um conceito muito encontrado é: “uma Ontologia pode ser definida como uma especificação formal e explícita de uma conceituação compartilhada” (GRUBER, 1995), onde conceituação refere-se a abstração de um domínio que identifica os conceitos relevantes desse domínio e compartilhada diz respeito ao conhecimento consensual

representado por uma ontologia. Algumas outras definições de ontologia no campo da computação, bem como uma ampla discussão sobre elas, podem ser encontradas em (GUARINO, 1988).

Atualmente, as tecnologias mais importantes da ontologia para a Web Semântica são as seguintes (YU, 2011):

- *Resource Description Framework*[6] (RDF): é um modelo de dados para objetos ("recursos") e relacionamento entre eles, que prevê uma semântica simples para este modelo de dados; estes modelos de dados podem ser representados em uma sintaxe XML.
- RDF *Schema* é uma linguagem de descrição de vocabulário para descrever propriedades e classes de recursos RDF, com uma semântica para hierarquias de generalização de tais propriedades e classes.
- *Ontology Web Language*[7] (OWL): é uma linguagem de descrição de vocabulário mais rico para descrever propriedades e classes, tais como as relações entre as classes (por exemplo, disjunção), cardinalidade (por exemplo, "exatamente um"), igualdade, definição de tipo de dados de propriedades mais rica, características das propriedades (por exemplo, a simetria), e classes enumeradas.

No contexto educacional, ontologias podem ser aplicadas para resolver uma variedade de problemas complexos, como a representação do conhecimento em sistemas inteligentes, o compartilhamento de conhecimentos, reuso de aplicações, anotação e busca de objetos de aprendizagem, personalização de conteúdo de aprendizagem, entre tantos outros desafios (BITTENCOURT *et al*., 2008).

2.4.2 **Dados Ligados**

O termo Dados Ligados ou *Linked Data* pode ser sintetizado como o uso da Web para criar ligações entre os dados originários de diferentes fontes, sendo que esses dados podem estar armazenados em diversos bancos de dados, mantidos por diferentes organizações

[6] Resource Description Framework (RDF), recomendação W3C de 22 Fevereiro de 1999. http://www.w3.org/TR/1999/REC-rdf-syntax-19990222

[7] Ontology Web Language (OWL), recomendação W3C de 10 de fevereiro de 2004. http://www.w3.org/TR/owl-features

e distribuídos em diferentes localizações geográficas (BIZER *et al*., 2009).

Um dos principais objetivos dos Dados Ligados é estender a Web que conhecemos, onde documentos HTML estão interconectados, através de hiperlinks, para uma Web onde os dados possam estar diretamente ligados, sem necessidade da intervenção de alguma aplicação que faça essa ligação. Esta extensão da Web também é conhecida por Web de Dados (BIZER *et al*., 2009).

Em 2006, Tim Berners-Lee esboçou os princípios dos Dados Ligados (BERNERS-LEE, 2006): (1) usar URIs como nomes para as coisas; (2) usar URIs HTTP para que as pessoas possam buscar por esses nomes; (3) quando alguém pesquisar por uma URI, prover informação útil, utilizando os padrões recomendados (RDF, SPARQL); (4) incluir links para outras URIs, para que mais coisas possam ser descobertas.

O RDF é o padrão básico da proposta de Dados Ligados, através do qual os recursos são descritos, por meio de uma ou mais assertivas sobre eles; conteúdos são publicados de forma estruturada, enquanto uns e outros são interconectados através de links, diferentes dos convencionais. Esses links expressam a natureza ou "semântica" da ligação entre os recursos que estão sendo ligados (BIZER *et al*., 2009). A SPARQL[8] (*Simple Protocol and RDF Query Language*) é a linguagem de consulta para RDF e, geralmente, para os Dados Ligados. Ela é projetada para consultar as representações de dados baseados em grafos. As consultas podem ser realizadas em SPARQL *Endpoints* que são serviços REST (CORPORATION, 2013) que recebem as consultas e retornam seus resultados.

O maior projeto de utilização de Dados Ligados é o *Linking Open Data* (LOD), cuja responsabilidade é da W3C (*Word Wide Web Consortium*). Esse projeto tem o objetivo de capturar dados, que estão sendo mantidos no modelo da web convencional e que possuem licenças públicas, de diferentes conjuntos de dados (*datasets*), e convertê-los para o padrão de dados ligados, contribuindo assim para o enriquecimento semântico desses dados e na rede de dados como um todo. A figura 2 mostra os conjuntos de dados que foram publicados em formato de Dados Ligados, até agosto de 2014, por membros da comunidade do projeto *Linking Open Data* e por outros indivíduos e organizações.

Outro exemplo de projeto usando dados ligados é a DBPEDIA (AUER *et al*., 2007). A DBPEDIA é um esforço da comunidade para extrair informações estruturadas da enciclopédia virtual Wikipedia[9] e disponibilizá-las na Web. A DBPEDIA permite fazer

[8] http://www.w3.org/TR/rdf-sparql-query/

[9] https://www.wikipedia.org/

diversas consultas em conjuntos de dados derivados da Wikipédia e vincular outros conjuntos de dados com os dados da Wikipedia. A figura 3 mostra uma página da DBPEDIA com propriedades e valores relacionados ao tema “Sistema de recomendação”, resultante da seguinte página http://en.wikipedia.org/wiki/Recommender_system da Wikipédia.

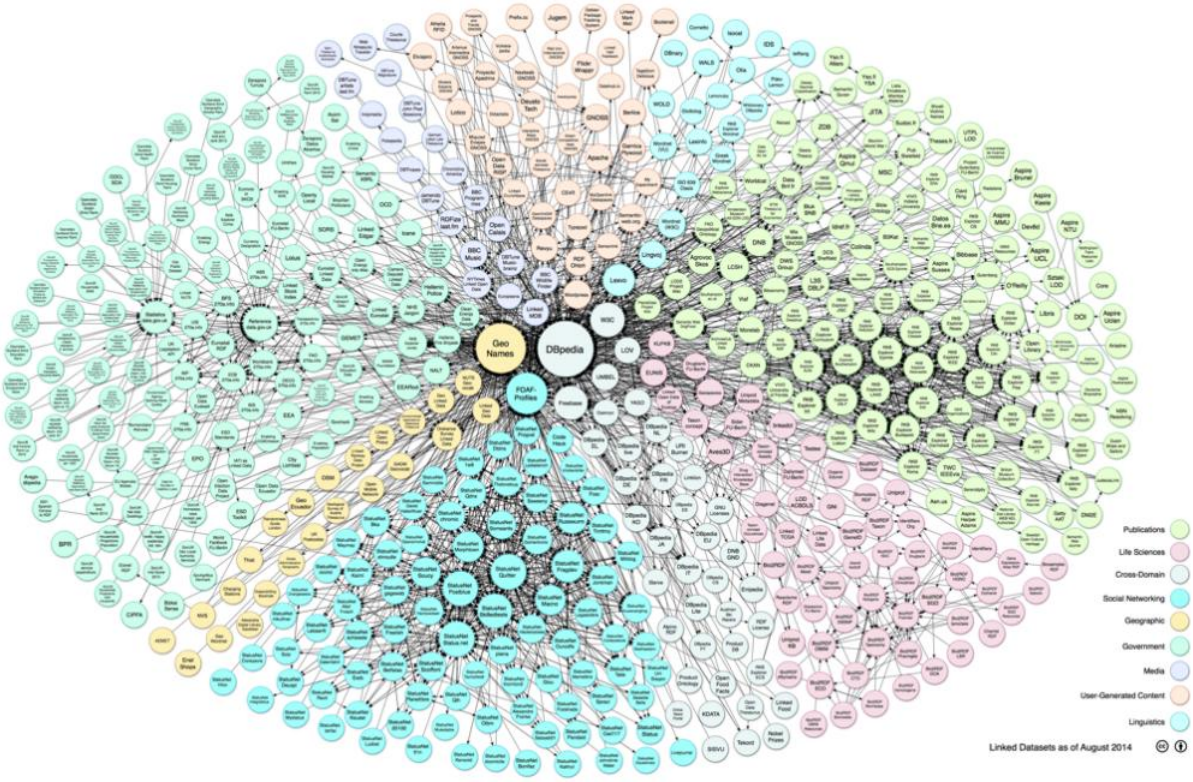
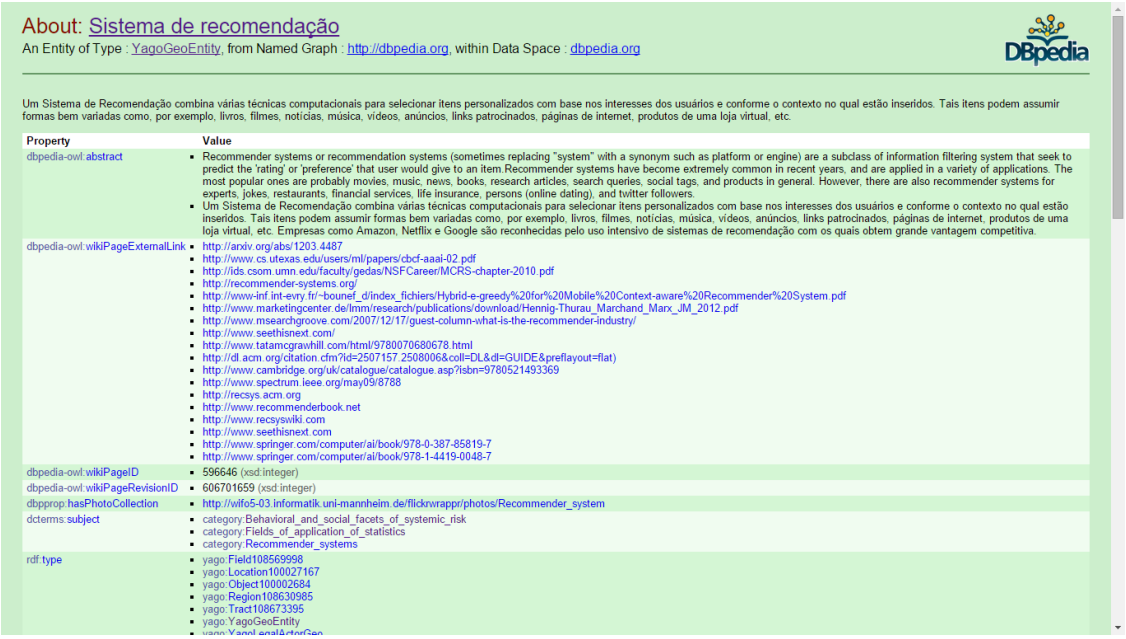


Figura 2 - Visão geral da rede de dados do *The Linking Open Data*. Fonte: CYGANIAK; JENTZSCH (2014).



About: Sistema de recomendação

An Entity of Type : YagoGeoEntity, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Um Sistema de Recomendação combina várias técnicas computacionais para selecionar itens personalizados com base nos interesses dos usuários e conforme o contexto no qual estão inseridos. Tais itens podem assumir formas bem variadas como, por exemplo, livros, filmes, notícias, música, vídeos, anúncios, links patrocinados, páginas de internet, produtos de uma loja virtual, etc.

Property	Value
dbpedia-owl:abstract	▪ Recommender systems or recommendation systems (sometimes replacing "system" with a synonym such as platform or engine) are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that user would give to an item.Recommender systems have become extremely common in recent years, and are applied in a variety of applications. The most popular ones are probably movies, music, news, books, research articles, search queries, social tags, and products in general. However, there are also recommender systems for experts, jokes, restaurants, financial services, life insurance, persons (online dating), and twitter followers. ▪ Um Sistema de Recomendação combina várias técnicas computacionais para selecionar itens personalizados com base nos interesses dos usuários e conforme o contexto no qual estão inseridos. Tais itens podem assumir formas bem variadas como, por exemplo, livros, filmes, notícias, música, vídeos, anúncios, links patrocinados, páginas de internet, produtos de uma loja virtual, etc. Empresas como Amazon, Netflix e Google são reconhecidas pelo uso intensivo de sistemas de recomendação com os quais obtem grande vantagem competitiva.
dbpedia-owl:wikiPageExternalLink	▪ http://arxiv.org/abs/1203.4487 ▪ http://www.cs.utexas.edu/users/ml/papers/cbcf-aaai-02.pdf ▪ http://ids.csom.umn.edu/faculty/gedas/NSFCareer/MCRS-chapter-2010.pdf ▪ http://recommender-systems.org/ ▪ http://www-inf.int-evry.fr/~bounef_d/index_fichiers/Hybrid-e-greedy%20for%20Mobile%20Context-aware%20Recommender%20System.pdf ▪ http://www.marketingcenter.de/lmm/research/publications/download/Hennig-Thurau_Marchand_Marx_JM_2012.pdf ▪ http://www.msearchgroove.com/2007/12/17/guest-column-what-is-the-recommender-industry/ ▪ http://www.seethisnext.com/ ▪ http://www.tatamcgrawhill.com/html/9780070680678.html ▪ http://dl.acm.org/citation.cfm?id=2507157.2508006&coll=DL&dl=GUIDE&preflayout=flat) ▪ http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=9780521493369 ▪ http://www.spectrum.ieee.org/may09/8788 ▪ http://recsys.acm.org ▪ http://www.recommenderbook.net ▪ http://www.recsyswiki.com ▪ http://www.seethisnext.com ▪ http://www.springer.com/computer/ai/book/978-0-387-85819-7 ▪ http://www.springer.com/computer/ai/book/978-1-4419-0048-7
dbpedia-owl:wikiPageID	▪ 596646 (xsd:integer)
dbpedia-owl:wikiPageRevisionID	▪ 606701659 (xsd:integer)
dbpprop:hasPhotoCollection	▪ http://wifo5-03.informatik.uni-mannheim.de/flickrwrappr/photos/Recommender_system
dcterms:subject	▪ category:Behavioral_and_social_facets_of_systemic_risk ▪ category:Fields_of_application_of_statistics ▪ category:Recommender_systems
rdf:type	▪ yago:Field108569998 ▪ yago:Location100027167 ▪ yago:Object100002684 ▪ yago:Region108630985 ▪ yago:Tract108673395 ▪ yago:YagoGeoEntity

Figura 3 - Página da DBPEDIA relacionada ao tema "Sistema de Recomendação".

O conceito de dados ligados vem sendo explorado na área educacional em alguns projetos. AQUIN (2012) explica que há potencial para o uso de dados ligados na educação em virtude da natureza aberta e acessível dos recursos educacionais produzidos por muitas universidades:

> "...enquanto o conteúdo de recursos educacionais abertos é, por definição, acessível e reutilizável, publicando os metadados correspondentes a eles como *Linked Data* pode tornar o conteúdo de diferentes repositórios mais detectável, acessível e conectável. O uso de vocabulários comuns para descrever os temas tratados é uma forma de tornar o conteúdo destes repositórios endereçável a nível global, e para recuperar os recursos em primeiro lugar pela sua relevância, sem ter que considerar a sua origem. " (AQUIN, 2012).

Além da publicação é importante ressaltar o poder de utilização de dados disponíveis através de dados ligados em sistemas educacionais. O aplicativo Talis Aspire Community Edition (HEATH *et al*., 2012) é um dos Sistemas de Recomendação que usam a abordagem de Dados Ligados. Ele permite que professores de universidades do Reino Unido criem e gerenciem listas de livros on-line para os cursos que lecionam. Estas listas são organizadas de tal forma que os metadados usados são padronizados e disponibilizados como Dados Ligados e os itens semelhantes são ligados uns aos outros. Feita essa organização e representação, o aplicativo consegue buscar itens semelhantes, a partir de propriedades de ligação dos dados, fornecendo a recomendação de itens semelhantes aos que estão sendo

consultados. Um exemplo de aplicativo, disponível no Facebook, que se beneficia das conexões possíveis com Dados Ligados é "*Course Profile*"[10] da *Open University*. Ele usa o catálogo de cursos da *Open University* como fonte de informações, catálogo esse que foi armazenado usando a abordagem de Dados Ligados e permite que um usuário se conecte a outros usuários, indicando se eles já cursaram, estão cursando ou têm interesse em cursar um determinado curso. A partir da indicação de pessoas torna-se possível também sugerir outras atividades ou cursos ligados.

2.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO

No contexto deste trabalho foi adotado o conceito de BURKE (2002) para Sistemas de Recomendação, sendo esse aplicado ao domínio específico da Educação. Na proposta detalhada nesta dissertação o Sistema de Recomendação possui também as características de um sistema sensível ao contexto, ou seja, que usa o contexto do usuário para fornecer informações e/ou serviços relevantes ao usuário, sendo, no contexto deste trabalho, oferecidas aos usuários algumas recomendações educacionais de forma individualizada e personalizada.

Os dados de entrada do sistema de recomendação foram selecionados após a realização de um levantamento de quais elementos contextuais podem ser relevantes em sistemas sensíveis ao contexto que possuem a finalidade de fazer recomendações educacionais aos seus usuários, são eles:

- dados de perfil: nome, idade, histórico educacional, tipo de mídia mais compartilhada pelo usuário, gênero e idiomas de domínio.
- elementos contextuais: interesses dos usuários, dia e horário de maior acesso, dispositivo usado e as conexões de um usuário com outros usuários.

Em relação às técnicas de recomendação foi escolhida a técnica baseada em conhecimento, buscando-se minimizar os problemas conhecidos em relação às outras técnicas de recomendação. Para obter o conhecimento sobre os usuários, necessário para o uso da técnica de recomendação adotada, são utilizadas informações disponíveis em redes sociais, resultantes das suas interações, que possam ajudar a caracterizar o seu perfil e o contexto em que eles estão inseridos, sem a necessidade de conhecer recomendações passadas ou avaliações de outros usuários a respeito de itens a serem recomendados.

O sistema de recomendação possui, ainda, o conhecimento sobre as características

[10] http://apps.facebook.com/courseprofiles/

dos recursos educacionais que serão recomendados e através do processo de recomendação é capaz de selecionar quais deles são mais relevantes considerando o conhecimento sobre características e preferências do usuário.

Algumas técnicas de Extração de Informação foram aplicadas aos textos associados às páginas curtidas e compartilhadas por usuários nas suas redes sociais para extrair informação estruturada que possa ser usada no processo de recomendação de recursos educacionais. As técnicas usadas foram:

- detecção de idioma, usado para determinar quais funções podem ser usadas no texto, já que algumas delas são restritas a um conjunto de idiomas. A identificação do idioma é relevante ainda para descobrir idiomas de domínio do usuário, caso essa informação não esteja explicitada no seu perfil;
- reconhecimento de entidades, que visa a extração e classificação de pessoas, empresas, organizações, cidades, características geográficas e outras entidades mencionadas no texto;
- identificação de conceitos, que é capaz de entender como conceitos se relacionam e, ainda, identificar conceitos que não são necessariamente diretamente referenciados no texto;
- identificação de palavras-chave para encontrar quais palavras foram mais relevantes no texto e classificá-las;
- classificação do conteúdo, que analisa o texto, buscando incluí-lo em categorias pré-definidas organizadas hierarquicamente.

Além de usar técnicas de Extração de Informação para obter informações estruturadas a partir de textos extraídos de redes sociais, algumas ontologias foram usadas para buscar interesses relacionados aos temas extraídos.

No contexto dessa pesquisa, os dados ligados foram usados como fonte de consulta de dados disponíveis para serem usados como recursos educacionais para recomendação.

3 PROPOSTAS DE SISTEMAS DE RECOMENDAÇÃO

Muitas propostas de Sistemas de Recomendação sensíveis ao contexto podem ser encontradas na literatura e, de alguma forma, todas elas estão relacionadas a esta pesquisa. No entanto, considerando a amplitude do tema, os trabalhos apresentados neste tópico se restringem a Sistemas de Recomendação educacionais que consideram o contexto dos seus usuários para gerar recomendação ou adaptação de conteúdo.

ALHARBI *et al*. (2012) descrevem uma arquitetura baseada em contexto para recomendação em Ambientes Pessoais de Aprendizagem ou *Personal Learning Environments* (PLEs). O foco da recomendação dessa proposta são serviços independentes (Facebook, YouTube, calendário pessoal e a biblioteca digital IEEE Xplore) condizentes com o interesse do usuário. A recomendação é feita a partir de informações contextuais dos serviços e do aluno. O contexto dos serviços é capturado através de sensores virtuais, e o contexto do usuário considera as informações do seu perfil e seu feedback em relação a recomendações anteriores. As informações de contexto do usuário são categorizadas em pessoais ou estáticas e preferências. As informações pessoais consideradas nesse estudo são: nome, endereço, idade, estado civil e e-mail. O resultado da recomendação é um conjunto de serviços e recursos sugeridos a partir do interesse atual do usuário, gerando notificações de links e páginas.

KLANJA-MILICEVIC *et al*. (2012) propõem uma arquitetura de um sistema de recomendação baseado em *tags* para ambientes de e-*learning*, nomeado Protus. O Protus é um sistema de tutoria projetado para ajudar os alunos no aprendizado dos fundamentos de linguagem de programação. O principal objetivo é recomendar materiais úteis e interessantes para alunos a distância, tendo como base suas atividades (como padrões sequenciais, páginas visitadas, resultados de testes e notas obtidas), diferentes experiências, preferências, objetivos de aprendizagem e outros atributos significativos. O modelo proposto usa um modelo de aluno para reconhecer os objetivos do aluno, seu estilo de aprendizagem, e para filtrar objetos de aprendizagem que possuam conteúdo apropriado para ele. O Protus utiliza técnicas de clusterização para agrupar os alunos pelo seu estilo de aprendizagem. Após a definição dos grupos, eles são usados para determinar atividades educacionais apropriadas.

WANG & WU (2011) apresentam uma proposta de recomendação sensível ao contexto para implementar um sistema educacional ubíquo (u-*leaning*) adaptativo. O objetivo é permitir ao aluno construir o seu ambiente de forma auto gerenciada e personalizada. Usando as transmissões da rede sem fio e a detecção das metas de aprendizagem em situações

reais através de sensores, o material didático apropriado é recomendado à plataforma de aprendizagem do aluno, considerando o seu contexto. O método de recomendação adotado pelo estudo utiliza a experiência dos usuários que possuem metas de aprendizagem semelhantes e a avaliação feita por eles aos recursos educacionais. Após a recomendação, o aluno avalia o item recomendado e aceito por ele. Este *feedback* dado pelo aluno permite o aprendizado do sistema e uma recomendação mais qualificada, uma vez que recursos bem aceitos e avaliados são mais recomendados a outros alunos. A definição do perfil do aluno é feita usando a técnica de Filtragem Colaborativa e Mineração dos registros de navegação na Web.

Em (FRITZEN *et al*., 2012) é apresentada uma proposta que busca modelar o contexto a partir da extração de mensagens trocadas entre alunos em grupos de redes sociais e usar esse contexto para melhorar a relevância na recuperação de documentos na Web, tendo como objetivo apoiar a aprendizagem colaborativa. A intenção da proposta é tornar a recuperação de informação sensível ao contexto das discussões (por intermédio do uso das mensagens para modelagem do contexto), oferecendo, portanto, resultados de busca contextualizados. Os autores propõem a captura do contexto a partir do enriquecimento das mensagens em dados abertos e o uso do contexto para melhorar as consultas do usuário, fornecendo-lhe conteúdo mais adequado a partir da Web.

DWIVEDI & BHARADWAJ (2012) apresentam uma proposta de recomendação baseada na avaliação de conteúdo em uma rede social educacional, onde os alunos compartilham uns com os outros as suas classificações para recursos, incluindo livros, materiais didáticos, etc. Sempre que um aluno deseja uma recomendação sobre um determinado conteúdo, ele envia uma consulta para o seu amigo imediato e para os amigos distantes. O amigo imediato é aquele que está ligado a ele por um salto (ligado diretamente), enquanto que o amigo distante é aquele que está ligado a ele por 2 ou mais saltos. Os alunos que usaram o recurso no passado respondem a consulta com suas classificações. Com base nas respostas e no histórico das classificações entre amigos, é feito o cálculo da pontuação para a recomendação. A rede social considera o estilo de aprendizagem e o nível de conhecimento de tal forma que os alunos que possuem um certo grau de similaridade nesses atributos são considerados amigos imediatos. Sendo assim, a recomendação feita por um amigo imediato será compatível com o estilo de aprendizagem e o nível de conhecimento do aluno.

CASAGRANDE *et al*. (2013) apresentam uma técnica de recomendação baseada na construção implícita de perfis de usuários através da observação dos valores de metadados

dos conteúdos que esses usuários acessam. A técnica de recomendação proposta utiliza a filtragem colaborativa combinada com uma técnica de agrupamento de usuários por interesses similares.

Em FERREIRA *et al*. (2013) é apresentado o UbiGroup, um modelo de recomendação ubíqua de conteúdo educacional para grupos de alunos. Esse modelo leva em conta o perfil dos alunos e o contexto no qual eles estão inseridos. O UbiGroup faz uso de agentes de software para realizar o processo de recomendação. Feita a associação são utilizados filtros considerando o perfil e o contexto do aluno, que são definidos manualmente.

Em FRADE *et al*. (2014) é apresentada uma arquitetura, chamada UniVirtual, baseada em agentes para recomendação de objetos de aprendizagem em um ambiente virtual 3D. A solução proposta objetiva tornar a aprendizagem virtual personalizada e adequada dependendo das necessidades de cada estudante. Nessa arquitetura, os alunos precisam, inicialmente, estarem cadastrados no ambiente virtual de aprendizagem Moodle[11]. As informações que caracterizam o perfil do aluno são extraídas das informações geradas nesse LMS (*Learning Management System*). Essa arquitetura pode ser aplicada em qualquer ferramenta de gestão de aprendizagem. Após realizar a autenticação, o usuário pode explorar um ambiente tridimensional através de um personagem que representa o estudante. Cada departamento possui um professor virtual que interage com o usuário através da oferta de objetos de aprendizagem de forma autônoma. O que caracteriza essa autonomia é a possibilidade de identificação da proximidade do estudante por meio de uma implementação básica de Inteligência Artificial (IA) associada ao sistema multiagentes que encontra-se integrado ao ambiente e que trata, por sua vez, da recomendação dos objetos de aprendizagem baseada na localidade e perfil do estudante em questão.

SUDHANA *et al*. (2013) descrevem uma proposta de recomendação de recursos educacionais que considera elementos do contexto do aluno fazendo uso de ontologias para representar este contexto. As informações contextuais usadas são divididas em contexto estático e contexto dinâmico. No contexto estático são consideradas as informações pessoais (ex. nome, qualificação, nacionalidade) e as informações referentes ao ambiente educacional do aluno, que, geralmente, não estão sujeitas a constantes mudanças. Já no contexto dinâmico, são relevantes as informações relacionadas à preferência e ao interesse do aluno, que podem mudar durante o processo educacional. Os dados do perfil do aluno são coletados através do seu registro que contém dados pessoais como qualificação, posição, preferências, entre outros.

[11] http://moodle.com.br/site/

O calendário e as atividades educacionais do aluno são armazenados na base de dados de atividades educacionais.

Em (REIS; BARRÉRE, 2014) é proposta uma arquitetura que permite que diversos conteúdos educacionais sejam recomendados a alunos. Esses alunos devem ser inicialmente categorizados pelo professor em relação ao nível de conhecimento sobre um determinado assunto (como conhecimento básico, médio ou avançado). A recomendação inicial ocorre conforme indicação do professor de qual conteúdo é sugerido para cada nível de conhecimento. Novos conteúdos poderão ser indicados pelo professor ou pelos alunos durante o andamento das atividades. A partir da interação dos alunos com os conteúdos, o mecanismo de recomendação levará em consideração um conjunto de informações como perfil do usuário, informações de contexto e qualidade de experiência (QoE), para compor uma recomendação adaptada. A arquitetura permite também que os alunos trabalhem de forma colaborativa, onde os mesmos possam indicar conteúdo.

RIBEIRO *et al*.(2013) propõem um sistema de recomendação baseado no conjunto das *hashtags* atribuídas nas postagens dos fóruns, fazendo uso do esquema da frequência de ocorrência do termo no documento e do inverso da frequência do termo entre documentos da coleção. Nessa proposta é necessário que os usuários atribuam *hashtags* em suas postagens. Essas *hashtags* têm finalidade classificatória, criando agrupamentos textuais, que serão utilizadas pelo sistema para fazer as recomendações.

Em (REZENDE *et al*., 2013) é proposta a arquitetura do BROAD-RS que permite a automação do processo de recomendação do objetos de aprendizagem adaptado aos interesses do aluno e de acordo com o modelo didático especificado pelo professor. Este modelo foi baseado em metadados pedagógicos retirados dos padrões BROAD (NERY *et al.*, 2012), LOM (DUVAL; HODGINS, 2003) e OBAA (VICARI *et al*. 2009), contextualizados para interferir no interesse do aluno, como, por exemplo: conteúdos de maior interesse, mídias prediletas, o nível de aprendizagem, entre outros. O foco desta pesquisa está na recomendação, no modelo didático e no gerenciamento de conteúdos personalizados. A abordagem proposta também faz uso de um sistema multi-agentes que leva em conta os modelos didáticos e o contexto do aluno para recomendar o OA mais adequado.

Muitas abordagens diferentes são usadas em Sistemas de Recomendação sensíveis ao contexto encontrados na literatura. As principais distinções estão nas técnicas usadas para determinar ou enriquecer o perfil do usuário (captura de informações através da navegação em sistemas, manualmente, interação com o conteúdo, similaridade entre perfis, etc.), na técnica de recomendação (filtragem colaborativa, baseada em conteúdo, híbrida, baseada em regras

semânticas, etc.) e alvo da recomendação (objetos de aprendizagem, ambientes educacionais, sites, cursos, etc.). A tabela 2 apresenta um resumo dos trabalhos relacionados apresentados, sendo que a primeira coluna se refere aos itens descritos na tabela 1.

Tabela 1 - Legendas usadas na análise das propostas de recomendação.

NÚMERO	ITENS DE ANÁLISE
1	Extração de características do perfil e do contexto do usuário
2	Alvo da recomendação
3	Técnica ou abordagem de recomendação
4	Características principais

Fonte: elaborada pelo autor.

Tabela 2 - Comparativo entre as propostas de recomendação.

(ALHARBI *et al*., 2012)	
1	Perfil traçado através de aplicação de formulário Registros de logs (uso do sistema)
2	Serviços independentes (Facebook, YouTube, Personal Calendar and IEEE Xplore *digital library*)
3	Baseada em conhecimento Baseada em conteúdo
4	Recomendação com o objetivo de criar um ambiente pessoal de aprendizagem.
(KLANJA-MILICEVIC *et al*., 2012)	
1	Registros de logs (AVA - padrões de navegação, páginas visitadas, resultados de testes e notas obtidas) Tags inseridas por usuários
2	Recursos educacionais
3	Clusterização para agrupar os alunos no mesmo nível de conhecimento Recomendação baseada em tags (busca por tags inseridas pelos usuários)
4	Agrupamento de usuários por nível de conhecimento, permitindo separação de recomendação do mesmo tema em níveis de dificuldade.
(WANG; WU, 2011)	
1	Identificação automática através de sinais de rádio (tecnologia RFID - *Radio-Frequency IDentification* (FINKENZELLER, 1999))
2	Recursos educacionais
3	Filtragem colaborativa Regras de associação (mineração de dados)
4	Sistema de recomendação Ubíquo, onde o aluno pode receber recomendações em qualquer lugar e em qualquer momento.
(FRITZEN *et al*., 2012)	
1	Mensagens trocadas entre alunos em grupos de redes sociais são extraídas através de API da rede social.
2	Documentos da Web
3	Busca contextualizada na Web Baseada em conhecimento
4	Definição automática de contexto em redes sociais Busca contextualizada Recomendação em grupos de discussão em redes sociais.
(DWIVEDI; BHARADWAJ, 2012)	

1	Perfil traçado através de aplicação de formulário
2	Recursos educacionais
3	Filtragem colaborativa
4	Criação de uma rede de recomendação Associação de alunos com características semelhantes como amigos imediatos Propagação da solicitação de recomendação pela rede.
(CASAGRANDE *et al*., 2013)	
1	Registro de metadados dos recursos educacionais acessados
2	Recursos educacionais de um repositório determinado
3	Filtragem colaborativa Agrupamento de usuários por interesses similares
4	Construção do perfil através da análise dos metadados de recursos acessados Simplicidade e escalabilidade da técnica de recomendação.
(FERREIRA *et al*., 2013)	
1	Registro de logs (Monitoramento do sistema através de agentes de software)
2	Objetos de aprendizagem
3	Baseada em conhecimento
4	Arquitetura multi-agente.
(FRADE *et al*., 2014)	
1	Registro de logs (Monitoramento dos registros do LMS Moodle)
2	Objetos de aprendizagem
3	Baseada em conhecimento Baseada em localização (técnica demográfica)
4	Ambiente virtual 3D.
(SUDHANA *et al*., 2013)	
1	Registro do aluno no sistema (formulário de cadastro)
2	Recursos educacionais
3	Baseada em conhecimento
4	Uso de ontologias para a representação do perfil e do contexto do aluno.
(REIS; BARRÉRE, 2014)	
1	Informações fornecidas pelo professor Registro de logs (monitoramento das ações do usuário no sistema de recomendação)
2	Recursos educacionais de um repositório determinado
3	Baseada em conhecimento Filtragem colaborativa
4	Ambiente m-*learning* Consideração de características dos dispositivos nos quais o aluno acessa o material.
(RIBEIRO *et al*., 2013)	
1	Informações de fóruns em ambientes virtuais de aprendizagem Inserção de *hashtags* no texto dos fóruns
2	Recursos educacionais
3	Filtragem de conteúdo por *hashtags*, priorizando através do peso de cada *hashtag*
4	Definição de peso para as *hashtags* inseridas em fóruns para priorização da recomendação.
(REZENDE *et al*., 2013)	
1	Registro do aluno no sistema (formulário de cadastro)
2	Objetos de aprendizagem
3	Regras semânticas
4	Uso de agentes para automatizar a recomendação

	Inserção do modelo didático no processo de recomendação.

Fonte: elaborada pelo autor.

3.1 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo foram apresentados e comparados alguns trabalhos relacionados a Sistemas de Recomendação na área educacional. Através da análise dos trabalhos foi possível perceber quais as dificuldades apresentadas por eles, bem com as diferentes soluções adotadas por cada um. Essa análise permitiu a tomada de decisão em relação à proposta apresentada nesta dissertação. O sistema de recomendação apresentado nesta dissertação evolui a proposta feita em (REZENDE *et al.*, 2013) e traz aspectos inovadores em relação a ele e aos demais trabalhos (ALHARBI *et al*., 2012) (DWIVEDI; BHARADWAJ, 2012) (REIS; BARRÉRE, 2014) (SUDHANA *et al*., 2013) no sentido de extrair e considerar informações geradas espontaneamente nas redes sociais, mais especificamente no Facebook, já que esses sistemas obtêm informações acerca do usuário através da aplicação manual de formulários.

Na proposta de FRITZEN *et al*. (2012) as redes sociais já haviam sido exploradas para extração automática de contexto a partir das discussões em grupos, sendo esse usado para recuperação de documentos na Web. A proposta apresentada nesta dissertação traz avanços em relação a essa proposta ao tornar possível gerar recomendações individuais, especificamente educacionais, coerentes com os interesses e as preferências de qualquer usuário na rede social, mesmo que este não esteja inserido em um grupo com discussões no âmbito educacional.

A possibilidade de envio das recomendações para a rede social do usuário, permitindo que ele possa usar parte do tempo que passa interagindo para ter acesso a conteúdo educacional e que possa, ainda, compartilhar tais conteúdos, recomendá-los para outros usuários, enfim, utilizar todos os recursos intrínsecos das redes sociais traz progressos em relação aos trabalhos de CASAGRANDE *et al*. (2013) FERREIRA *et al*. (2013) FRADE *et al.* (2014); KLANJA-MILICEVIC *et al.* (2012); RIBEIRO *et al*. (2013); WANG; WU (2011), uma vez que esses utilizam informações extraídas de sistemas acadêmicos, tais como ambientes virtuais de aprendizagem, restringindo assim o acesso às recomendações para usuários que estejam de alguma forma inseridos em um contexto acadêmico.

A proposta apresentada neste trabalho traz avanços em relação a todos os trabalhos no que se refere à recomendação de pessoas com interesses semelhantes e ao uso de conjuntos de dados ligados para buscar conteúdos educacionais.

Na tabela 3 é feito um resumo da proposta apresentada nesta dissertação, seguindo os mesmos itens de análise da tabela 1.

Tabela 3 - Principais características da proposta de recomendação BROAD-RSI.

1	Extração de informações de perfil e contexto através da interação do usuário em redes sociais (páginas curtidas, grupos dos quais participa, itens de perfil, etc)
2	Objetos de aprendizagem Páginas Web Pessoas com interesses semelhantes Vídeos
3	Baseada em conhecimento
4	Extração do perfil dos usuários através de informações geradas automaticamente em redes sociais Recomendação baseada em repositório e em dados ligados.

Fonte: elaborada pelo autor.

4 BROAD-RSI

Neste capítulo apresentamos a solução proposta nomeada BROAD-RSI (Sistema de Recomendação Baseado nas Interações Sociais) (PEREIRA *et al.*, 2014a) (PEREIRA *et al.*, 2014b) (PEREIRA *et al.*, 2014), seus objetivos e sua arquitetura. O BROAD-RSI tem três objetivos principais: (1) a identificação de características do perfil e do contexto do usuário a partir de informações geradas, espontaneamente, através das interações nas redes sociais; (2) o uso do perfil e do contexto do usuário para auxiliar no processo de seleção e recomendação de recursos educacionais adequados à ele; (3) o oferecimento de recomendações educacionais, baseadas em interesses e preferências, a usuários que não são alunos regulares de nenhum curso ou instituição, no entanto possuem interesse em se capacitar; (4) a indicação de pessoas com interesses educacionais semelhantes, buscando aproximá-las na intenção de formarem uma rede de interesses em torno de um determinado assunto.

Os objetivos principais podem ser expandidos em outros objetivos secundários, a citar:

- Explorar o longo tempo dispendido por usuários nas redes sociais, levando até eles recomendações de recursos educacionais;
- Explorar o interesse dos usuários em estarem atualizados através das redes sociais;
- Usar recursos disponíveis em redes sociais, tais como “recomendar”, “curtir”, “compartilhar”, “comentar” para as recomendações feitas, aumentando assim a capacidade de propagação das recomendações educacionais, bem como a qualidade de conteúdos educacionais distribuídos;
- Avançar em relação ao projeto BROAD-RS (REZENDE *et al.*, 2013) acrescentando a definição do perfil e do contexto dos usuários, através da aquisição de informações disponibilizadas por eles nas redes sociais.

Vale esclarecer que, no escopo deste trabalho, interesses educacionais são considerados como todos os temas para os quais o usuário deseja obter informação, ou seja, tema que ele deseja aprender mais, obter conhecimento, sem que eles necessariamente estejam relacionados às áreas de sua formação acadêmica.

Para atingir os objetivos citados, foi escolhida a rede social Facebook, pelos motivos já citados na motivação, que incluem seu grande uso entre estudantes e entre usuários de forma geral.

O processo simplificado de extração das informações do usuário e do seu uso na recomendação, conforme está sendo adotado pelo BROAD-RSI, pode ser visualizado na figura 4. Ele tem início através do acesso do usuário ao BROAD-RSI, usando sua conta do Facebook. Em seguida, o usuário será questionado sobre a autorização para que suas informações sejam usadas pelo BROAD-RSI. Diante da autorização dada pelo usuário, é feita a extração das informações, que incluem dados de perfil, lista de amigos, páginas curtidas, postagens e grupos. Na etapa seguinte, as informações são filtradas, buscando aquelas que definem interesses educacionais, preferências sobre dias e horários de acesso, idiomas, mídias e, ainda, os dados de perfil, tais como nome, e-mail, idade e gênero. Uma vez filtradas as informações necessárias, os interesses identificados passam por uma etapa de enriquecimento, usando algumas técnicas de Extração de Informação e de Web Semântica, através de ontologias, para identificar interesses relacionados.

Após a etapa de enriquecimento dos interesses é feita a representação semântica desses e também do perfil e do contexto do usuário. Por fim, o perfil e contexto são usados para determinar os parâmetros de busca e relevância dos recursos educacionais e pessoas com interesses semelhantes. A etapa posterior à busca inclui a priorização dos recursos encontrados e envio dessas recomendações para o usuário, através de um aplicativo na rede social.

O BROAD-RSI é apresentado em dois níveis de detalhamento de representação da arquitetura proposta. Neste capítulo a sua arquitetura é detalhada e no capítulo 5 são apresentados os aspectos relacionados ao desenvolvimento da proposta.

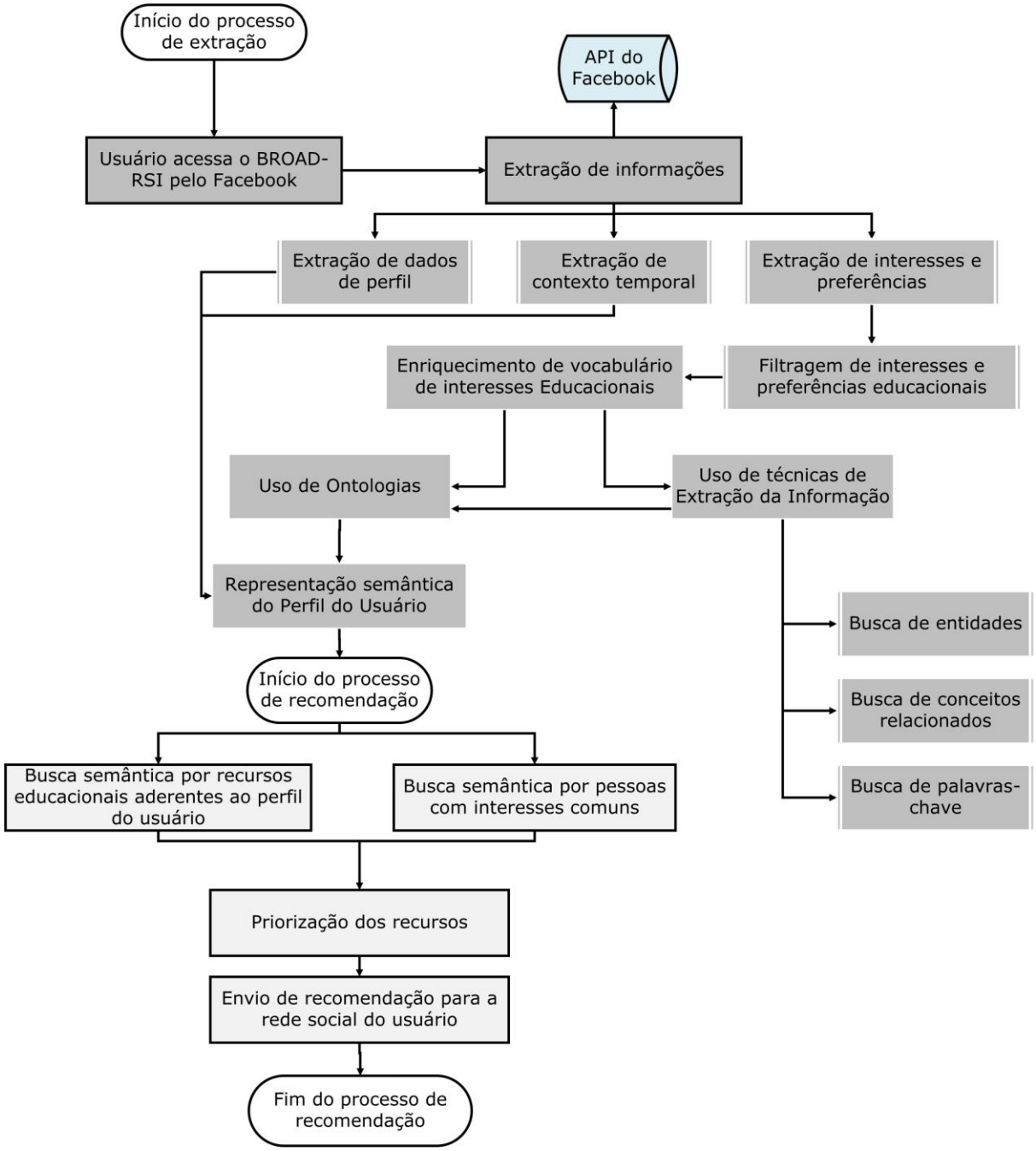


Figura 4- Fluxo simplificado de recomendação do BROAD-RSI.

4.1 ARQUITETURA

A arquitetura do BROAD–RSI é dividida em 4 módulos: Extração de Informação; Enriquecimento do Perfil; Representação Semântica e Sistema de Recomendação. Pretende-se através dela: (1) especificar uma estratégia para extrair informações acerca do interesse educacional e informações básicas de perfil dos usuários, usando para isso a rede social Facebook; (2) determinar itens de perfil educacional do usuário a partir das informações

extraídas; (3) enriquecer o perfil extraído, buscando tópicos de interesses implícitos; (4) gerar a representação semântica desse perfil e (5) considerar as informações extraídas para sugerir recursos educacionais apropriados.

Para processar as informações extraídas, definir características de perfil e de contexto do usuário e gerar recomendações educacionais aderentes a essas características definiu-se a arquitetura do BROAD-RSI, apresentada na figura 5.

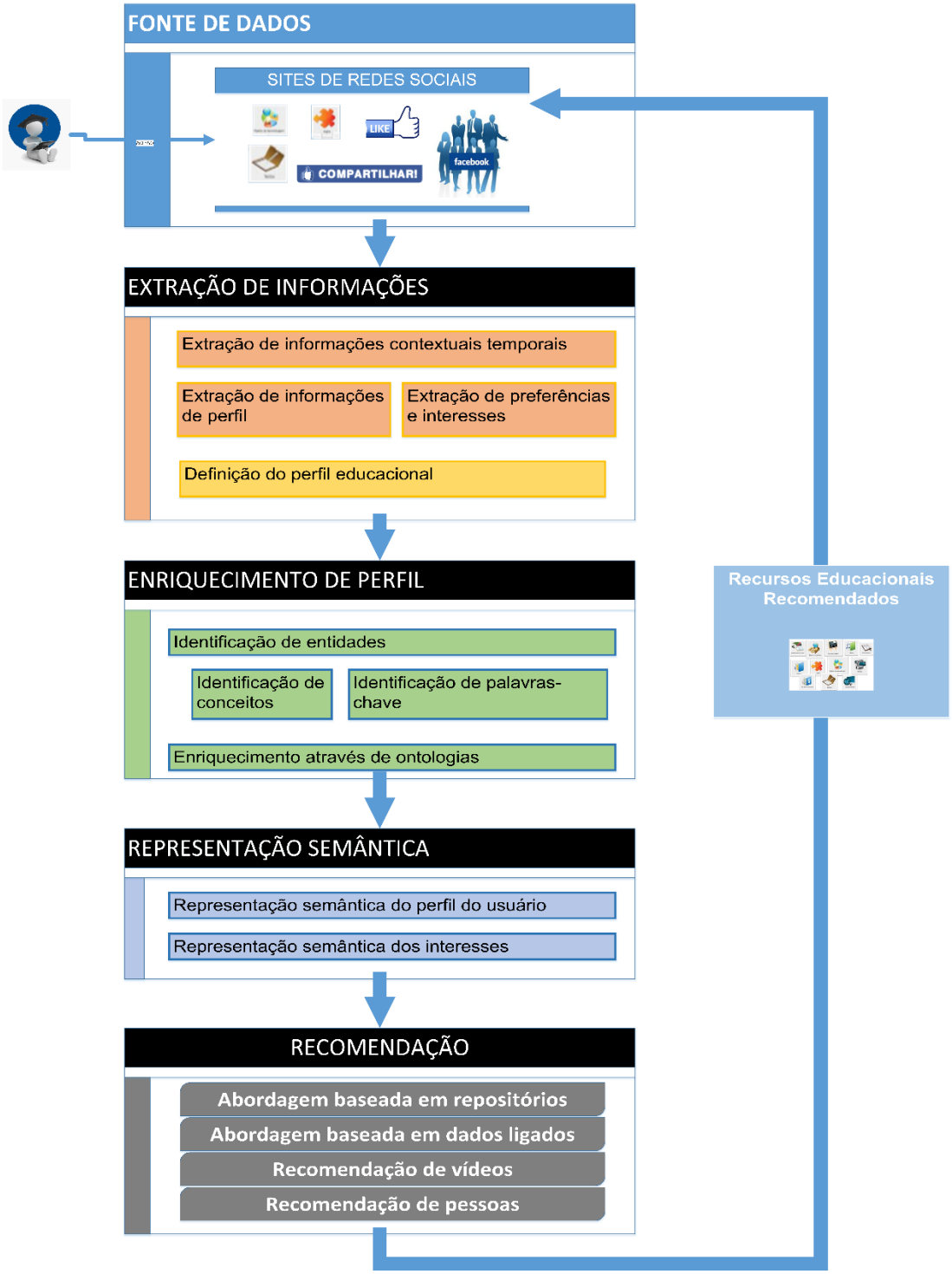


Figura 5 – Representação da arquitetura do BROAD-RSI.

4.2 EXTRAÇÃO DE INFORMAÇÕES

O módulo de extração de informações é responsável por identificar as informações disponíveis através do perfil e das interações do usuário, selecionar quais são relevantes para a recomendação e filtrar tópicos relacionados aos seus interesses educacionais.

O processo de extração de informações inicia-se a partir do acesso do usuário pela sua conta do Facebook. Conforme dito anteriormente, após realizar o acesso, o usuário deve aceitar as permissões necessárias do aplicativo para que as informações sejam extraídas (figura 6).

Figura 6 - Permissão necessária para a extração de dados do usuário no Facebook.

O Facebook disponibiliza, através de sua API[12], informações referentes às conexões de cada usuário, informações públicas do seu perfil (nome, idade, cidade de nascimento e residência, dados educacionais, dados profissionais, idiomas falados), seus interesses e preferências (livros, músicas, filmes, jogos, páginas, entre outros) e ainda os grupos dos quais ele participa.

A definição de quais informações foram consideradas no BROAD-RSI foi feita em duas etapas. Na primeira, buscou-se identificar quais eram as informações relevantes para Sistemas de Recomendação educacional, ou seja, quais elementos do perfil e do contexto do usuário seriam úteis para gerar recomendações relevantes. A investigação dessa questão foi feita através de um mapeamento sistemático (PEREIRA *et al.*, 2014), cujo objetivo foi identificar os elementos contextuais encontrados em trabalhos relacionados a Sistemas de

[12] https://developers.facebook.com/

Recomendação educacionais sensíveis ao contexto e fazer uma análise quantitativa das categorias de contexto, além de uma lista abrangente de informações em cada categoria contextual.

Na segunda etapa foi feito um filtro no conjunto de informações identificadas na primeira etapa, considerando apenas aquelas possíveis de serem extraídas da rede social Facebook. Ao final das duas etapas, foram definidos os elementos contextuais indicados na tabela 4. As informações foram categorizadas de acordo com o framework definido em (VERBERT *et al*., 2012), conforme detalhado em (PEREIRA *et al.*, 2014).

Tabela 4 - Elementos contextuais usados no BROAD-RSI.

CATEGORIA	SUBCATEGORIA	ELEMENTOS DE CONTEXTO
Computacional	Hardware	Tipo de dispositivo (*notebook, tablet, desktop*, celular)
Tempo	Horário	Horários das postagens nas redes sociais
	Intervalo de tempo	Dia da semana Período do dia com maior número de acesso
Atividades	Tópicos	Interesses identificados através das interações do usuário
Usuário	Informações básicas	Nome Qualificação Cidade onde reside Histórico educacional Idade
	Conhecimento	Conhecimento de idiomas Competências educacionais
	Preferências Interesses	Interesses gerais e educacionais Preferência (mídia, comunicação, música, notícias, filmes) Dispositivos utilizados
Relacionamentos Pessoais		Interesses comuns entre amigos Similaridade entre pessoas Grupos de discussão nos quais o usuário participa

Fonte: elaborada pelo autor.

4.2.1 **Definição do Perfil Educacional**

Depois de extraídas as informações citadas anteriormente, faz-se necessário filtrá-las, buscando quais, dentre os diversos interesses do usuário, podem ser considerados de cunho Educacional. Para isso são consideradas: as áreas de interesse definidas no Perfil Educacional do usuário e as páginas curtidas, cuja categoria esteja relacionada a temas ou interesses educacionais, tais como, páginas referentes a faculdades, escolas, linhas de

pesquisa, sites educacionais, entre outras. É extraído também o nível educacional do usuário (Ensino Médio, Graduação, Pós-Graduação). Cabe recordar que são considerados temas ou interesses educacionais, no escopo deste trabalho, todos aqueles assuntos que o usuário tem vontade de conhecer e aprender sobre, não estando esse conjunto restrito a temas acadêmicos.

Apesar das informações educacionais não serem obrigatórias, elas estão presentes no perfil de muitos usuários. Segundo um estudo feito no escopo deste trabalho com uma amostra de 251 usuários, 92 deles possuíam, em seu perfil público, informações sobre áreas educacionais de interesse, ou seja, 36,65%. Os participantes do estudo não necessariamente estavam inseridos em um ambiente acadêmico. Eles eram amigos na rede social Facebook de alguns alunos voluntários do mestrado em Ciência da Computação da Universidade Federal de Juiz de Fora. Considerando o amplo uso do Facebook, o quantitativo de pessoas beneficiadas por uma recomendação educacional pode ser significativo.

Além das informações relacionadas aos interesses educacionais dos usuários, são capturadas também aquelas relativas à sua preferência em relação aos dias, horários e períodos (manhã, tarde e noite) de acesso à rede social. Pretende-se, com isso, enviar recomendações educacionais nos horários mais oportunos de acordo com o perfil de acesso do usuário. São extraídas, ainda, o tipo de dispositivo que o usuário utiliza para acessar a rede social e os tipos diferentes de mídias, a partir dos quantitativos dos diferentes tipos de recursos compartilhados e curtidos pelo usuário no Facebook, tais como Livros, Músicas, Vídeos, Imagens e Jogos. Essa informação ajuda a priorizar recursos de acordo com o seu formato, não sendo usada, portanto, para eliminar uma recomendação, mas apenas para dar mais importância e prioridade a um subconjunto delas.

4.3 ENRIQUECIMENTO DE PERFIL

O módulo de enriquecimento de perfil é responsável por identificar interesses relacionados àqueles explicitamente informados nos perfis das redes sociais, através de técnicas de Extração de Informação e Web Semântica.

As páginas, postagens, grupos e demais informações extraídas do perfil do usuário, geralmente, possuem um nome e um texto que as descreve, conforme mostrado na figura 7. Em uma primeira versão deste trabalho, foram considerados apenas o nome, título ou o texto de uma área de interesse, item ou página. Foi possível perceber que não haviam elementos suficientes para compreender os tópicos de interesse de um usuário e gerar recomendações educacionais relevantes. O nome ou título, geralmente, trazem informações

escassas ou muito simplificadas, impedindo o seu uso para, por exemplo, gerar recomendações relacionadas. Já o texto de cada página traz informações mais completas que permitem algumas associações. No entanto, tais informações encontram-se dispersas ao longo do texto sem uma estruturação pré-definida que permita a sua extração.

Sendo assim, fez-se necessário explorar técnicas de Extração de Informação e Web Semântica, através de ontologias, para obter informações relevantes e enriquecer os itens de perfil e contexto extraídos da rede social, obtendo-se então um melhor entendimento dos temas de interesse e das suas associações a interesses correlatos.

Figura 7 - Exemplos de textos descritivos de páginas no Facebook.

As principais tarefas de Extração de Informação usadas foram: detecção de entidades, identificação de conceitos, identificação de palavras-chave e detecção de idioma do texto. Já o enriquecimento com Web Semântica foi usado através de ontologias de domínio específicas para cada área de conhecimento. Ambas as abordagens possuem vantagens e desvantagens e o uso das duas se complementam no objetivo de enriquecer o perfil do usuário.

A abordagem de enriquecimento usando tarefas de Extração de Informação possui a vantagem de não depender de um especialista em cada uma das áreas de conhecimento para criar ontologias e de poder ser usada em qualquer domínio de conhecimento. Porém, por outro lado, as associações são mais amplas, trazendo resultados que, em alguns casos, podem estar

distantes dos interesses explicitados pelo usuário. A solução adotada para minimizar o problema da identificação de termos distantes dos reais interesses do usuário foi a intervenção do mesmo que pode confirmar ou refugar os interesses identificados.

Já na abordagem de enriquecimento usando ontologias há mais precisão na identificação de termos correlatos, já que a ontologia é construída por especialistas com conhecimento de cada área de conhecimento. Outra vantagem do uso de ontologias é a hierarquia e o sequenciamento de termos, possíveis de serem representados através delas e que podem enriquecer as recomendações, uma vez que permitem a apresentação de recursos em uma sequência definida, preservando os requisitos necessários para o entendimento dos diversos assuntos.

As duas abordagens são utilizadas tanto de forma isolada, quanto em conjunto com a extração de entidades, conceitos e palavras-chave, sendo ambas complementares, uma vez que uma não sendo eficiente para uma determinada situação, a outra abordagem é usada para melhorar os resultados. Caso as páginas não apresentem texto ou o seu texto não resulte em entidades, conceitos e palavras-chave, a ontologia é usada de forma isolada, sendo feitas consultas, cujo termo de busca é o título das páginas. Caso contrário, tanto os títulos, quanto as entidades, conceitos e palavras-chaves são passados como parâmetros para a consulta de termos correlatos nas ontologias.

4.4 REPRESENTAÇÃO SEMÂNTICA

O módulo de representação semântica é responsável por criar uma ontologia do perfil do usuário e dos interesses encontrados, associando as informações extraídas às classes e propriedades pertinentes. Além disso, esse módulo é responsável pelo armazenamento das informações em uma base de dados semântica.

Após a definição do perfil e a identificação dos interesses do usuário, bem como o enriquecimento desses tópicos de interesse, é feita a representação semântica desse perfil, onde são incluídas propriedades relacionadas ao perfil do usuário, propriedades que relacionam o usuário aos seus interesses e às pessoas com as quais ele se relaciona.

A representação semântica se torna importante para futuras inferências, associações e buscas semânticas que podem auxiliar tanto no enriquecimento do perfil do usuário, quanto no enriquecimento dos tópicos de interesses.

Os interesses educacionais, bem como os interesses enriquecidos, também dão origem a uma base semântica de interesses educacionais extraídos do Facebook, relacionados

entre si através de propriedades semânticas. Por exemplo, quanto um interesse X é extraído da página de um determinado usuário, esse interesse é armazenado, já com uma relação semântica de interesse com o usuário, e, após ser enriquecido, os interesses descobertos também são armazenados e associados ao interesse X através de propriedades semânticas. Essa representação semântica é gerada automaticamente como consequência do uso do BROAD-RSI.

4.5 RECOMENDAÇÃO

O módulo de recomendação é responsável por realizar buscas em fontes diferentes (repositórios de objetos de aprendizagem, conjuntos de dados ligados e repositórios de vídeos), estabelecendo a relação entre o perfil do usuário e os recursos educacionais armazenados nessas fontes. Nesse módulo é definida também a prioridade de recomendação a partir da aderência dos recursos educacionais ao perfil e ao contexto do usuário. Além da recomendação de recursos educacionais, é responsabilidade deste módulo recomendar pessoas que possuem interesses semelhantes aos do usuário.

O foco da recomendação do projeto BROAD-RSI são os recursos educacionais, armazenados em seu próprio repositório, e catalogados, através de um conjunto de metadados, conforme detalhado em (NERY *et al*., 2012). No entanto, a dificuldade de localização dos recursos educacionais em repositórios, a falta de catalogação correta, o baixo índice de atualização, comuns em diversos repositórios, têm grande impacto na eficiência dos Sistemas de Recomendação e de personalização de conteúdo educacional, já que esses dependem da quantidade e da qualidade dos recursos educacionais catalogados. Nesse sentido, buscamos explorar uma alternativa para as limitações dos repositórios de recursos educacionais, usando outras fontes de recursos educacionais. A seguir são apresentadas três abordagens: baseada em repositórios de objetos de aprendizagem; baseada em conjuntos de dados ligados; e uma para recomendação de vídeos, baseada no Canal de Vídeos Youtube[13].

4.5.1 **Abordagem baseada em Repositórios de Objetos de Aprendizagem**

A recomendação dos recursos educacionais é feita pelo BROAD-RSI a partir da relação estabelecida entre as características do perfil do usuário e os metadados dos recursos

[13] https://www.youtube.com/

educacionais. Geralmente os recursos educacionais são catalogados usando padrões de metadados, como por exemplo: IEEE LOM (DUVAL; HODGINS, 2003) e OBAA (VICARI *et al*., 2010), permitindo assim a busca e reutilização desses recursos.

Na tabela 5 são apresentados os metadados considerados, bem como a relação estabelecida entre eles e as informações extraídas.

Tabela 5 - Relação entre metadados e informações extraídas.

METADADOS LOM	INFORMAÇÕES EXTRAÍDAS
General.Title	Interesses identificados e interesses inferidos
General.Keyword	
General.Description	
General.Language	Idiomas falados
Educational.InteractivityType *Technical.Format*	Preferência por tipo de mídias
Educational.Context	Nível de escolaridade
Educational.TypicalAgeRange	Idade
Technical.Requirement	Dispositivo utilizado

Fonte: elaborada pelo autor.

O primeiro passo da recomendação é buscar recursos educacionais relacionados aos interesses capturados do usuário e aos interesses inferidos a partir da etapa de enriquecimento semântico. Essa busca retorna um conjunto de recursos que, então, é processado visando determinar a prioridade das recomendações, ou seja, quais dos recursos são mais ou menos aderentes ao perfil do usuário. A aderência é determinada a partir da comparação entre os valores dos metadados *General.Language*, *Educational.InteractivityType*, *Technical.Format*, *Educational.Context*, *Educational.TypicalAgeRange*, *Technical.Requirement* com os Idiomas falados, mídia preferida, nível de escolaridade, idade e dispositivo utilizados, respectivamente.

4.5.2 **Abordagem baseada em Dados Ligados**

A abordagem de recomendação usando dados ligados aumenta a possibilidade de encontrar conteúdos educacionais em diferentes fontes que estão sendo constantemente atualizadas, evitando o problema de escassez de novas recomendações. Muitos repositórios estão começando a disponibilizar seus recursos educacionais no formato de Dados Ligados, ampliando as possibilidades de localização e recomendação de recursos educacionais em um domínio amplo de interesses (Dietze, 2013). Neste trabalho foram usadas duas fontes de

Dados Ligados como fonte de busca de conteúdo: DBPEDIA[14] e *Open University*[15]. A DBPEDIA foi escolhida por três motivos: (1) a grande quantidade de informações disponíveis; (2) a possibilidade de recuperar uma informação passando termos em diferentes idiomas como parâmetros de busca; (3) a diversidade de temas para os quais são disponibilizadas informações. Já a *Open University* traz a vantagem de disponibilizar conteúdo com objetivo especificamente educacional, o que não acontece com a DBPEDIA.

Dentre as muitas informações disponíveis nas duas fontes, algumas podem contribuir com o aprendizado do usuário, como por exemplo, uma breve explicação sobre um determinado assunto, links para conferências, periódicos, páginas diversas que estejam associados ao tema de interesse, apresentações, livros, *podcasts*, vídeos, palestras, entre outros. Na figura 8 são mostrados alguns recursos disponíveis no conjunto de dados ligados da *Open University* para o tema "*Project management*".

Figura 8 - Recursos disponíveis no conjunto de dados ligados da *Open University*.

Uma dificuldade encontrada nessa abordagem é a priorização das recomendações, já que não existe um conjunto de metadados que possa ser associado ao perfil do usuário para então determinar o grau de aderência da recomendação. Buscando minimizar essa limitação, foram usados os recursos das redes sociais da seguinte forma: qualquer recurso educacional é apresentado ao usuário junto com opções para que ele possa curtir, recomendar ou compartilhar e, ainda, o número de pessoas que já executou essas ações para aquele

[14] http://wiki.dbpedia.org/Datasets

[15] http://data.open.ac.uk/

determinado recurso. Dessa forma, o usuário tem através do BROAD-RSI uma indicação da quantidade de pessoas que já indicou ou aprovou determinado conteúdo, sendo esse um critério de priorização da recomendação. A apresentação das recomendações, conforme explicado, pode ser vista nas figuras 9 e 10.

Tema	Links Externos	Facebook
Semantic_Web	http://computemagazine.com/man-who-invented-world-wide-web-gives-new-definition/	Recomendar Compartilhar 1
Semantic_Web	http://esw.w3.org/topic/ConverterToRdf	Recomendar Compartilhar 4
Semantic_Web	http://www.amazon.com/Developers-Guide-Semantic-Web/dp/3642159699/ref=sr_1_1?ie=UTF8&qid=1321027111&sr=8-1	Recomendar Compartilhar 5
Semantic_Web	http://www.semanticoverflow.com/questions/1/where-can-i-learn-about-the-semantic-web	Recomendar Compartilhar 1
Semantic_Web	http://www.worldwidewebsize.com/	Recomendar Compartilhar 1 mil
Semantic_Web	http://portal.acm.org/citation.cfm?id=900051.900063&coll=ACM&dl=ACM&CFID=29933182&CFTOKEN=24611642	Recomendar Compartilhar 0
Semantic_Web	http://www.amazon.com/Explorers-Guide-Semantic-Thomas-Passin/dp/1932394206/	Recomendar Compartilhar 0
Semantic_Web	http://www.amazon.com/Introduction-Semantic-Web-Services/dp/1584889330/	Recomendar Compartilhar 0
Semantic_Web	http://www.amazon.com/Semantic-Primer-Cooperative-Information-Systems/dp/0262012421/	Recomendar Compartilhar 0
Semantic_Web	http://www.amazon.com/Semantic-Web-Technologies-Research-Ontology-based/dp/0470025964/	Recomendar Compartilhar 6

1 / 2 [1 - 10 / 17]

Figura 9 - Recomendação de Links recuperados da DBPEDIA (Tema *Semantic Web*).

Tema	Links Externos	Facebook
E-learning	http://technologysource.org/article/technologyenhanced_learning_in_industry_and_higher_education/	Recomendar Compartilhar 0
E-learning	http://www.efquel.org/	Recomendar Compartilhar 252
E-learning	http://www.montereyinstitute.org/nroc/	Recomendar Compartilhar 61
E-learning	http://www.talentlms.com/elearning	Recomendar Compartilhar 219
E-learning	http://books.google.com/books?id=pFzxLUAnSS8C&printsec=frontcover	Recomendar Compartilhar 0

Figura 10 - Recomendação de Links recuperados da DBPEDIA (Tema *E-learing*).

4.5.3 **Recomendação de vídeos**

Buscando diversificar as fontes de busca de conteúdo educacional usadas pelo BROAD-RSI, é feita também uma consulta ao site de vídeos Youtube[16]. O Youtube oferece uma grande quantidade de vídeos, tendo 300 horas de vídeo sendo enviadas a cada minuto em mais de 61 idiomas. Dentre os inúmeros vídeos disponibilizados, existem importantes iniciativas de produção e disponibilização de conteúdo educacional como o Youtube EDU[17]. Por isso, explora-se no BROAD-RSI a possibilidade de usar este canal de vídeos, disponível gratuitamente, de forma ampla e que engloba um conjunto de temas diversificados para enriquecer as recomendações educacionais.

[16] https://www.youtube.com/

[17] https://www.youtube.com/channel/UCs_n045yHUiC-CR2s8AjIwg

Através da API disponibilizada para desenvolvedores[18], o Youtube oferece a possibilidade de diversas ações, como por exemplo, busca e envio de vídeos, criação de *playlist*, inscrições, etc. No escopo desse trabalho é feita uma busca, onde são considerados os interesses e o idioma falado pelo usuário. A busca é feita no conjunto de todos os vídeos disponíveis pelo Youtube No entanto, vale ressaltar a possiblidade de fazer pesquisas específicas em um determinado Canal, o que permite, por exemplo, que uma instituição, um professor ou um grupo de educadores produzam recursos educacionais no formato de vídeo, criem um canal específico para um determinado assunto e o BROAD–RSI faça buscas específicas nesse canal, aumentando assim a qualidade e precisão das recomendações educacionais.

Assim como no caso da abordagem baseada em dados ligados, as recomendações de vídeos oriundos do Youtube também são apresentadas aos usuários acompanhadas de recursos disponíveis nas redes sociais, como “curtir”, “recomendar”, “compartilhar” e o número de pessoas que realizou tais ações, indicando aos usuários quais vídeos foram mais aceitos pelos usuários da rede social (figura 11).

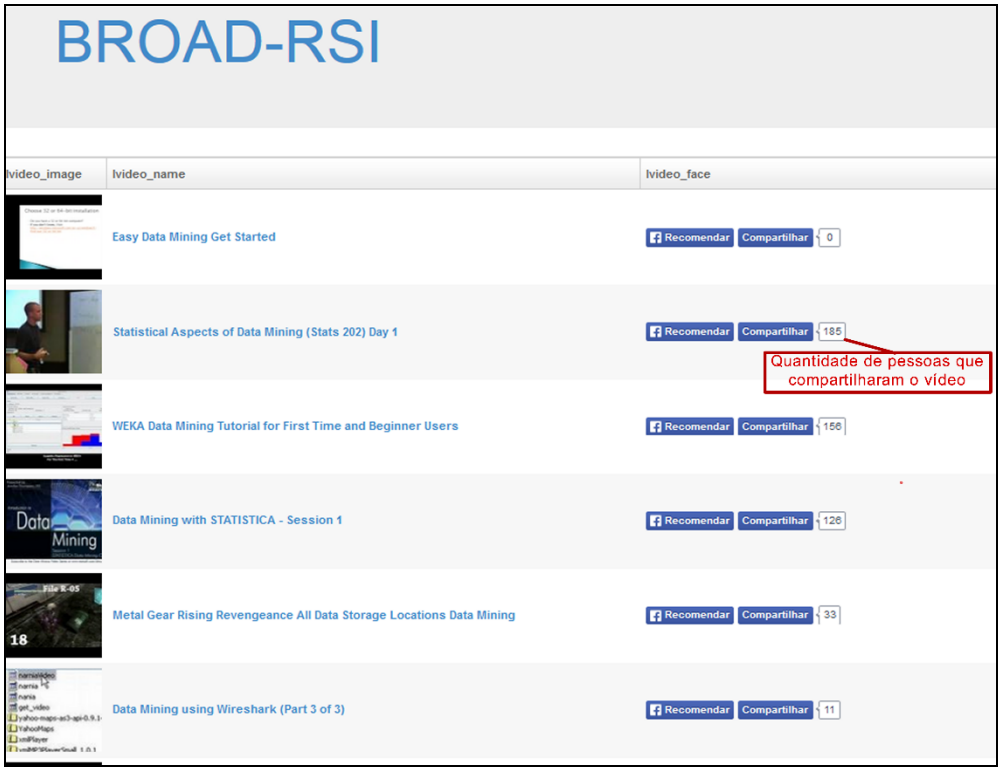


Figura 11 - Recomendação de vídeos do Youtube (Tema *Data Mining*).

4.5.4 **Recomendação de pessoas**

[18] https://www.youtube.com/yt/dev/pt-BR/api-resources.html

A recomendação de pessoas é feita após a identificação dos temas de interesse do usuário, quando então é possível extrair redes de interesses e indicar a ele pessoas que tenham interesses comuns a um determinado usuário. A formação das redes de interesse propicia um ambiente de discussão dos temas de interesse, onde as contribuições dos seus membros enriquecem o conhecimento de todos que fazem parte dela. Mais que a própria discussão que pode ser gerada dentro das redes de interesse, é possível, futuramente, explorar a recomendação de recursos educacionais pela técnica de filtragem colaborativa.

Na figura 12 pode ser vista a formação de uma Rede de Interesse no Tema Ciência da Computação, gerada a partir de informações extraídas do Facebook. O usuário representado pelo rótulo (A) é aquele para o qual serão feitas as recomendações. Os nós representados pelo rótulo (B) são os amigos diretamente ligados a ele, portanto, pessoas que possuem o mesmo interesse e já fazem parte do círculo de amizades do usuário. Já os nós representados pelo rótulo (C), são usuários que possuem o mesmo interesse e não estão ligados diretamente a ele. Através das conexões mostradas é possível indicar ao usuário os seus amigos que possuem interesses comuns e, ainda, quais são os amigos que podem conectá-lo a outras pessoas, que possuem o mesmo interesse, mas ainda não fazem parte do seu círculo de amizade.

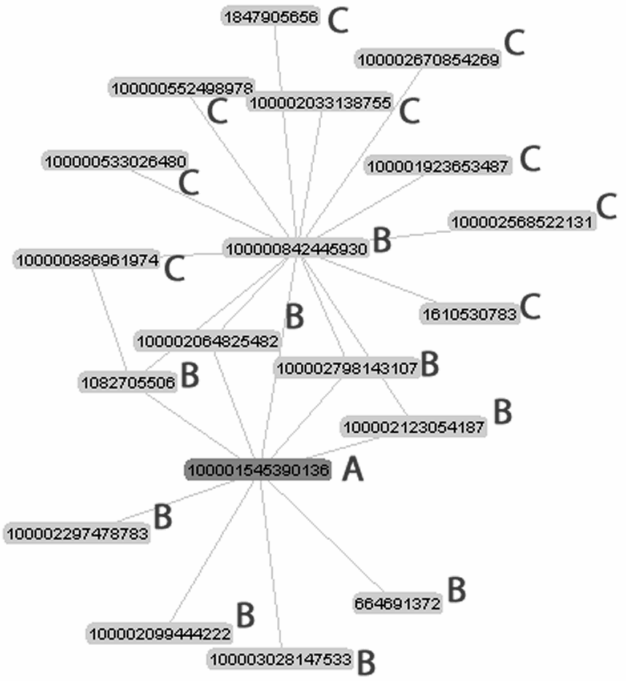


Figura 12 - Rede de pessoas com um determinado interesse comum.

Essa lista de pessoas recomendadas é adquirida a partir de uma busca semântica

na base de usuários do BROAD-RSI. É feito também um cálculo de similaridade entre esses usuários, cálculo esse que será detalhado no capítulo 5, de tal forma que quanto mais interesses comuns os usuários possuem, mais similares eles são.

4.5.5 **Armazenamento e entrega da recomendação aos usuários**

Após realizadas todas as etapas das abordagens apresentadas, o resultado é um conjunto de recursos educacionais (objetos de aprendizagem, links, vídeos, textos) aderentes e priorizados de acordo com o perfil e o contexto do usuário. As recomendações são armazenadas em uma base de dados semântica, onde poderão ser acessadas futuramente para melhoria da própria abordagem de recomendação usada pelo BROAD-RSI.

A entrega da recomendação pode ser feita ao usuário de duas formas: (1) através da solicitação do usuário e (2) através do envio automático da recomendação na rede social do usuário.

Na primeira abordagem, o usuário que deseja receber uma recomendação aderente a um dos seus interesses pode acessar o BROAD-RSI, escolher o interesse e, então, ele receberá uma lista com as recomendações ordenadas por aderência ao perfil e terá a possibilidade de avaliar, curtir, compartilhar ou recomendar cada uma delas. Já na segunda abordagem, o usuário receberá automaticamente em seu perfil na rede social uma recomendação, selecionada a partir dos critérios de priorização, sendo que as primeiras recomendações a serem enviadas são aquelas mais aderentes ao seu perfil, juntamente com um link para realizar a avaliação desta recomendação, conforme mostrado nas figuras 13 e 14. As opções de compartilhar, curtir e recomendar, neste caso, ficam a cargo dos próprios recursos oferecidos pelas redes sociais.

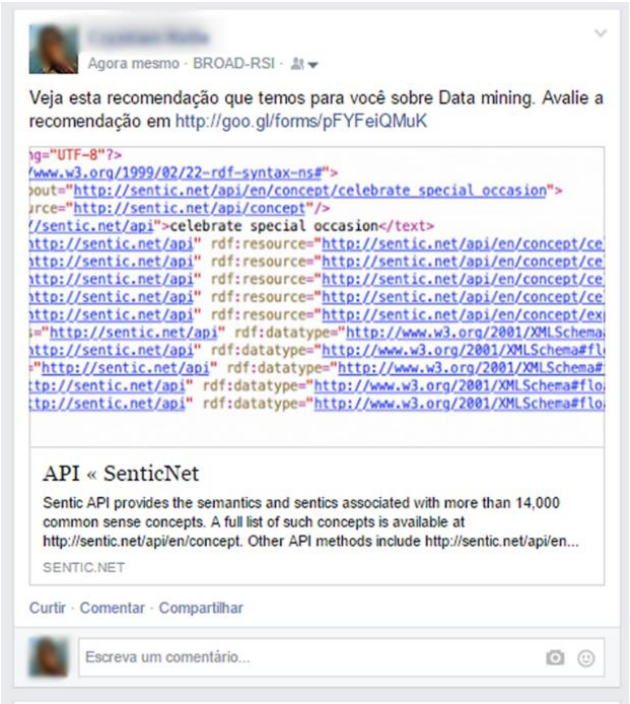


Figura 13 - Recomendação enviada para a rede social do usuário sobre o tema "*Data Mining*".

Figura 14 - Recomendação enviada para a rede social do usuário sobre o tema "*Software Engineering*".

4.6 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo foi apresentada a arquitetura do BROAD-RSI de forma a elucidar as etapas envolvidas no processo de extração de informações dos usuários a partir de suas interações na rede social Facebook e de recomendação de recursos educacionais.

Através dessa arquitetura, buscou-se especificar alternativas para identificar características básicas de perfil e de contexto, interesses e preferências dos usuários, através

da extração de informações na rede social Facebook. Do mesmo modo, buscou-se alternativas para explorar, para fins educacionais, as possibilidades de interações das redes sociais, bem como o tempo dispendido pelos usuários nestes ambientes, levando até eles recomendações educacionais relevantes, oriundas de diferentes fontes de recursos educacionais. Na tabela 6 é feita uma comparação entre as diferentes fontes usadas na proposta, buscando tornar claro suas peculiaridades, vantagens e desvantagens.

Tabela 6 - Comparação das fontes de recursos educacionais usadas na recomendação.

	Repositório de objetos de aprendizagem	Dados Ligados	Vídeos do Yotube
Atualização	Atualizado com pouca frequência	Atualizado com grande frequência (depende do conjunto de dados ligados)	Atualizado com grande frequência
Priorização	Mais precisa	Pouco precisa	Pouco precisa
Especificidade de tema	Muito específico	Dependente do conjunto -DBPEDIA: pouco específica -*Open University*: mais específica	Pouco específco
Especificidade de tipo de recurso	Pouco específico	Dependente do conjunto -DBPEDIA: pouco específica -*Open University*: mais específica	Muito específico
Catalogação	Boa estruturação	Boa estruturação	Pouco estruturada
Principais vantagens	Catalogação geralmente feita por profissionais da educação; Recursos específicos para fins educacionais; Descrição detalhada através de metadados;	Diversidade de temas; Diversidade de mídias; Atualização constante (em alguns casos); Estruturação semântica;	Quantidade de recursos disponíveis; Diversidade de temas e idiomas; Atualização permanente;
Principais desvantagens	Dificuldade de atualização; Repositórios específicos para determinadas instituições.	Nem sempre o conteúdo tem fins educacionais; Falta de metadados na descrição dos recursos.	Nem sempre o conteúdo tem fins educacionais; Falta de metadados na descrição dos recursos; Especificidade do tipo de mídia

			(vídeo).

Fonte: elaborada pelo autor.

Através da explicação do processo de recomendação e das decisões tomadas ao longo do processo é possível atingir os principais objetivos descritos no início do capítulo. Vale ressaltar que o processo de recomendação proposto pode ser expandido para uso em outros sistemas que capturam, através do seu uso, informações relevantes para identificar interesses e preferências dos usuários, como por exemplo, ambientes virtuais de aprendizagem, sistemas acadêmicos, blogs, entre outros. Para tal, seriam necessárias algumas adaptações no mapeamento das informações que darão subsídio à definição do perfil e do contexto e, ainda, na entrega das recomendações, determinando qual a interface a ser usada para apresenta-las ao usuário. Os outros módulos descritos na arquitetura podem ser usados para qualquer um dos sistemas.

A arquitetura do BROAD-RSI continua sendo discutida no capítulo 5, onde serão apresentados detalhes da sua implementação e das tecnologias usadas no seu desenvolvimento.

5 DESENVOLVIMENTO DO PROTÓTIPO

O desenvolvimento do primeiro protótipo do BROAD-RSI exigiu o uso de diferentes tecnologias e soluções Web. Na figura 15 é detalhada a arquitetura apresentada no capítulo anterior e são apresentadas as principais tecnologias empregadas em cada módulo. A solução proposta está baseada na linguagem de programação e plataforma computacional Java.

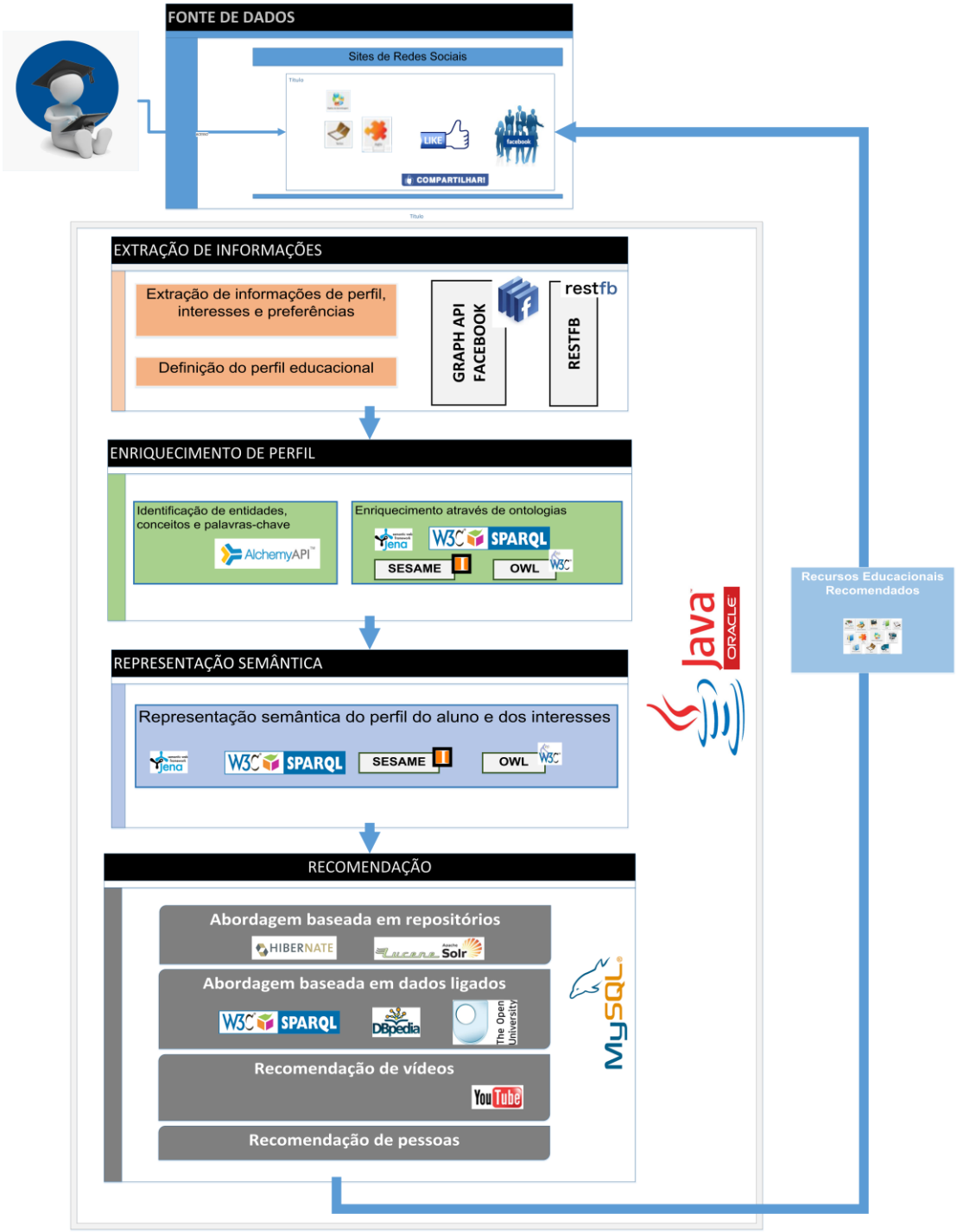


Figura 15 - Representação da arquitetura do BROAD-RSI (Tecnologias).

5.1 EXTRAÇÃO DE INFORMAÇÕES

Para acessar as informações do usuário no Facebook foi usada a API fornecida pelo Facebook para desenvolvedores de aplicativos, chamada GraphAPI[19], na versão 2.0. Essa

[19] https://developers.facebook.com/docs/graph-api?locale=pt_BR

API é uma plataforma para a criação de aplicativos e está disponível para os membros dessa rede social. Ela permite aos aplicativos usar as conexões sociais e informações de perfil para o desenvolvimento de aplicações próprias com diferentes propósitos. Todos os aplicativos desenvolvidos estão sujeitos a aceitação dos termos de privacidade. A GraphAPI usa o protocolo RESTful (CORPORATION, 2013) e as respostas são dadas em formato XML. Algumas das informações disponibilizadas através da GraphAPI podem ser vistas na figura 16.

A exploração da GraphAPI foi feita através da biblioteca chamada RestFB[20] (Allen, n.d.). A escolha dessa biblioteca foi feita por ela ser escrita em Java, possuir ampla documentação, contendo exemplos de utilização, ter licença *Open Source* e ser muito utilizada por aplicativos desenvolvidos em Java que utilizam a API do Facebook.


```
{
  "id": "XXXXXXXXXX",
  "bio": "XXXXXXXX!",
  "birthday": "XX/XX/XXXX",
  "education": [
    {
      "school": {
        "id": "179527005422275",
        "name": "Coluni - UFV"
      },
      "type": "High School",
      "year": {
        "id": "194878617211512",
        "name": "2002"
      }
    },
    {
      "concentration": [
        {
          "id": "112076745475038",
          "name": "Database Design"
        },
        {
          "id": "107626725934132",
          "name": "Computer Science"
        },
        {
          "id": "105595706140624",
          "name": "Software Engineering"
        }
      ],
      "school": {
        "id": "104044756299355",
        "name":   "Universidade   Federal   de
XXXXXXXXXXX "
      },
      "type": "College",
      "year": {
        "id": "140617569303679",
        "name": "2007"
      }
    },
{
      "concentration": [
        {
          "id": "183327321706293",
          "name": "Informática na Educação"
        {
          "concentration": [
            {
              "id": "112195622130706",
              "name": "Data mining"
            }
          ],
          "degree": {
            "id": "410182562425137",
            "name": "Ciência Da Computação"
          },
          "school": {
            "id": "104044756299355",
            "name":    "Universidade    Federal    de
XXXXXXXXXXX"
          },
          "type": "Graduate School",
          "year": {
            "id": "136328419721520",
            "name": "2009"
          }
        }
      ],
"school": {
            "id": "104044756299355",
            "name":    "Universidade    Federal    de
XXXXXXXXXXX"
          },
          "type": "Graduate School",
"year": {
            "id": "142963519060927",
            "name": "2010"
          }
        },
"email": "xxxxxxxxxx@gmail.com",
      "favorite_athletes": [
        {
          "id": "116898848372905",
          "name": "Nova Origem"
        }
      ],
      "languages": [
        {
          "id": "159730814098851",
          "name": "Portuguese"
        },
        {
```


[20] http://restfb.com/

```
        },
        {
          "id": "110383649037339",
          "name": "E-Learning"
        }
      ],
      "school": {
        "id": "104044756299355",
        "name":   "Universidade   Federal   de
XXXXXXXXXXX "
      },
      "type": "Graduate School"
    },
    {
      "concentration": [
        {
          "id": "112105258801226",
          "name": "Project Management"
        }
      ],
          "id": "105820926125566",
          "name": "Espanhol"
        }
      ],
      "locale": "pt_BR",
      "timezone": -2,
      "updated_time": "2014-12-15T16:25:48+0000",
      "verified": true,
      "work": [
        {
          "employer": {
            "id": "109561392399490",
            "name":    "Universidade    Federal    de
XXXXXXXXXXX"
          },
          "position": {
            "id": "211547055525858",
            "name": "ANALISTA DE SISTEMAS -TI"
          }
        }
      ]
    }
```


Figura 16 - Exemplo de informações disponíveis através da API do Facebook.

Neste módulo são extraídas as informações de perfil, os interesses e, ainda, são calculadas duas características contextuais do usuário: dia e horário que ele mais interage e a mídia preferida. A primeira característica é calculada a partir dos campos *created_time* e *updated_time* de cada postagem. Para o cálculo são consideradas as 50 últimas postagens do usuário, sendo que o dia da semana e o horário nos quais o usuário fez o maior número de postagens são escolhidos como seu dia e horário preferidos. O cálculo da segunda característica considera a quantidade de páginas relativas a livros lidos (*book.reads*), exemplificado na figura 17, filmes assistidos (*video.watches*), músicas compartilhadas (*music.listens*), jogos (*games.plays*) e o tipo de mídia usada nas postagens compartilhadas pelo usuário, sendo que aquele tipo de mídia mais postada pelo usuário é selecionado como seu tipo de mídia preferido. Todos os cálculos explicados são atualizados regularmente, sendo adotado nessa proposta o intervalo de sete dias.

```
{
  "data": [
    {
      "id": "814133828624691",
      "from": {
        "id": "100000842445930",
        "name": "xxxxxx"
      },
      "start_time":"2014-12-15T15:12:44+0000",
      "publish_time":"2014-12-15T15:12:44+0000",
      "application": {
        "name": "Books",
        "id": "174275722710475"
      },
      "data": {
        "book": {
          "id": "106160446088057",
          "url":
"https://www.facebook.com/pages/The-Innocent-Man-
Murder-and-Injustice-in-a-Small-
Town/106160446088057",
          "type": "books.book",
          "title": "The Innocent Man: Murder and
Injustice in a Small Town"
        }
      },
"type": "books.reads",
 "no_feed_story": false,
      "likes": {
        "count": 0,
        "can_like": true,
        "user_likes": false
      },
      "comments": {
        "count": 0,
        "can_comment": true,
        "comment_order":
"chronological"
      }
    },
    {
      "id": "803037976400943",
      "from": {
        "id": "100000842445930",
        "name": "xxxxxx"
      },
      "start_time":        "2014-11-
24T02:07:58+0000",
      "publish_time":      "2014-11-
24T02:07:58+0000",
      "application": {
        "name": "Books",
        "id": "174275722710475"
      },
```


Figura 17 - Informações sobre os livros lidos por um determinado usuário.

5.2 ENRIQUECIMENTO DE PERFIL

Conforme mencionado anteriormente, o enriquecimento de perfil é feito a partir de duas abordagens: (1) o enriquecimento através de técnicas de extração de informações (identificação de conceitos, entidades e palavras-chave) usando os textos descritivos das páginas favoritas do usuário e (2) o enriquecimento através de ontologias, buscando interesses relacionados àqueles extraídos.

Na primeira abordagem é usada a *AlchemyAPI*[21], uma ferramenta que oferece um conjunto de funcionalidades sobre o texto fornecido, como análise de sentimentos, extração de *tags*, detecção de idioma, extração de entidades, categorização de texto, extração de conceitos relacionados, extração de palavras-chave, entre outros. A maioria destas funcionalidades usa métodos estatísticos de processamento de linguagem natural (CHOWDHURY, 2003) e algoritmos de aprendizagem de máquina (SEBASTIANI, 2002).

A escolha da *AlchemyAPI* foi feita apoiada em resultados apresentados em (FERNANDES, 2013) e (KOVACIC, 2012) nos quais a *AlchemyAPI* apresentou bons

[21] http://www.alchemyapi.com/

resultados em relação a outras APIs, tais como F-EXT[22] , *Ltasks*[23] e *Boilerpipe*[24]. Além disso, a *AlchemyAPI* possui suporte a diferentes idiomas, fator que foi determinante para a sua escolha, já que, por se tratar de dados extraídos de redes sociais, resultantes da colaboração de pessoas de diferentes nacionalidades, é possível que páginas escritas em diferentes idiomas sejam encontradas.

O uso da *AlchemyAPI* é feito depois de extraídas as conexões do usuário com as suas páginas favoritas. Nesse momento, é feita a filtragem do texto descritivo de cada página e, então, esse texto é submetido aos métodos da *AlchemyAPI* para a detecção do idioma do texto, extração de entidades, conceitos e palavras-chave, conforme esquema mostrado na figura 18.

Figura 18 - Esquema de uso da *AlchemyAPI*.

A detecção de idioma é usada para determinar quais demais funções podem ser aplicadas ao texto, já que algumas delas são restritas a um conjunto limitado de idiomas. A identificação do idioma é relevante ainda para descobrir idiomas de domínio do usuário, caso essa informação não esteja explicitada no seu perfil.

Na figura 19 é apresentado um exemplo de extração de entidades, conceitos e palavras-chave de um determinado texto (mostrado no início da figura 19) sobre o assunto "DBPEDIA". As entidades, palavras-chaves e conceitos extraídos são retornados juntamente com um grau de relevância. O grau de relevância do termo em relação ao texto mostra quão importante é aquele termo dentro do texto analisado. Com o apoio de cinco especialistas, professores de diferentes áreas de conhecimento, foi definido que, para este trabalho, seriam considerados apenas os termos que possuem relevância maior que 0,50, já que eles identificaram que os temas com grau de relevância menor que 0,5 possuíam pouca relação

[22] Site: http://ltasks.com

[23] Site: http://www.opencalais.com

[24] Site: https://code.google.com/p/boilerpipe

com o tema principal.Optou-se pela filtragem buscando priorizar temas de interesse, evitando assim apresentar para o usuário recomendações de temas fracamente relacionados ao seu interesse real.

LANGUAGE: English

DBpedia is the Semantic Data Web interface to Wikipedia. It exposes Wikipedia as a database and provides a live demonstration of power of Structured Web Data (aka. RDF-based Linked Data). DBpedia includes the following: 1. REST-based API (SPARQL) 2. Query Language (SPARQL) 3. Ability to serialize results in a myriad of formats - XML, JSON, RDF, N3 4. Full Text Search Interface 5. Lots of examples!

Click here to learn more about **concepts**.

Visual | JSON | API

- Entities
- Keywords
- Taxonomy
- Concepts
- Document Sentiment
- Targeted Sentiment
- Relations
- Language
- Title
- Author
- Text
- Feeds
- Microformats

Concept	Relevance	Linked Data
Semantic Web	0.944142	dbpedia freebase yago website
World Wide Web	0.657546	dbpedia freebase yago
Serialization	0.619192	dbpedia freebase
Web 2.0	0.57727	dbpedia freebase yago
Web service	0.550802	dbpedia freebase
Java	0.517979	dbpedia freebase yago website

Entity	Relevance	Sentiment	Type	Subtypes	Linked Data
Wikipedia	0.880065	neutral	Company		
Query Language	0.576089	neutral	FieldTerminology		

Keyword	Relevance	Sentiment
RDF-based Linked Data	0.93056	neutral
Semantic Data Web	0.879067	neutral
Data (aka.	0.829531	neutral
API (SPARQL)	0.805341	neutral
Text Search Interface	0.756956	neutral
DBpedia	0.69445	neutral
live demonstration	0.649874	neutral
Query Language	0.636308	neutral
JSON	0.490252	neutral
N3	0.486681	neutral
myriad	0.470562	positive
Wikipedia	0.466059	neutral
Structured	0.456586	neutral

Figura 19 - Extração de entidades, conceitos e palavras-chave usando a AlchemyAPI.

O exemplo de enriquecimento usando o apoio de Técnicas de Extração de Informação pode ser visto na figura 19, onde a princípio era possível saber apenas que o usuário possuía interesse pelo “DBPEDIA” e ainda que existia um texto descrevendo este interesse. No entanto, fazia-se necessário extrair deste texto termos que pudessem identificar melhor esse interesse. A partir da aplicação dos métodos citados, o interesse “DBPEDIA” foi

expandido trazendo áreas como *Semantic Web*, *World Wide Web*, *Serialization*, Web 2.0, *Web Service*, Java, *Resource Description Framework*.

A segunda abordagem para o enriquecimento de perfil faz uso de ontologias. Para trabalhar com as ontologias, criando-as, realizando consultas e manipulando-as foram usados o Apache Jena, o banco de dados RDF Sesame, e a linguagem SPARQL para consultas.

O Apache Jena[25] (ou somente Jena) é um framework Java, *open source*, para a construção de aplicações web semânticas e aplicações de dados ligados. O framework é composto por diferentes APIs que interagem para processar dados RDF (FOUNDATION, 2011).

O Sesame[26] é um framework para armazenamento e consulta de grande quantidade de metadados em RDF e RDF *Schema*, que oferece suporte para armazenamento, controle de concorrência, exportação de dados e consultas feitas em SPARQL. Através dele é possível inserir, apagar, manipular e consultar dados RDF. Os dados RDF de todas as ontologias usadas como apoio para o enriquecimento de perfil, o perfil do usuário e as recomendações feitas são armazenados através do Sesame. Consultas, na linguagem SPARQL, são realizadas para recuperar dados armazenados e interesses inferidos.

As ontologias usadas pelo BROAD-RSI através do protótipo desenvolvido para demonstrar a proposta deste trabalho foram geradas com o apoio de uma aplicação, também desenvolvida no escopo deste trabalho, que extrai os termos e suas dependências a partir da ementa de disciplinas da Universidade Federal de Juiz de Fora. Nas figuras 20 e 21 são mostradas algumas ontologias criadas a partir da ementa de disciplinas.

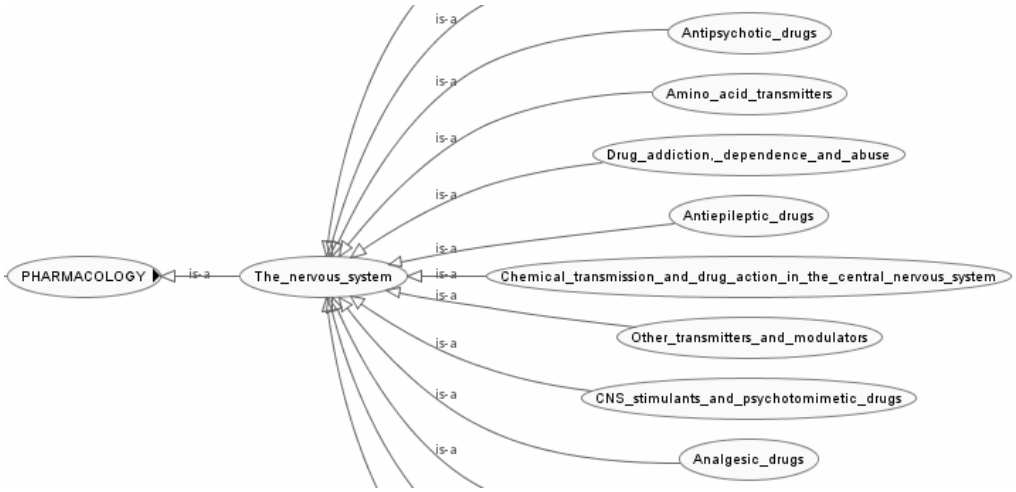


Figura 20 - Ontologias a respeito de conceitos sobre o tema "*Pharmacology*".

[25] https://jena.apache.org/

[26] http://rdf4j.org/

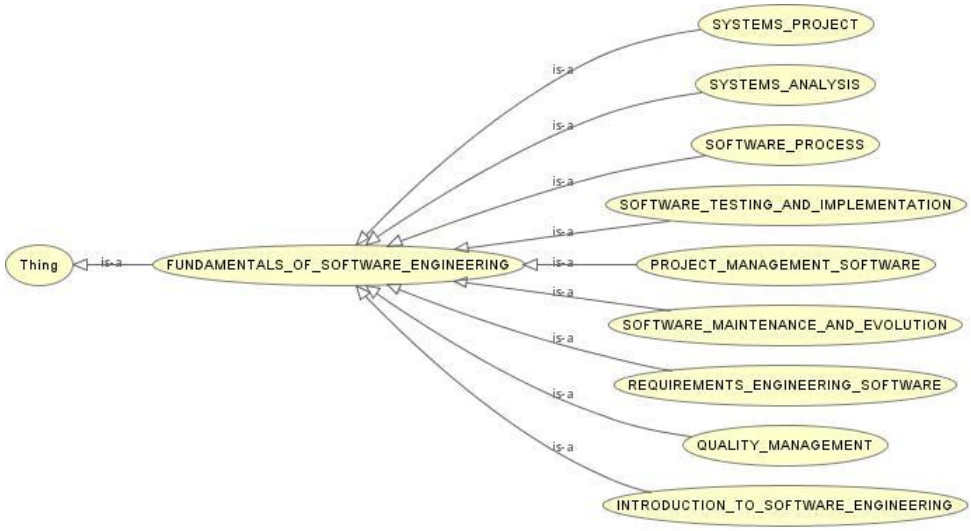
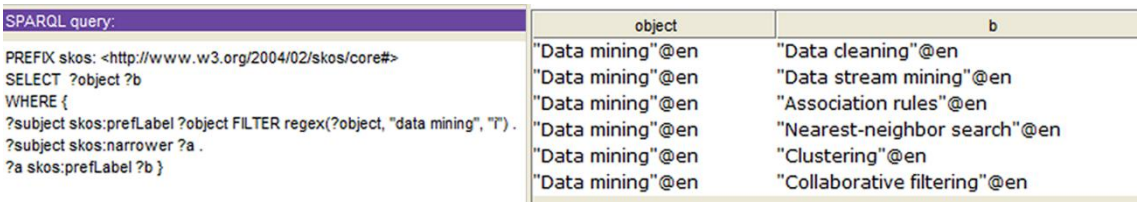


Figura 21 - Ontologias a respeito de conceitos sobre o tema "*Software Engineering*".

Foi utilizada também a taxonomia na área de computação criada a partir do sistema de classificação da ACM[27] (ASSOCIATION FOR COMPUTING MACHINERY, 2012). Ela se baseia em um vocabulário semântico como a única fonte de categorias e conceitos que refletem o estado da arte da área de computação. Um exemplo de consulta SPARQL feita nesta taxonomia pode ser vista na figura 22. Nessa consulta (esquerda) é feito um enriquecimento do termo "*Data Mining*", retornando itens relacionados a esse termo (direita).



SPARQL query:

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT  ?object ?b
WHERE {
?subject skos:prefLabel ?object FILTER regex(?object, "data mining", "i") .
?subject skos:narrower ?a .
?a skos:prefLabel ?b }
```

object	b
"Data mining"@en	"Data cleaning"@en
"Data mining"@en	"Data stream mining"@en
"Data mining"@en	"Association rules"@en
"Data mining"@en	"Nearest-neighbor search"@en
"Data mining"@en	"Clustering"@en
"Data mining"@en	"Collaborative filtering"@en

Figura 22 - Exemplo de SPARQL feita na taxonomia da ACM.

Desta forma, caso um usuário expresse interesse, por exemplo, pela área "*Data Mining*", seja através da ação de determinar este como um tema do seu interesse educacional no seu perfil ou curtir uma página (*https://www.facebook.com/pages/Minera%C3%A7%C3%A3o-de-dados/109281725756837?fref=ts&rf=112195622130706*), então o seu interesse será estendido para os temas mostrados na figura 22 (direita), já que todos eles estão relacionados ao tema. Com esse enriquecimento o usuário não só receberá recomendações a respeito do tema "*Data Mining"*, mas também de todos os temas associados.

[27] http://dl.acm.org/ft_gateway.cfm?id=2371137&ftid=1290922&dwn=1

5.3 REPRESENTAÇÃO SEMÂNTICA

No módulo de representação semântica são usadas as mesmas tecnologias descritas no módulo anterior. Para a representação do perfil e do contexto do usuário foram usados os vocabulários FOAF (BRICKLEY; MILLER, 2014) e SIOC (BRESLIN *et al.*, 2006), com algumas extensões necessárias.

FOAF (*Friend of Friend*) é uma ontologia, que utiliza a tecnologia RDF, para descrever pessoas, seus interesses, e suas ligações em rede social de forma semântica. Todas as características de perfil extraídas do perfil do usuário no Facebook, a ligação desses usuários com outros usuários e os tópicos de interesse são representados usando a FOAF. Algumas propriedades extras foram criadas em virtude de necessidades específicas da arquitetura do BROAD-RSI, tais como a relevância de um interesse inferido em relação a um interesse extraído e link da descrição semântica do item de interesse em repositórios de dados ligados como DBPEDIA e YAGO[28]. Os índices de relevância, bem como os links para a DBPEDIA e YAGO são obtidos como o auxílio da API Alchemy.

A ontologia SIOC (*Semantically-Interlinked Online Communities*), que fornece os principais conceitos e propriedades necessárias para descrever, de forma semântica, informações de comunidades on-line, tais como fóruns, wikis, blogs e grupos, foi usada para representar todos os grupos dos quais o usuário participa. O armazenamento é feito no Sesame. As consultas ao perfil do usuário são feitas através de consultas SPARQL.

Conforme mencionada na arquitetura do BROAD-RSI, as entidades, conceitos e palavras-chave identificadas também dão origem a uma base semântica onde temas de interesses são correlacionados e, posteriormente também podem ser usados para enriquecer os perfis de usuários. Todos os interesses inferidos, seja pela abordagem de enriquecimento por Extração de Informação, seja pela abordagem de enriquecimento por ontologias, são relacionados através de propriedades semânticas, de tal forma que quando um usuário expressa interesse em alguma área é feita também uma consulta na base do BROAD-RSI para buscar interesses que em outro momento tenham sido relacionados e possam, então, enriquecer os interesses do usuário. A geração e armazenamento desta base semântica são feitas usando Jena, Sesame e a linguagem SPARQL para consultas.

[28] http://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/

5.4 RECOMENDAÇÃO

Conforme dito anteriormente, no módulo de recomendação são usadas três abordagens: (1) recomendação baseada em repositórios de objetos de aprendizagem, (2) recomendação baseada em dados ligados e (3) recomendação de vídeos.

5.4.1 **Abordagem baseada em repositórios de objetos de aprendizagem**

Na recomendação de objetos de aprendizagem é estabelecido uma relação entre as características de perfil e do contexto do usuário e os metadados dos recursos educacionais.

5.4.1.1 Passo 1 – Determinação do interesse que será atendido pela recomendação.

Considerando a dinamicidade das interações nas redes sociais, foi adotado no processo de recomendação uma priorização por interesses recentes, já que acredita-se que tão logo o usuário tenha expressado interesse por algum tema, seja mais relevante lhe enviar recomendações a respeito deste tema. Sendo assim, os interesses mais recentemente manifestados através das redes sociais são considerados mais relevantes no envio de recomendações.

A diversificação das recomendações enviadas é outra questão que é tratada no processo de recomendação. É desejável nesta proposta que o usuário receba recomendações sobre os diversos assuntos pelos quais ele tem interesse, não ficando a recomendação restrita a um ou outro interesse. Sendo assim, em resumo, a priorização da recomendação deve seguir dois requisitos:

a. os interesses mais recentes sempre estarão em primeiro lugar na priorização;
b. os interesses para os quais (pelo menos) uma recomendação já tenha sido enviada vão para o fim da lista de priorização.

Para atender tais requisitos o processo de priorização dos interesses segue os seguintes passos:

i. sempre que um novo interesse for detectado ele é armazenado com a data que o usuário o adicionou às suas preferências;
ii. sempre que uma recomendação for enviada, ela é armazenada com a data em que foi enviada ao usuário;

iii. a lista de interesses sempre é ordenada considerando as duas datas: data que o usuário adicionou o interesse à sua preferência e a data do envio da recomendação para aquele item. A ordenação se dará com a data do envio da recomendação em ordem crescente (do mais antigo para o mais novo) e a data que o usuário adicionou o interesse às suas preferências em ordem decrescente (do mais novo para o mais antigo).

Através da execução desses passos é possível garantir que o usuário recebe recomendações sempre para o assunto que ele está mais interessado no momento e ainda que ele sempre é atendido com recomendações para todos os seus interesses.

5.4.1.2 Passo 2 – Filtragem dos objetos de aprendizagem de acordo com um determinado interesse.

A filtragem consiste basicamente de uma busca pelo interesse e pelos temas relacionados à ele na base de metadados que descreve os objetos de aprendizagem.

Para aumentar a eficiência e precisão da busca na base de metadados, é usada uma busca por índice usando o Lucene[29]. O Lucene é uma biblioteca *Open Source* para indexação e consulta de textos. Essa biblioteca trabalha em duas etapas: (1) indexação, onde os dados originais são indexados gerando uma estrutura de dados relacionados para pesquisa baseada em palavras-chave; (2) pesquisa, que é feita no índice criado, por palavras digitadas pelo usuário, gerando resultados ordenados pela similaridade do texto com a consulta. A similaridade do documento é usada na priorização da recomendação.

A geração do índice no Lucene foi feita da seguinte forma: para cada registro do banco de dados, que contém os valores dos metadados, é gerado um *Document Lucene* distinto, composto por cinco campos (*Fields*): identificador, título, descrição, palavras-chave e um campo (Field) não indexado que armazena o endereço de acesso ao documento.

5.4.1.3 Passo 3 – Determinação da aderência do objeto de aprendizagem ao perfil e ao contexto do usuário.

Após a filtragem é feita a etapa de determinação da aderência do objeto de aprendizagem ao perfil e contexto do usuário, onde é verificado se as características desses

[29] http://lucene.apache.org/core/

objetos são aderentes aos itens de perfil e de contexto do usuário. A tabela 7 apresenta as relações de porcentagem para construção das regras para definição dos níveis de aderência. Os valores V (verdadeiro) e F (falso) são atribuídos a partir da acareação entre a característica do usuário e o valor do metadado do objeto de aprendizagem. Por exemplo, se o objeto de aprendizagem é escrito em Português e o usuário possui domínio do idioma Português será atribuído valor V para a relação estabelecida e, assim sucessivamente para os outros metadados.

Tabela 7 – Parte da tabela usada no cálculo da aderência dos recursos educacionais.

INFORMAÇÕES EXTRAÍDAS					
Idiomas falados	**Preferência por mídias**	**Nível de escolaridade**	**Idade**	**Dispositivo**	
V	V	V	V	V	100%
F	V	V	V	V	80%
V	F	V	V	V	80%
V	V	F	V	V	80%
V	V	V	F	V	80%
V	V	V	V	F	80%
V	V	V	F	F	60%
V	V	F	V	F	60%
V	V	F	F	V	60%
F	V	V	V	F	60%
F	V	V	F	V	60%
F	F	V	V	V	60%
V	V	F	F	F	40%
V	F	V	F	F	40%
V	F	F	V	F	40%
V	F	F	F	V	40%
F	V	V	F	F	40%
F	V	F	V	F	40%
F	V	F	F	V	40%
F	F	V	V	F	40%
F	F	V	F	V	40%
F	F	F	V	V	40%
V	F	F	F	F	20%
F	V	F	F	F	20%
F	F	V	F	F	20%
F	F	F	V	F	20%
F	F	F	F	V	20%
F	F	F	F	F	0%

Fonte: elaborada pelo autor.

Depois de calculada a aderência dos recursos educacionais é aplicado o grau de relevância do documento a partir do índice de similaridade resultante da etapa de filtragem usando o Lucene. Por fim, os recursos são ordenados de acordo com o seu grau de aderência e

relevância do documento em relação à busca, respectivamente. Uma demonstração da ordenação dos recursos educacionais pode ser vista na tabela 8.

Tabela 8 - Exemplo de ordenação dos recursos educacionais.

	Interesse	Aderência	Relevância do documento
Recurso Educacional 12	Data Mining	100%	1.657
Recurso Educacional 1	Data Mining	100%	1.643
Recurso Educacional 8	Data Mining	100%	1.643
Recurso Educacional 10	Clustering	80%	1.657
Recurso Educacional 6	Data Warehouse	80%	1.643
Recurso Educacional 15	Data Mining	80%	1.643
Recurso Educacional 3	Data Warehouse	80%	1.537
Recurso Educacional 5	Data Mining	80%	1.085
Recurso Educacional 9	Data Mining	60%	1.986
Recurso Educacional 4	Data Mining	60%	1.786
Recurso Educacional 11	Association rules	60%	1.786
Recurso Educacional 7	Clustering	60%	1.475
Recurso Educacional 13	Data Mining	40%	1.786
Recurso Educacional 14	Association rules	20%	1.643
Recurso Educacional 2	Data Mining	0%	1.657

Fonte: elaborada pelo autor.

A visão geral do processo de recomendação do BROAD-RSI é apresentada na figura 23.

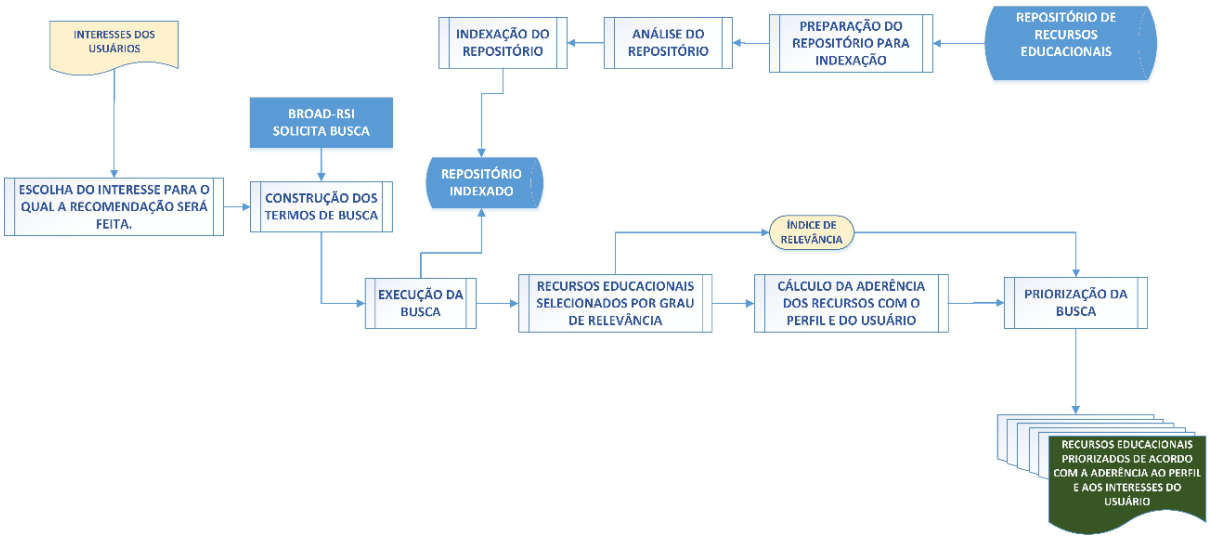


Figura 23 - Visão geral do processo de recomendação do BROAD-RSI.

Todo o processo descrito é usado no caso em que as recomendações são enviadas automaticamente para a rede social do usuário. Caso ele queira consultar as recomendações disponíveis para um determinado interesse, o passo 1 é ignorado, já que o próprio usuário já determinou qual o tema para o qual ele deseja ver as recomendações. Os passos 2 e 3 permanecem sendo adotados.

5.4.2 Abordagem baseada em Dados Ligados

Nessa abordagem foram usadas duas fontes de Dados Ligados como fonte de busca de recursos educacionais: o *dataset* da DBPEDIA e o *dataset* da Open University.

5.4.2.1 Passo 1 – Determinação do interesse que será atendido pela recomendação

O passo 1 descrito no item anterior é válido também para esta abordagem.

5.4.2.2 Passo 2 - Filtragem dos recursos educacionais de acordo com um determinado interesse.

i. Dataset da DBPEDIA

Depois de determinado o tema para o qual a recomendação será feita, o próximo passo para a extração de recomendações educacionais através da DBPEDIA é a busca por páginas relacionadas ao tema de interesse selecionado no passo anterior. Para isso é feita uma busca usando a propriedade *rdfs:label*. Os rótulos (*labels*) de um recurso da DBPEDIA são criados a partir dos títulos das páginas da Wikipédia. Na figura 24 é apresentado um exemplo de consulta por um recurso relacionado ao interesse "*Data Mining*".

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX rdfs: < http://www.w3.org/2000/01/rdf-schema#>
SELECT ?area where {?area rdfs:label ?t . FILTER (lcase(str(?t)) = 'data
mining')}";
```

Figura 24 - SPARQL usada na recuperação de recursos na DBPEDIA.

Uma vez realizada a consulta, o serviço da DBPEDIA[30] retorna todos os recursos cujo rótulo contenha o termo indicado. A página da DBPEDIA referente ao recurso "Data Mining" é mostrada na figura 25.

Após recuperado o recurso, o BROAD-RSI é responsável por realizar outras consultas por propriedades específicas que possam trazer conteúdos educacionais relevantes ao tema de interesse. Dentre as propriedades disponíveis, são utilizadas as seguintes: *dbpedia-owl:abstract, dbpedia-owl:wikiPageExternalLink, foaf:isPrimaryTopicOf, dbpedia-owl:academicDiscipline, dbpprop:discipline of, dbpedia-owl:literaryGenre, dcterms:subject, dbpprop:hasPhotoCollection.* A escolha dessas propriedades foi feita após uma pesquisa a

[30] http://dbpedia.org/sparql

respeito de quais propriedades poderiam retornar um conteúdo educacional relevante para o usuário. No entanto, não se pode afirmar que a lista de propriedades, cujo retorno possa trazer recursos educacionais, tenha sido inteiramente considerada.

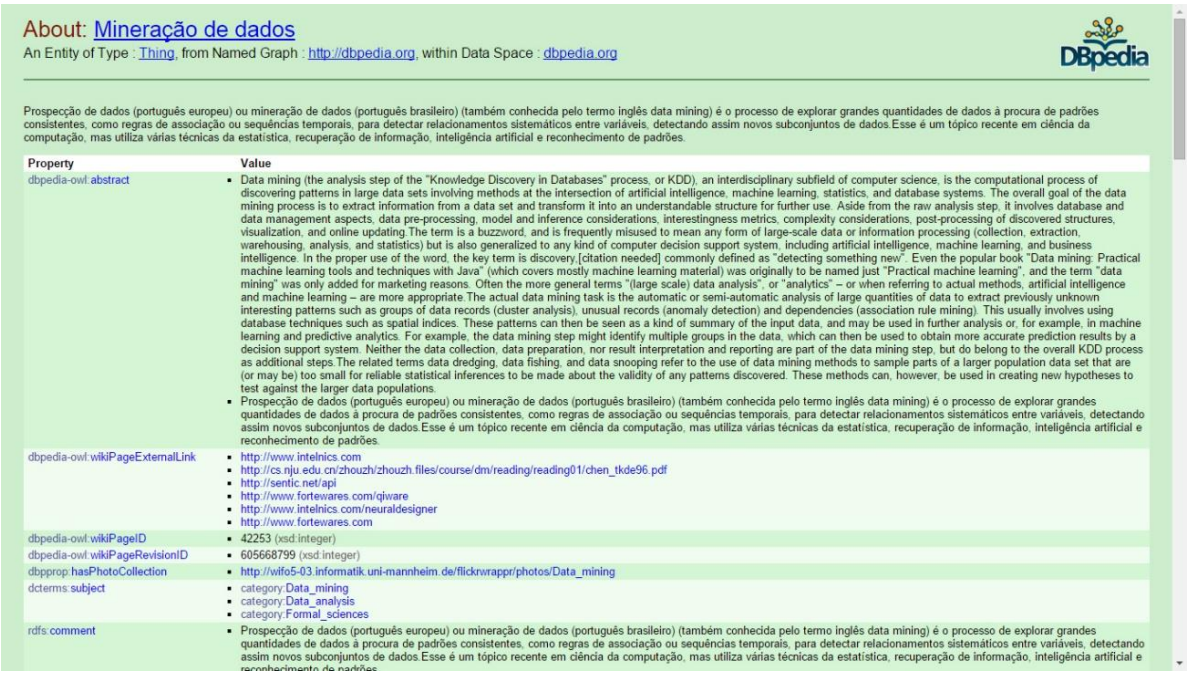
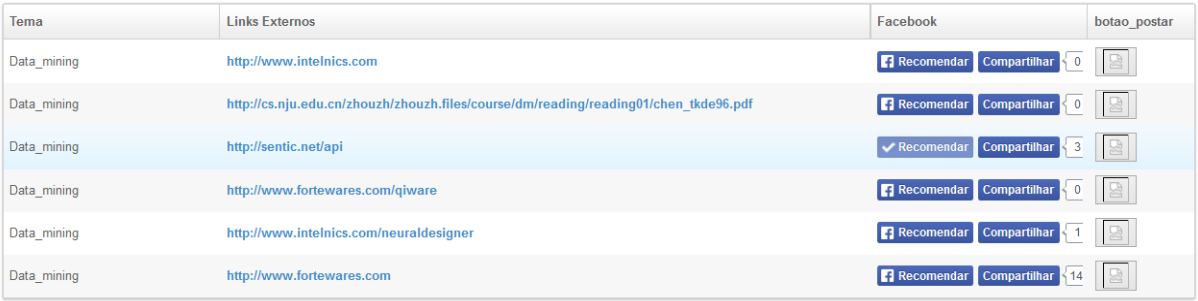


Figura 25 - Página da DBPEDIA referente ao tema "*Data Mining*".

Usando essas propriedades, são recuperadas e apresentadas para o usuário algumas informações que podem contribuir com o seu aprendizado, como, por exemplo, uma breve explicação sobre o assunto, links para conferências, periódicos e outras páginas diversas que estejam associadas ao tema de interesse. Na figura 26 é mostrada uma tela que é apresentada para o usuário com links sobre o tema "*Data Mining*".



Tema	Links Externos	Facebook	botao_postar
Data_mining	http://www.intelnics.com	Recomendar Compartilhar 0	
Data_mining	http://cs.nju.edu.cn/zhouzh/zhouzh.files/course/dm/reading/reading01/chen_tkde96.pdf	Recomendar Compartilhar 0	
Data_mining	http://sentic.net/api	Recomendar Compartilhar 3	
Data_mining	http://www.fortewares.com/qiware	Recomendar Compartilhar 0	
Data_mining	http://www.intelnics.com/neuraldesigner	Recomendar Compartilhar 1	
Data_mining	http://www.fortewares.com	Recomendar Compartilhar 14	

Figura 26 - Links recuperados através da DBPEDIA.

ii. *Dataset* da Open University

Um processo semelhante é utilizado para a busca de recursos educacionais no *dataset* da Open University. No entanto, nesse caso é possível usar a propriedade *rdfs:type* para agrupar os recursos apresentando-os ao usuário em grupos, tais como Livros, *Podcasts*, Documentos em geral, Vídeos, Cursos, etc. Na Figura 27 é mostrada a consulta SPARQL

executada no serviço da *Open University*[31] para recuperar recursos do tipo *Book* para o tema "*Project Management*".

```
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX dbpediap: <http://dbpedia.org/property/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT DISTINCT ?s ?t ?q WHERE {?s dcterms:description  ?t . FILTER
regex(lcase(str(?t)), 'project management')?s dbpediap:url ?q . ?s
rdf:type <http://purl.org/ontology/bibo/Book>}
```

Figura 27 – SPARQL executada no *Dataset* da DBPEDIA.

Os recursos recuperados trazem um conjunto de outros dados ligados que podem ser úteis ao usuário. Por exemplo, um recurso do tipo *Book* retornado pode ser parte de um curso ou estar associado a outro recurso de outro tipo e, dessa forma, despertar o interesse em conhecer conteúdos que vão além dos recuperados pelo BROAD-RSI.

5.4.2.3 Passo 3 – Determinação da aderência do objeto de aprendizagem ao perfil e ao contexto do usuário.

Tanto no caso da DBPEDIA, quanto no caso da Open University, existe uma grande dificuldade para priorizar os recursos retornados para cada tipo de interesse, já que os mesmos não possuem uma catalogação que possa ser usada para determinar quais são mais ou menos relevantes para os usuários. Esse fato dificulta o envio da recomendação automaticamente para o usuário e, por isso, optou-se por deixar essa funcionalidade disponível apenas para as solicitações manuais de recomendações feitas pelos usuários.

5.4.3 **Abordagem baseada em vídeos**

5.4.3.1 Passo 1 – Determinação do interesse que será atendido pela recomendação

O passo 1 descrito nos itens anteriores é válido também para esta abordagem.

5.4.3.2 Passo 2 – Filtragem dos recursos educacionais de acordo com um determinado interesse.

No passo 2 é feita a busca de vídeos com o apoio da API do canal de vídeos

[31] http://data.open.ac.uk/query

Youtube[32]. Para tal fez-se necessário registrar o aplicativo BROAD-RSI no Youtube, que a partir disso, retorna um código que deve ser usado pelo aplicativo para enviar solicitações.

A busca por vídeos foi feita apenas através de palavras-chave, indicando quantos vídeos devem ser retornados. No momento da configuração da busca, é possível informar quais serão os campos considerados na busca. Foram escolhidos: identificador, tipo, título, URL e descrição do vídeo.

A relevância dos vídeos é determinada pela própria API. Neste caso não são consideradas outras características de perfil do usuário para priorização dos recursos, já que não existem metadados definidos para comparação nos vídeos do Youtube. As opções de "curtir", "compartilhar" e "recomendar" também estão disponíveis para cada um dos vídeos recomendados. A forma que os vídeos são apresentados para o usuário pode ser vista na Figura 28.

No caso dos vídeos, a ordem de envio da recomendação considera a relevância dos vídeos retornados pela própria API. Neste caso, o vídeo mais relevante será o primeiro a ser enviado.

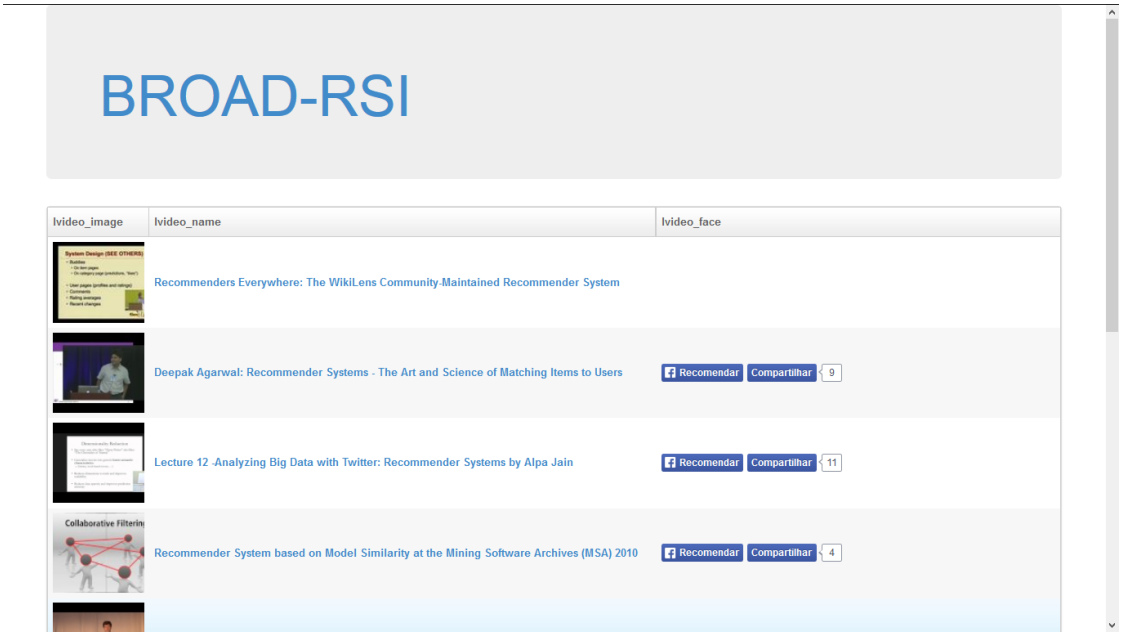


Figura 28 - Recomendação de vídeos do Youtube.

5.4.4 Recomendações de pessoas

A recomendação de pessoas consiste em sugerir a um determinado usuário outras pessoas que possuem interesses similares aos seus. Para isso são considerados todos os

[32] https://www.youtube.com/yt/dev/pt-BR/api-resources.html

interesses do usuário para o cálculo de similaridade, sendo que quanto mais interesses comuns os usuários têm, mais similares eles são. A recomendação de pessoas é feita por área de interesse, no entanto, para cálculo da similaridade são considerados todos os interesses. Por exemplo, caso o usuário **A** peça uma recomendação de pessoas que possuem interesse na área de Gerenciamento de Projetos, serão apresentados a ele outras pessoas que possuem tal interesse e, para cada uma dessas pessoas, será mostrado um valor numérico que determina o grau de similaridade com o usuário. Para este cálculo são considerados todos os interesses comuns, conforme mostra a fórmula abaixo: seja **A** o usuário que solicitou a recomendação e **B** o usuário recomendado, temos que:

$$Grau\ de\ similaridade = \frac{N^{\underline{o}}\ de\ interesses\ comuns\ entre\ \boldsymbol{A}\ e\ \boldsymbol{B}}{\text{Nº de interesses de } \mathbf{A}}$$

A rede de usuários com interesses comuns é formada considerando todos os usuários do BROAD-RSI funcionando da seguinte forma: caso o usuário A queira ver a rede de interesses em torno do tema "Engenharia de software" será feita uma busca semântica por todos os usuários com interesse nesse tema, que já são amigos dele na rede social Facebook. Logo após, são consultados os demais usuários com interesse em Engenharia de software que não estão diretamente ligados ao usuário A. Tendo todos os usuários com o mesmo interesse reconhecidos, são identificadas as conexões entre eles de tal forma que seja possível mostra ao usuário A como ele pode alcançar todos as pessoas que possuem interesse em Engenharia de Software a partir dos seus contatos na rede social.

A implementação da recomendação de pessoas foi feita com o apoio de um conjunto de ferramentas de software chamadas Prefuse[33], que é uma arquitetura para a visualização de grafos e árvores desenvolvida na linguagem de programação Java. Ela possui um conjunto de recursos para modelagem de dados, visualização e interação. Sendo que, neste trabalho, ela foi usada na visualização das redes conforme mostrado anteriormente na Figura 12.

5.4.5 Armazenamento e entrega da recomendação aos usuários

A recomendação é enviada ao usuário no Facebook, através da própria API disponibilizada por essa rede social. Um trecho do envio pode ser visto na Figura 29. Após realizadas todas as etapas descritas são determinadas duas mensagens, que serão enviadas ao

[33] http://prefuse.org/

usuário na sua página no Facebook: uma contendo um resumo do tema de interesse (extraído da DBPEDIA) e um objeto de aprendizagem referente ao tema (extraído do repositório) e outra com um vídeo (extraído do Youtube). As recomendações são enviadas sempre no dia e horário de maior acesso do usuário.

```
public void publishRecommendation(String id_usuario, String interesse, String recomendacao){
    facebookClient  = new DefaultFacebookClient("[illegible]", "↵
[illegible]"){
    @Override protected String getFacebookGraphEndpointUrl() { return "https://graph.facebook.com/v2.0"; }
};
    FacebookType publishMessageResponse = facebookClient .publish(id_usuario+"/feed", FacebookType.class,Parameter.↵
with("message", "Veja esta recomendação que temos para você sobre "+ interesse+". Avalie a recomendação em http://goo.gl/↵
forms/pFYFeiQMuK"), Parameter.with("link", recomendacao));
}
```

Figura 29 - Trecho do código para envio da recomendação para o Facebook.

Após o envio de qualquer recomendação, a mesma é retirada da lista de possíveis itens a serem recomendados, garantindo assim que o mesmo usuário não receberá recomendações duplicadas.

Caso o usuário queira consultar recomendações sobre qualquer um dos temas do seu interesse, ele poderá, a qualquer momento, acessar a página principal do aplicativo e consultar recomendações por temas. Todas as recomendações recebidas pelos usuários podem ser avaliadas por eles através do link que acompanha a recomendação, no caso do envio automático, ou através do próprio aplicativo BROAD-RSI, no caso de consultas manuais. Esta avaliação embasará futuras recomendações e melhorias futuras no próprio sistema. O histórico de todas as recomendações é armazenado em um banco de dados MySQL. Este armazenamento visa, principalmente, impedir que recomendações sejam enviadas em duplicidade ao usuário, mas também registrar a avaliação do usuário para cada recurso recomendado e criar um histórico que poderá contribuir com o aprimoramento do protótipo em versões futuras.

5.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo foram detalhados os aspectos relacionados às tecnologias usadas no desenvolvimento do protótipo BROAD-RSI. Para analisar a viabilidade da proposta, definir quais as tecnologias necessárias para atingir os objetivos propostos e permitir a avaliação da proposta, foi desenvolvido um protótipo, no qual as etapas detalhadas no capítulo 4 foram implementadas.

O processo de definição de tecnologias passou por inúmeras etapas e análise de diferentes alternativas para obter os melhores resultados, sendo que alguns princípios nortearam as escolhas feitas, tais como, uso de ferramentas e serviços gratuitos e escolha de

opções escaláveis, que possam atender um grande número de usuários.

A exploração de recursos educacionais oriundos de datasets como a DBPEDIA e *Open University* foi definida como uma iniciativa de usar dados ligados para possibilitar uma ampliação da proposta do BROAD-RSI sem alterações impactantes, já que, geralmente, os dados ligados usam vocabulários e propriedades semânticas comuns, permitindo assim incluir novos *datasets,* usando as mesmas propriedades semânticas na recuperação dos recursos. A mesma possibilidade de ampliação é válida para repositórios de recursos educacionais que usam os metadados do padrão LOM para a sua catalogação, já que as regras de recuperação dos recursos estão definidas para esse padrão.

A implementação da recomendação de pessoas mostrou ser viável associar pessoas em torno de um interesse comum, além de calcular quão similares são os usuários em relação aos seus interesses educacionais. A recomendação de pessoas dentro do BROAD-RSI também explora os recursos das redes sociais, como por exemplo a possibilidade de seguir a pessoa recomendada ou visualizar o seu perfil na rede social Facebook. O conjunto de pessoas recomendadas é composto apenas por usuários do BROAD-RSI por questões de privacidade dos dados, já que faz-se necessário o usuário permitir a extração das suas informações, através do login no BROAD-RSI, para, então, ser possível ter acesso aos seus interesses e realizar as associações e cálculos necessários.

6 AVALIAÇÃO

Ao longo deste capítulo são descritas as avaliações realizadas na proposta de recomendação de recursos educacionais do BROAD-RSI, usando informações geradas espontaneamente nas redes sociais dos usuários. A definição e o planejamento das avaliações das hipóteses formuladas nessa dissertação são apresentados a seguir. A avaliação da viabilidade da proposta se deu por três provas de conceito e um estudo de caso.

Para verificar as hipóteses apresentadas foi desenvolvido um protótipo, conforme detalhado no capítulo 5, cujos objetivos são: (1) extrair características de perfil e de contexto do usuário, através da rede social Facebook; (2) enriquecer os interesses educacionais extraídos; (3) exibir aos usuários recomendações educacionais aderentes ao seu perfil e ao contexto; (4) apoiar o processo de avaliação da proposta por parte do usuário, ou seja, possibilitar que os usuários avaliem as características de perfil e de contexto extraídas e as recomendações feitas dentro do próprio protótipo.

Após o desenvolvimento do protótipo, foram realizadas três provas de conceitos (*Proof of Concept – PoC*), com objetivos específicos, que serão detalhados a seguir. O objetivo de cada prova de conceito, bem como o do estudo de caso, foram definidos de acordo com a abordagem *Goal/Question/Metric* (GQM) (BASILI, 1992). Segundo a abordagem GQM, os objetivos devem ser definidos conforme o *template* a seguir:

- *"**Analisar o** <objeto do estudo> **com a finalidade de** <objetivo> **com respeito à** <foco da qualidade> **do ponto de vista de** <perspectiva> **no contexto de** <contexto>".*

Depois da realização das provas de conceito, algumas adaptações foram feitas no protótipo, objetivando ajustar o instrumento de avaliação desenvolvido e notou-se ainda a necessidade de aplicação de um questionário aos usuários com o intuito de esclarecer algumas questões que surgiram na aplicação da prova de conceito. Nesse sentido, posteriormente à realização das três provas de conceito, um estudo de caso também foi conduzido.

6.1 PROVAS DE CONCEITO

A opção por iniciar a avaliação da proposta apresentada nesta dissertação usando uma prova de conceito justifica-se, neste contexto, pela necessidade de demonstrar a viabilidade técnica do modelo proposto, dos conceitos e das tecnologias envolvidas no projeto. Apesar de não apresentar o formalismo de um estudo experimental, a utilização de

provas de conceitos neste domínio contribuiu também para a avaliação preliminar das hipóteses formuladas, o ajuste do instrumento de avaliação, a amenização de dificuldades técnicas pontuais e o refinamento dos passos que foram, posteriormente, aplicados no estudo de caso.

6.1.1 Escolha dos participantes

As provas de conceito foram realizadas com usuários reais, identificados pelos pseudônimos: Pesquisador A e Pesquisador B, ambos estudantes do programa de mestrado em Ciência da Computação da UFJF. Os dois usuários indicaram utilizar a rede social Facebook com frequência diária, estando mais de 4 horas por dia conectados, tendo a mesma como uma de suas fontes de informações sobre assuntos diversos, inclusive buscando através dela informações que contribuem com a sua formação intelectual.

Apesar da diversidade de usuários das redes sociais e o aspecto não restritivo do modelo proposto, que permitiria a realização das provas de conceito em um contexto mais amplo, fez-se necessário restringir o escopo de realização das provas de conceito a usuários com interesses relacionadas à área de Computação, uma vez que o repositório de objetos de aprendizagem usado como fonte para as recomendações contém recursos específicos catalogados para essa área.

6.1.2 Coleta dos dados

Durante a realização das provas de conceito, os usuários foram convidados a responder algumas perguntas com o objetivo de avaliar diferentes questões, que serão detalhadas na apresentação de cada uma das três provas de conceito. A escala usada na coleta dos dados foi a escala de Likert. A Escala Likert mede atitudes e comportamentos utilizando opções de resposta que variam de um extremo a outro. Ao contrário de uma simples pergunta de resposta "sim ou não", uma Escala Likert permite descobrir níveis de opinião. A escala de Likert é definida em (MCIVER; CARMINES, 1981) como um conjunto de itens, composto por aproximadamente um número igual de declarações favoráveis e desfavoráveis relativas ao objeto de estudo. Um grupo de indivíduos é convidado a responder cada afirmação de acordo com o grau de concordância ou discordância. Normalmente, eles são instruídos a escolherem uma das cinco respostas: concordo totalmente, concordo, indeciso, discordo ou discordo totalmente.

No caso do questionário aplicado nas provas de conceito e no estudo de caso, a legenda usada foi concordo plenamente, concordo parcialmente, não concordo nem discordo, discordo parcialmente, discordo plenamente; sendo associada aos valores numéricos 2, 1, 0, -1,-2, respectivamente.

6.1.3 Primeira PoC: características de perfil e de contexto do usuário

O objetivo da primeira prova de conceito é definido por:

- ***Analisar*** *a extração de características de perfil e de contexto do usuário* ***para o propósito de*** *identificar e avaliar as informações sobre o perfil e os interesses educacionais do usuário* ***com relação à*** *corretude das características de perfil e interesses extraídos d****o ponto de vista do*** *usuário n****o contexto de*** *usuários da rede social Facebook, que possuem interesses relacionados à área de Computação e expressam tais interesses através das suas interações nesta rede social.*

O Pesquisador A e o Pesquisador B estão interessados em se manterem atualizados em relação a temas abordados nas comunidades científicas relacionadas às suas respectivas pesquisas. Para isso, ambos participam de diversas comunidades e grupos em redes sociais, além de curtir e compartilhar páginas relacionadas a temas de seus interesses.

Os usuários Pesquisador A e Pesquisador B decidiram utilizar o BROAD-RSI para terem acesso a recomendações educacionais aderentes ao seu perfil e aos seus interesses educacionais.

Ambos os usuários deram as permissões necessárias ao aplicativo do BROAD-RSI para que fossem feitas as devidas extrações para a identificação de características de perfil e do contexto e, logo após, foram direcionados à uma página onde foi apresentado um conjunto de características dos seus respectivos perfis e alguns dos seus interesses são identificados.

A partir da extração das informações disponíveis na rede social Facebook e do processamento feito pelo BROAD-RSI para filtrar informações úteis para o contexto de recomendações educacionais, foi apresentado para o Pesquisador A o conjunto de características de perfil apresentado na figura 30.

Característica	Valor	Avaliação
Nome	[illegible]	
Email	[illegible]	
Idade	31	
Cidade	[illegible]	
Gênero	Feminino	
Dia de maior acesso	Domingo	
Horário de maior acesso	11	
Nível Educacional	Pós Graduação	
Mídia	Impressa	
Idioma	Portuguese	
Idioma	Espanhol	

Figura 30 - Características de perfil do Pesquisador A.

Através das características de perfil extraídas foi possível obter informações relevantes para a seleção de recursos educacionais. Por exemplo, em uma futura recomendação para o Pesquisador A, podem ser priorizados recursos educacionais textuais, escritos nos idiomas português ou espanhol, cuja dificuldade seja coerente a cursos de pós-graduação. A recomendação pode ser enviada ao perfil do usuário na rede social uma vez por semana, sendo domingo o dia escolhido, por volta das 11h da manhã, já que as informações extraídas dão indicativos de que nesse dia e horário o usuário provavelmente estará conectado e poderá receber a recomendação enquanto interage na rede social.

O Pesquisador B, após acessar pela primeira vez o BROAD-RSI e dar as permissões devidas também tem um perfil inicial traçado, conforme mostrado na figura 31.

Característica	Valor	Avaliação
Nome	[illegible]	
Email	[illegible]	
Idade	29	
Cidade	Juiz de Fora	
Gênero	Masculino	
Dia de maior acesso	Terça-feira	
Horário de maior acesso	8	
Nível Educacional	Pós Graduação	
Mídia	Audio	

Figura 31 - Características de perfil do Pesquisador B.

Para o Pesquisador B já existem indícios de que seriam mais apropriados para futuras recomendações os recursos educacionais no formato de áudio, que também possuem linguagem e conteúdo adaptados aos cursos de pós-graduação e que o melhor dia e horário para enviá-las ao usuário é às terças-feiras no horário de 8h da manhã. Ele não explicitou em seu perfil quais são os idiomas que ele domina, no entanto, através da identificação dos idiomas de páginas curtidas e compartilhadas é possível obter indícios a respeito dessa informação.

Além das características de perfil são extraídos pelo BROAD-RSI interesses educacionais e interesses gerais dos usuários. O primeiro passo da descoberta de interesses do usuário é analisar todas as suas páginas favoritas e extrair delas quais são as áreas abordadas. Para isso, foram usadas técnicas de Extração de Informação com o objetivo de categorizar as páginas. O resultado da categorização de todas as páginas do usuário Pesquisador A foi a descoberta de três grandes áreas de interesse conforme mostrado na figura 32.

Interesses Gerais	/science/computer_science/artificial_intelligence
Interesses Gerais	/technology_and_computing
Interesses Gerais	/technology_and_computing/software

Figura 32 - Grandes áreas de interesse do Pesquisador A.

Após a descoberta de interesses gerais, que representam a categorização de todos os interesses do Pesquisador A, são extraídos os interesses educacionais específicos para cada uma das páginas favoritas, conforme mostrado na figura 33. Esses interesses serão a base para

a seleção dos temas dos recursos educacionais que podem interessar ao usuário em futuras recomendações.

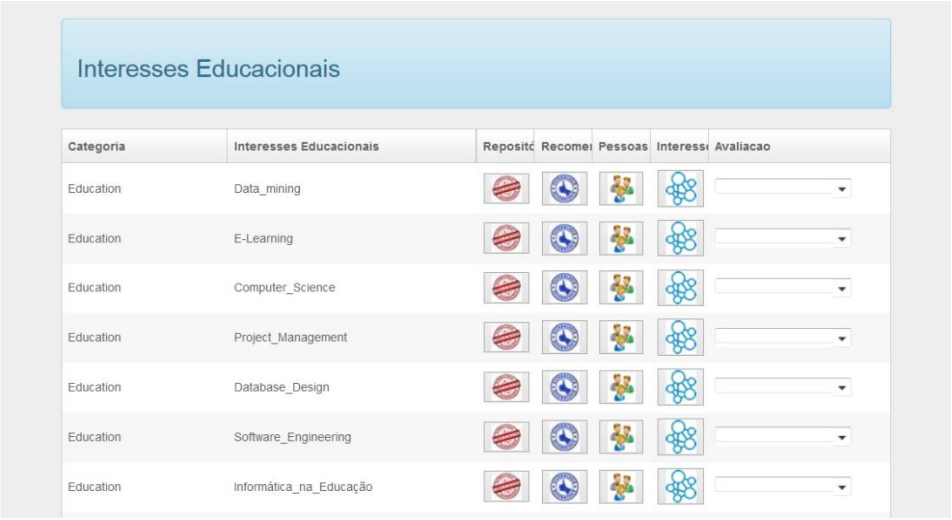


Figura 33 - Interesses especificamente específicos do Pesquisador A.

Além dos interesses especificamente educacionais do Pesquisador A, é possível identificar outros interesses diversos, filtrados por categorias pré-selecionadas. Cada página no Facebook possui uma categoria ou um conjunto de categorias. A seleção das páginas a serem consideradas na extração de interesses é feita com base nas suas categorias associadas, onde são desconsideradas aquelas que não possuem nenhuma relação com assuntos educacionais. Na figura 34 são mostrados alguns dos interesses diversos do Pesquisador A.

Figura 34 - Interesses diversos do Pesquisador A.

O mesmo processo de identificação de interesses é feito para o Pesquisador B. Na figura 35 são mostradas as grandes áreas de interesse do Pesquisador B, obtidas através da categorização das páginas favoritas.

Interesses Gerais	/law,_govt_and_politics/government/courts_and_judiciary
Interesses Gerais	/technology_and_computing

Figura 35 - Grandes áreas de interesse do Pesquisador B.

Na figura 36 são apresentados os interesses especificamente educacionais extraídos. Já na figura 37 podem ser visualizados os interesses diversos do Pesquisador B, seguindo a mesma filtragem por categorias descrita para o Pesquisador A.

Interesses Educacionais

Categoria	Interesses Educacionais	Repositóric	Recomend	Pessoas	Interesses	Avaliacao
Education	Network_Computing					
Education	Wireless_network					

Figura 36 - Interesses especificamente educacionais do Pesquisador B.

Interesses Diversos

Categoria	Interesses Educacionais	Repositóri	Recomend	Pessoas	Interesses	Avaliacao
Interest	IPV6					
technology	OPEN_COMPUTE_J					
Magazine	O_GIBRA_INDICA					
software	CLOUD_MARKET					
Magazine	INTERNET_OF_THINGS					
Non-profit organization	IEEE_CLOUD_COMPUTING					
internet website	WORKSHOP_DE_SEGURANÇA_DA_ _SEGINFO					
Interest	COMPUTER_NETWORK					
School	IFF_-_CAMPUS_BOM_JESUS_DO_I					
software	CLOUD_FOUNDRY					
technology	_CLOUD_COMPUTING_CONCEPTS_					
software	INTERNET_OF_THINGS_BASED_ON					

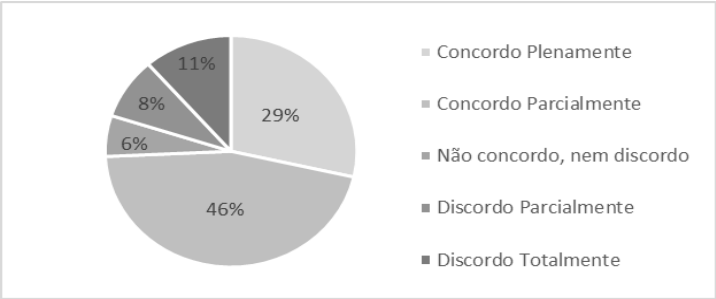
Figura 37 - Interesses diversos do Pesquisador B.

6.1.3.1 Análise quantitativa das características de perfil e de contexto do usuário

Através da extração de informações do Pesquisador A são detectados 26 possíveis interesses educacionais. O pesquisador A é convidado a avaliar tais interesses, respondendo, para cada um dos interesses, a seguinte pergunta: este tema de interesse identificado através das suas interações na rede social Facebook representa de fato um interesse educacional para você?

As respostas do Pesquisador A são apresentadas no gráfico 1. Se somadas as avaliações onde o usuário concordou plenamente (29%) e parcialmente (46%) com os interesses extraídos, temos um total de 75% dos interesses coerentes com os reais interesses educacionais do usuário. O Pesquisador A indicou discordar totalmente de 11% e discordar parcialmente de 8% dos interesses. Ele ainda se declarou neutro em relação a 6% deles.

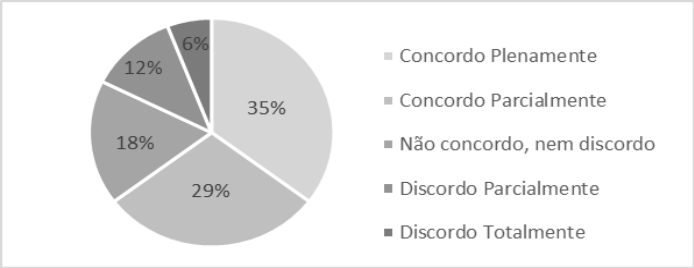
Gráfico 1 - Avaliação dos interesses extraídos do Pesquisador A.



Fonte: elaborado pelo autor.

No perfil do Pesquisador B foram identificados 17 interesses. Ele foi convidado a responder a mesma questão colocada para o Pesquisador A para cada um dos temas. O resultado dessa avaliação é apresentado no gráfico 2. Se somados os interesses em que ele concordou totalmente (35%) e concordou parcialmente (29%), tem-se o total de 64% dos temas coerentes com os seus interesses educacionais. O Pesquisador B discordou parcialmente, discordou totalmente e se manteve neutro em respectivamente 12%, 6% e 18% dos temas.

Gráfico 2 - Avaliação dos interesses extraídos do Pesquisador B.



Fonte: elaborado pelo autor.

Em relação às características de perfil extraídas tanto o usuário Pesquisador A, quanto o usuário Pesquisador B concordaram plenamente com todos os itens extraídos.

6.1.3.2 Evidências observadas na primeira PoC

Esta primeira prova de conceito mostra a viabilidade de extrair informações relacionadas aos usuários através de interações na rede social Facebook, filtrar tais informações de forma que sejam identificadas características de perfil e interesses educacionais. Dessa forma, tem-se indícios de que a primeira hipótese apresentada nesta dissertação pode ser confirmada, já que a extração de informações, geradas espontaneamente a partir da interação dos usuários em redes sociais, permitiu identificar características de perfil e de contexto do usuário.

Vale observar que mesmo a prova de conceito sendo realizada em um grupo de estudantes da pós-graduação em Ciência da Computação, foi possível identificar interesses relacionados a outras áreas, como, por exemplo, no caso do Pesquisador A que possui interesses em temas como História do Mundo, História do Brasil, Geografia e Fotografia. Tais interesses dificilmente seriam identificados em um sistema educacional convencional, como ambientes virtuais de aprendizagem e sistemas acadêmicos, por exemplo. Isso indica que se o repositório do aplicativo BROAD-RSI for expandido, dispondo de recursos educacionais em áreas diversas, seus usuários poderão ser atendidos nas mais diversas áreas de conhecimento.

A partir da avaliação feita pelos dois usuários foi possível perceber que a identificação de interesses educacionais obteve um bom percentual de aceitação. No entanto, fez-se importante observar quais fatores contribuíram para a rejeição de alguns dos interesses. A observação dos itens que não representam o interesse do usuário trouxe indícios de que as categorias que possuem maior abrangência, como por exemplo, Comunidades, Livros, Revistas, Website culturais, Universidades, entre outras trouxeram temas distantes e não representativos para interesses educacionais, apesar de estarem, em alguns casos, associadas a temas educacionais.

6.1.4 Segunda PoC: enriquecimento de interesses educacionais extraídos

Considerando que os interesses extraídos foram enriquecidos e novos interesses foram descobertos, foi realizada uma segunda prova de conceito, especificamente para analisar esses interesses descobertos. O objetivo da segunda prova de conceito é definido por:

- ***Analisar o*** *enriquecimento dos interesses educacionais extraídos das redes sociais* ***para o propósito de*** *avaliar os novos interesses identificados a partir do enriquecimento dos interesses extraídos* ***com relação à*** *compatibilidade dos interesses identificados com os reais interesses do usuário* ***do ponto de vista do*** *usuário* ***no contexto*** *de usuários da rede social Facebook, que possuem interesses relacionados à área de tecnologia e expressam tais interesses através das suas interações na rede social.*

Os interesses extraídos na primeira prova de conceito para os usuários Pesquisador A e Pesquisador B foram enriquecidos e alguns interesses descobertos foram apresentados a eles.

O pesquisador A demonstrou na sua rede social possuir interesse, por exemplo, na área de "*Data Mining*". Sabe-se que a área de Data Mining possui diversos outros tópicos relacionados e, conforme observado na primeira PoC, esses tópicos não estão presentes no perfil do usuário. Para realizar o enriquecimento do interesse "Data Mining" e oferecer ao Pesquisador A recursos educacionais relacionados a temas afins, é feito o enriquecimento desse interesse.

Conforme explicado nos capítulos 4 e 5, o enriquecimento é feito de duas formas. Na primeira abordagem, o enriquecimento é resultado do uso de técnicas de Extração de Informação, para reconhecimento de entidades, conceitos e palavras-chave, a partir do texto descritivo da página de interesse no Facebook. Um exemplo é mostrado na figura 38, onde são apresentados termos extraídos a partir das técnicas de Extração de Informação para o texto relacionado ao tema "*Data mining*" no Facebook. Na segunda abordagem, foi usada uma ontologia e através de uma busca pelo termo "*Data Mining*" identifica-se outros temas relacionados apresentados na figura 39. Esse processo de enriquecimento é feito para cada um dos interesses do usuário.

Figura 38 - Termos extraídos do texto descritivo da área de "*Data Mining*" no Facebook.

Figura 39 - Termos encontrados através da consulta à uma ontologia usando o termo "*Data Mining*".

O usuário Pesquisador B que demonstrou, através da sua rede social, ter interesse na área de “Wireless Network”, também teve esse interesse enriquecido, seguindo as duas abordagens explicadas anteriormente. Na figura 40 é mostrado o enriquecimento usando as técnicas de Extração de Informação. No caso do enriquecimento através da ontologia, não foram encontrados novos termos relacionados.

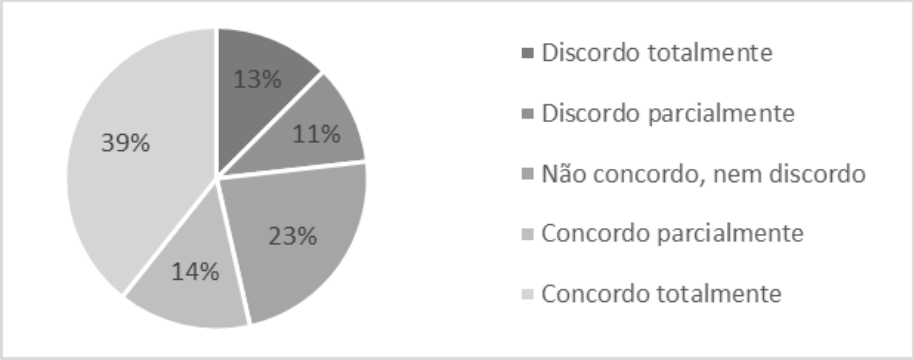
Figura 40 - Termos extraídos do texto da página descritiva da área de interesse "*Wireless Network*" no Facebook.

6.1.4.1 Análise quantitativa do enriquecimento de interesses educacionais extraídos

Após o processo de enriquecimento, os usuários Pesquisador A e Pesquisador B foram convidados a avaliar alguns dos interesses enriquecidos, respondendo à seguinte

pergunta: os interesses descobertos ou associados ao interesse extraído da sua rede social representam de fato seus reais interesses educacionais? Os usuários indicaram concordar totalmente (39%) e parcialmente (14%) com 53% dos interesses sugeridos, se mantiveram neutros em 23% dos casos e discordaram parcialmente (11%) e totalmente (13%) de 24% dos interesses sugeridos, conforme mostrado no gráfico 3.

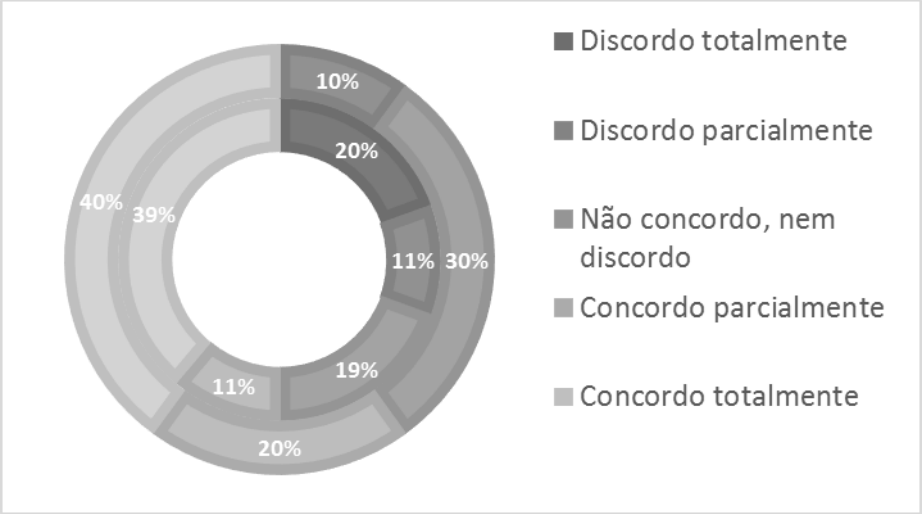
Gráfico 3 - Avaliação dos interesses descobertos dos Pesquisadores A e B.



Fonte: elaborado pelo autor.

Durante a realização desta prova de conceito foi verificada também a aceitação dos usuários em relação às duas abordagens: enriquecimento através de técnicas de Extração de Informação e enriquecimento através de ontologias. O resultado é apresentado no gráfico 4. Através dessa verificação foi possível perceber que a aceitação dos interesses sugeridos através da abordagem de enriquecimento por ontologias foi maior (60%) que a aceitação dos interesses sugeridos através da abordagem de enriquecimento por técnicas de Extração de Informação (50%).

Gráfico 4 - Comparativo entre o enriquecimento por ontologia (círculo externo) e o enriquecimento por técnicas de Extração de Informação (círculo interno).



Fonte: elaborado pelo autor.

No que se refere aos interesses educacionais gerais identificados a partir da categorização das páginas favoritas, o Pesquisador A indicou concordar plenamente com 100% dos interesses, ao passo que o Pesquisador B concordou plenamente com 50% dos interesses e discordou parcialmente de 50% deles.

6.1.4.2 Evidências observadas na segunda PoC

Através da realização da segunda prova de conceito foi possível verificar a viabilidade de extração de interesses que não foram explicitados pelos usuários nas suas redes sociais. Esses interesses podem enriquecer e diversificar as recomendações de recursos educacionais enviadas ao usuário.

Os resultados obtidos nesta prova de conceito reforçaram os indícios obtidos na primeira PoC de que a primeira hipótese apresentada nesta dissertação pode ser confirmada, já que além de ser possível extrair informações das redes sociais, foi possível também enriquecer interesses educacionais extraídos das redes sociais do usuário através da descoberta de interesses implícitos.

Foi possível observar que os usuários desconheciam alguns dos temas sugeridos, mesmo esses estando diretamente associados a um dos seus temas de interesse. Nesses casos, os usuários indicaram não concordar, nem discordar da sugestão do tema, já que o desconheciam. Avalia-se aqui que apesar de, a princípio, esses temas não terem sido aceitos pelos usuários, a recomendação de recursos educacionais que os abordem pode trazer um conhecimento importante para o usuário a respeito de um tema relacionado a um dos seus interesses que até o momento ele desconhece.

Verificou-se uma aceitação um pouco maior dos usuários em relação aos interesses inferidos com o apoio de uma ontologia em relação aos interesses encontrados através de técnicas de Extração de Informação. Acredita-se que essa diferença seja válida para outros contextos e cenários, já que a ontologia geralmente é mais precisa, uma vez que depende do conhecimento de um especialista na área para o seu desenvolvimento, ao passo que a extração de informações é feita de forma automática, a partir da análise do texto, e pode sugerir termos que não expressem de fato uma correlação com o interesse do usuário.

Foi possível observar também através da segunda PoC que o enriquecimento por técnicas de extração de texto foi mais eficiente para interesses descritos na língua inglesa, já que a base de entidade, conceitos e suas relações e palavras-chave é muito mais extensa para termos em inglês. O uso de ontologias pode amenizar essa situação, através da criação de

ontologias de domínio em cada uma das áreas de conhecimento em diferentes idiomas.

A concordância dos usuários com apenas 53% dos interesses descobertos trouxe um indicativo de que a automatização dessa parte do processo de recomendação pode não ser a melhor abordagem, já que a não validação, por parte dos usuários, em relação a interesses descobertos pode gerar recomendações que não sejam condizentes aos reais interesses do usuário, fazendo com que o propósito de gerar recomendações aderentes ao seu perfil e aos seus interesses não seja alcançado com a qualidade desejada.

6.1.5 Terceira PoC: recomendações aderentes ao perfil e ao contexto do usuário

O objetivo da terceira prova de conceito é definido por:

- ***Analisar a*** *viabilidade de gerar recomendações educacionais através do BROAD-RSI* ***para o propósito de*** *constatar a viabilidade de geração das recomendações e avaliar tais recomendações* ***com relação à*** *aderência ao perfil do usuário e à sua relevância para o usuário* ***do ponto de vista do*** *usuário* ***no contexto*** *de recomendações geradas a partir do BROAD-RSI para usuários da rede social Facebook, que possuem interesses relacionados à área de tecnologia e expressam tais interesses através das suas interações na rede social.*

Para o melhor entendimento desta prova de conceito, vale esclarecer que aderência deve ser entendida como a relação entre as características dos recursos educacionais e as preferências do usuário, sendo que as características dos recursos educacionais são recuperadas através de metadados que os descrevem, conforme explicado no capítulo 4.

A relevância da recomendação é entendida como a sua capacidade de atender a uma necessidade do usuário que busca informação sobre um determinado tema, ou seja se a recomendação oferecida a ele terá relevância em seu contexto de alguma forma.

Esta prova de conceito foi realizada com os mesmos usuários descritos na PoC 1 e na PoC 2. Após a identificação de características de perfil, dos interesses educacionais dos usuários e da descoberta de novos interesses, o BROAD-RSI contém algumas informações úteis para gerar recomendações coerentes para cada um dos usuários.

Para a realização desta prova de conceito foi solicitado ao usuário que ele escolhesse alguns temas do seu interesse dentre aqueles extraídos e enriquecidos pelo BROAD-RSI e avaliados por ele como plenamente ou parcialmente condizentes com os seus reais interesses educacionais para receber recomendações. Conforme já explicado nos

capítulos 4 e 5, o usuário pode receber a recomendação de duas formas: na primeira é feito o envio automático da recomendação para a página do usuário na rede social Facebook e, na segunda, o usuário faz a solicitação de uma recomendação dentro do aplicativo BROAD-RSI, selecionando um dos seus interesses. Visando otimizar e facilitar o processo de avaliação, já que grande parte do questionário de avaliação está implementado dentro do aplicativo BROAD-RSI, optou-se pela segunda forma de apresentação da recomendação. No entanto, foi constatada a viabilidade técnica da primeira, buscando avaliar e garantir as duas alternativas.

Após executadas as etapas de extração de características de perfil e de contexto, definição do perfil educacional e enriquecimento dos interesses educacionais, mostradas nas provas de conceito apresentadas anteriormente, a etapa de recomendação usa todas as informações obtidas para gerar recomendações aderentes ao perfil e ao contexto dos usuários.

As recomendações feitas aos usuários estão divididas em três grupos em relação à fonte de pesquisa:

- Grupo 1: recomendações de objetos de aprendizagem armazenados no repositório do BROAD-RSI, catalogados com o conjunto de metadados do LOM (DUVAL; HODGINS, 2003);
- Grupo 2: recomendação de links, extraídos da DBPEDIA (AUER *et al*., 2007);
- Grupo 3: vídeos, extraídos do canal de vídeo Youtube[34].

Conforme mostrado na primeira PoC, o tema "Data Mining" é um dos temas de interesse do Pesquisador A. Este interesse foi identificado através de uma das páginas curtidas pelo usuário na rede social Facebook. O interesse "Data Mining" passou pela etapa de enriquecimento de interesse proposta no BROAD-RSI e gerou novos temas relacionados, conforme mostrado na segunda PoC. O Pesquisador A, seguindo as etapas propostas no modelo de recomendação, finalmente solicita ao BROAD-RSI recomendações sobre o tema "Data Mining". A partir dessa solicitação ele recebe recursos educacionais oriundos dos três grupos de recomendações.

Na figura 41 são mostrados os recursos educacionais recuperados através do repositório do BROAD-RSI recomendados ao Pesquisador A para o tema "*Data Mining*". Vale observar que a recomendação já engloba interesses recuperados pelo enriquecimento através das técnicas de extração de texto e ontologias. A aderência, mostrada na quarta coluna, é calculada conforme explicado no capítulo 5.

Pode-se perceber que alguns recursos educacionais são mostrados ao usuário mais

[34] www.youtube.com

de uma vez, no entanto com relevâncias distintas para temas diferentes, por exemplo o recurso, cujo título é "*Data Warehouse*", é mostrado ao usuário duas vezes, sendo que na primeira ele está associado ao tema "*Data Warehouse*", com relevância aproximada de 1,25, já para o tema "Data" ele tem relevância aproximada de 0,67.

Tema	Título	Relevancia	Aderência	Link
DATA WAREHOUSE	Data Warehouse	1.257743716	100.0	[illegible]
"Data mining"		0.892605483	75.0	[illegible]
DATA MINING	Processo de Web	0.891088008	100.0	[illegible]
DATA MINING		0.744943380	75.0	[illegible]
DATA	Data Warehouse	0.673409461	100.0	[illegible]
MACHINE LEARNING	E-COurse Ontology for Developing E-learning Courses	0.407287359	75.0	[illegible]

Figura 41 - Recomendações extraídas do repositório do BROAD-RSI para o tema "*Data Mining*" e seus temas relacionados.

Além das recomendações de recursos educacionais armazenados no repositório do BROAD-RSI, estão disponíveis para o Pesquisador A algumas recomendações de vídeos, mostradas na figura 42 e algumas recomendações de links, mostrados na figura 43.

Tema		Nome	Facebook
Data_mining		Easy Data Mining Get Started	Recomendar Compartilhar
Data_mining		Statistical Aspects of Data Mining (Stats 202) Day 1	Recomendar Compartilhar
Collaborative_filtering		Web 2.0 Expo NY: Jay Adelson (digg), Organizing Chaos: The Growth of Collaborative Filters	Recomendar Compartilhar
Collaborative_filtering		Collaborative Filtering	Recomendar Compartilhar
Data_cleaning		How to Use SAS - Lesson 5 - Data Reduction and Data Cleaning	Recomendar Compartilhar
		SPSS for newbies: Select	

Figura 42 - Recomendações de vídeos associados ao tema "*Data mining*" e seus temas relacionados.

Tema	Links Externos	Facebook	Avaliação
Data_mining	http://www.intelnics.com	Recomendar Compartilhar	
Data_mining	http://cs.nju.edu.cn/zhouzh/zhouzh.files/course/dm/reading/reading01/chen_tkde96.pdf	Recomendar Compartilhar	
Data_mining	http://sentic.net/api	Recomendar Compartilhar	
Data_mining	http://www.fortewares.com/qiware	Recomendar Compartilhar	
Data_mining	http://www.intelnics.com/neuraldesigner	Recomendar Compartilhar	
Data_mining	http://www.fortewares.com	Recomendar Compartilhar	
Collaborative_filtering	http://www.doi.org/	Recomendar Compartilhar	
Collaborative_filtering	http://agents.media.mit.edu/projects.html	Recomendar Compartilhar	
Collaborative_filtering	http://ectrl.itc.it/home/laboratory/meeting/download/p5-l_herlocker.pdf	Recomendar Compartilhar	
Collaborative_filtering	http://research.yahoo.com/pub/2435	Recomendar Compartilhar	
Collaborative_filtering	http://webpages.uncc.edu/~asaric/ISMIS09.pdf	Recomendar Compartilhar	
Collaborative_filtering	http://www.grouplens.org/papers/pdf/rec-sys-overview.pdf	Recomendar Compartilhar	
Collaborative_filtering	http://www.ieor.berkeley.edu/~goldberg/pubs/eigentaste.pdf	Recomendar Compartilhar	

Figura 43 - Recomendação de links sobre o tema "*Data Mining*" e seus temas relacionados.

Em todas as opções de recomendação, o Pesquisador A tem a opção de Recomendar, Compartilhar ou Curtir a recomendação na sua página na rede social Facebook. Ele poderá também ver quantas pessoas já curtiram a mesma recomendação. Desse modo, ele poderá usar também a informação acerca do número de compartilhamentos para definir a prioridade na sua escolha pela recomendação.

Conforme mostrado na primeira prova de conceito, o usuário Pesquisador B

possui interesse no tema “Wireless Network”. Este usuário deseja encontrar algumas recomendações educacionais sobre esse tema e, para isso, acessa o BROAD-RSI. Esse interesse é enriquecido, gerando outros temas relacionados, e algumas recomendações são geradas para o usuário. Na figura 44 são mostradas as recomendações de recursos educacionais armazenados no BROAD-RSI, relacionadas ao tema “Wireless Network”.

Tema	Título	Relevancia	Aderência	Link
WIRELESS	Wireless sensor networks: a survey	0.44361767	75.0	
TELECOMMUNICATIONS NETWORKS	Wireless sensor networks: a survey	0.614054322	75.0	
WIRELESS NETWORK	Wireless sensor networks: a survey	0.148615747	75.0	

Figura 44 - Recomendações de recursos educacionais armazenados no repositório do BROAD-RSI a respeito do tema "*Wireless Network*" e seus temas relacionados.

Na figura 45 são mostradas as recomendações de alguns links relacionados ao tema e na figura 46 são mostrados alguns vídeos recomendados ao Pesquisador B.

Tema		WikiPe
Wireless_network	International_Journal_of_Wireless_Information_Networks	Re
Wireless_network	Foundations_and_Trends_in_Communications_and_Information_Theory	Re
Wireless_network	A wireless network is any type of computer network that uses wireless data connections for connecting network nodes.Wireless networking is a method by which homes, telecommunications networks and enterprise (business) installations avoid the costly process of introducing cables into a building, or as a connection between various equipment locations. Wireless telecommunications networks are generally implemented and administered using radio communication. This implementation takes place at the physical level (layer) of the OSI model network structure.Examples of wireless networks include cell phone networks, Wi-Fi local networks and terrestrial microwave networks.@en	

Tema	Links Externos	Facebook	Avaliação
Wireless_network	http://wndw.net/pdf/wndw2-en/wndw2-ebook.pdf	Recomendar Compartilhar	
Wireless_network	http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_8-1/wireless_networks.html	Recomendar Compartilhar	

Figura 45 - Recomendação de links sobre o tema "*Wireless Network*" e seus temas relacionados.

Figura 46 - Recomendação de vídeos sobre o tema "*Wireless Network*" e seus temas relacionados.

A viabilidade técnica de envio das recomendações para a página do usuário no Facebook foi constatada através de um envio para a página do Pesquisador A, conforme mostrado na figura 47, na qual foi feita uma recomendação sobre *Data Mining*.

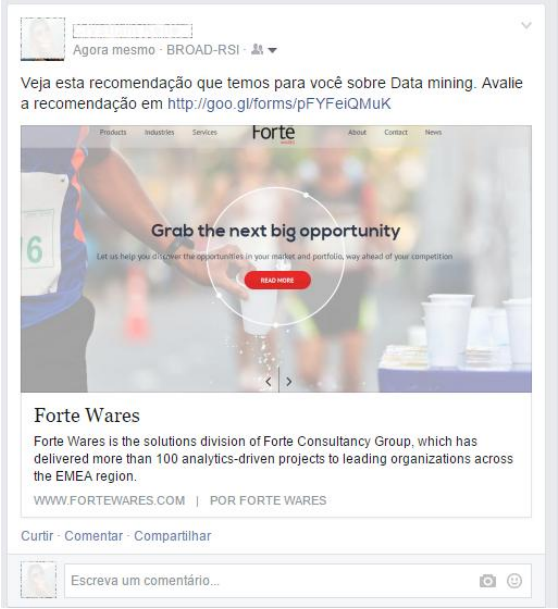


Figura 47 - Envio de recomendação na página do Facebook do Pesquisador A.

6.1.5.1 Análise quantitativa relacionada às recomendações feitas aos usuários

Os usuários Pesquisador A e Pesquisador B foram convidados a avaliar as recomendações geradas para os temas que foram avaliados por eles como plenamente coerentes ou parcialmente coerentes com os seus interesses educacionais. Para realizar a avaliação eles responderam às seguintes perguntas:

- Considerando o seu interesse neste tema, esta recomendação é relevante para você?
- Você compartilharia essa recomendação na sua rede social?

- Essa recomendação está aderente ao seu perfil?

Através da primeira pergunta, deseja-se verificar se as recomendações geradas pelo BROAD-RSI a partir das características de perfil e de contexto extraídas têm relevância para o usuário, ou seja, se elas são capazes de atender à sua necessidade de busca por informação sobre um determinado tema.

O questionamento feito aos usuários na segunda pergunta visa averiguar a capacidade de alcance das recomendações a outros usuários, através das redes sociais, uma vez, que se o usuário indicar interesse e intenção de compartilhar uma recomendação em sua rede social, possivelmente esse recurso educacional atingirá outros usuários, aumentando assim a possibilidade de disseminação e utilização desse recurso.

A terceira pergunta tem o objetivo de investigar se as recomendações propostas pelo BROAD-RSI foram coerentes com as características de perfil e contexto do usuário. Busca-se através desse questionamento encontrar indícios acerca da utilização de todas as informações extraídas e inferidas nas etapas anteriores para a recomendação de recursos educacionais. Para a terceira pergunta vale esclarecer que foi explicado aos usuários que a aderência está relacionada ao tema de interesse, ao idioma do recurso recomendado, ao tipo de recurso (áudio, vídeo, impresso), ao nível de escolaridade (ensino médio, graduação, pós-graduação, etc) e a possibilidade do tipo do material ser executado no seu dispositivo.

Os quantitativos apresentados a seguir foram separados para cada grupo de recomendação de acordo com os três grupos já explicados neste capítulo, a citar: grupo 1, cujos recursos foram extraídos do repositório do BROAD-RSI; grupo 2 cujos recursos são links extraídos da DBPEDIA; grupo 3, cujos recursos são vídeos extraídos do Youtube.

Os dois usuários foram convidados a analisar o conjunto de recomendações. O quantitativo obtido na avaliação do Pesquisador A para a questão "Considerando o seu interesse neste tema, esta recomendação é relevante para você?" é mostrado na tabela 9.

Tabela 9 – Respostas do Pesquisador A para as recomendações feitas (Pergunta 1).

CONSIDERANDO O SEU INTERESSE NESTE TEMA, ESTA RECOMENDAÇÃO É RELEVANTE PARA VOCÊ?				
Escala	Grupo 1	Grupo 2	Grupo 3	Total
Concordo plenamente	15	10	9	34
Concordo parcialmente	2	2	6	10
Não concordo, nem discordo	1	2	3	6
Discordo parcialmente	5	0	0	5
Discordo plenamente	0	0	0	0
Total	23	14	18	55

Fonte: elaborada pelo autor.

Fazendo-se uma análise geral, onde são consideradas as recomendações dos três grupos, pode-se perceber que 80% das recomendações foram avaliadas pelo Pesquisador A como plenamente (62%) ou (18%) parcialmente úteis, se considerado o interesse que o mesmo estava analisando.

O pesquisador B foi convidado a responder a mesma pergunta feita ao Pesquisador A para todos os interesses que ele avaliou serem plenamente ou parcialmente condizentes com os seus interesses educacionais. O quantitativo obtido na avaliação dele é mostrado na tabela 10.

Tabela 10 - Respostas do Pesquisador B para as recomendações feitas (Pergunta 1).

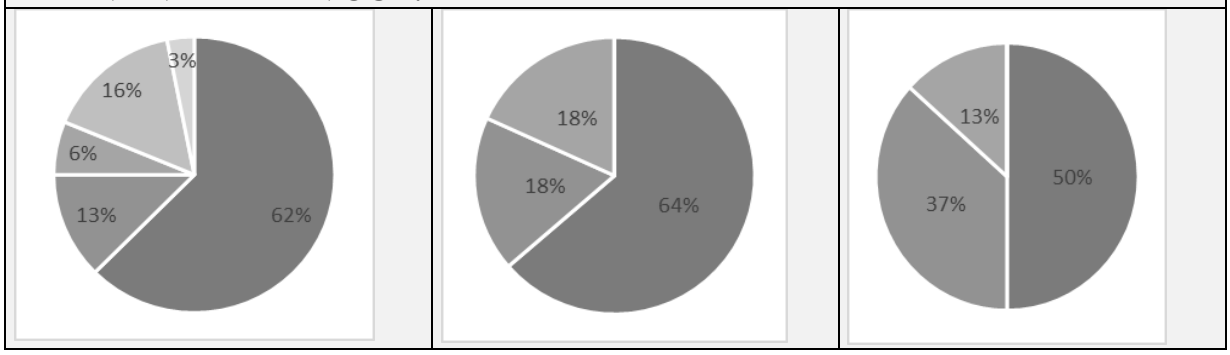
CONSIDERANDO O SEU INTERESSE NESTE TEMA, ESTA RECOMENDAÇÃO É RELEVANTE PARA VOCÊ?				
Escala	Grupo 1	Grupo 2	Grupo 3	Total
Concordo Plenamente	5	4	6	15
Concordo Parcialmente	2	2	5	9
Não concordo, nem discordo	1	2	1	4
Discordo Parcialmente	0	0	0	0
Discordo Plenamente	1	0	0	1
Total	9	8	12	29

Fonte: elaborada pelo autor.

Fazendo uma análise geral, onde são consideradas as recomendações dos três grupos, pode-se perceber que 82,75% das recomendações foram avaliadas pelo Pesquisador B como plenamente (51,72%) ou (31,03%) parcialmente úteis, se considerado o interesse que o mesmo estava analisando.

Na figura 48 são apresentados gráficos nos quais as respostas do Pesquisador A e do Pesquisador B foram separadas para cada um dos grupos de recomendação. Os gráficos são apresentados na seguinte ordem da esquerda para a direita: grupo 1, grupo 2, grupo 3, que são respectivamente recomendações oriundas do repositório do BROAD, da DBPEDIA e do canal de vídeos do Youtube.

CONSIDERANDO O SEU INTERESSE NESTE TEMA, ESTA RECOMENDAÇÃO É RELEVANTE PARA VOCÊ?



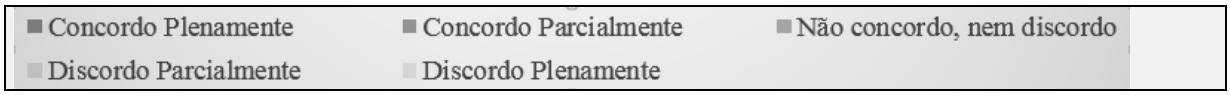


Figura 48 - Avaliações dos usuários para as recomendações feitas a eles (Pergunta 1).

Se comparados os três grupos, pode-se perceber que o grupo 1, cujo os itens recomendados foram extraídos do repositório do BROAD-RSI, foi o que apresentou menor percentual de aceitação (plena + parcial), 75%. Os recursos do grupo 2 e do grupo 3 receberam aceitação (plena + parcial) superiores a 80%.

Os dois usuários foram convidados a responder também à seguinte questão: "Você compartilharia esta recomendação na sua rede social?" As respostas do Pesquisador A podem ser vistas na tabela 11 e as respostas do Pesquisador B estão apresentadas na tabela 12.

Tabela 11 - Respostas do Pesquisador A para as recomendações feitas (Pergunta 2).

VOCÊ COMPARTILHARIA ESSA RECOMENDAÇÃO NA SUA REDE SOCIAL? (PESQUISADOR A)				
Escala	Grupo 1	Grupo 2	Grupo 3	Total
Concordo plenamente	7	6	11	24
Concordo parcialmente	10	6	5	21
Não concordo, nem discordo	0	2	2	4
Discordo parcialmente	6	0	0	6
Discordo plenamente	0	0	0	0
Total	23	14	18	55

Fonte: elaborada pelo autor.

Tabela 12 - Respostas do Pesquisador B para as recomendações feitas (Pergunta 2).

VOCÊ COMPARTILHARIA ESSA RECOMENDAÇÃO NA SUA REDE SOCIAL? (PESQUISADOR B)				
Escala	Grupo 1	Grupo 2	Grupo 3	Total
Concordo Plenamente	0	4	8	12
Concordo Parcialmente	7	2	3	12
Não concordo, nem discordo	2	2	1	5
Discordo Parcialmente	0	0	0	0
Discordo Plenamente	0	0	0	0
Total	9	8	12	29

Fonte: elaborada pelo autor.

Fazendo-se uma análise geral, onde são consideradas as recomendações dos três grupos, pode-se perceber que o Pesquisador A concorda (plena + parcial) em compartilhar 81,81% das recomendações, enquanto o Pesquisador B concorda (plena + parcial) em compartilhar 82,75% das recomendações.

Na figura 49 são apresentados gráficos nos quais a avaliação dos dois usuários foi separada para cada um dos grupos. Os gráficos são apresentados na seguinte ordem da

esquerda para a direita: grupo 1, grupo 2, grupo 3.

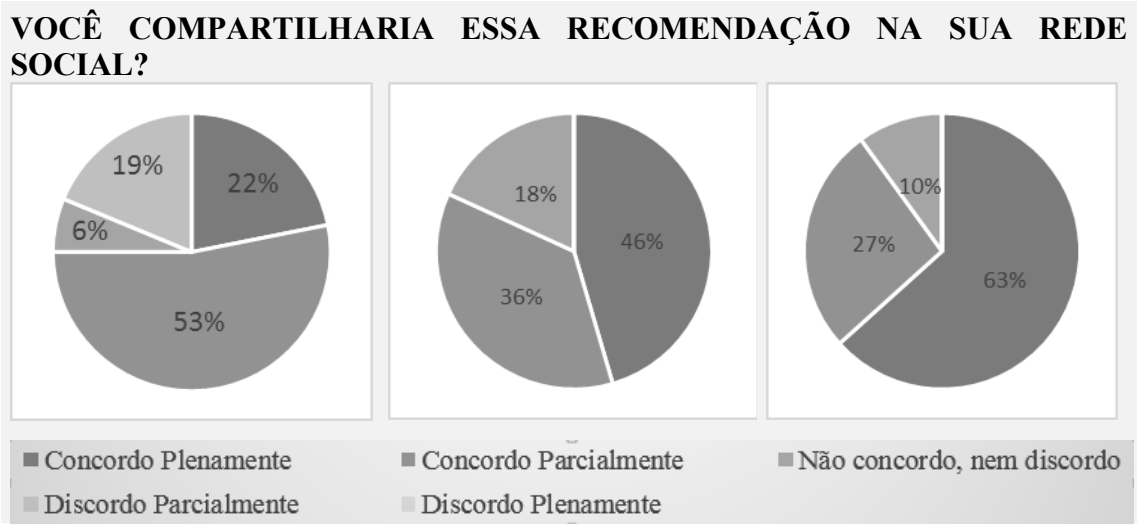


Figura 49 - Avaliação em relação ao compartilhamento das recomendações nas redes sociais.

Se comparados os três grupos, pode-se perceber que o grupo 1, cujos os itens recomendados foram extraídos do repositório do BROAD-RSI, foi o que apresentou menor percentual de aceitação (plena + parcial) dos usuários em relação a possibilidade de compartilhamento nas redes sociais, 75%. Os recursos do grupo 2 e do grupo 3 receberam aceitação (plena + parcial) superiores a 80%.

A resposta positiva dos pesquisadores em relação ao compartilhamento das recomendações mostrou que o envio dos recursos educacionais através de redes sociais tem a capacidade de alcançar pessoas que não são usuários do BROAD-RSI através de usuários que recebem a recomendação de um recurso educacional e a compartilha no seu perfil.

A terceira pergunta feita aos usuários foi a seguinte: Essa recomendação está aderente ao seu perfil? O quantitativo das respostas dadas pelo Pesquisador A pode ser visualizado na tabela 13, enquanto as respostas do Pesquisador B estão apresentadas na tabela 14.

Tabela 13 - Respostas do Pesquisador A para as recomendações feitas (Pergunta 3).

ESTA RECOMENDAÇÃO ESTÁ ADERENTE AO SEU PERFIL? (PESQUISADOR A)				
Escala	Grupo 1	Grupo 2	Grupo 3	Total
Concordo plenamente	14	1	1	16
Concordo parcialmente	3	12	15	30
Não concordo, nem discordo	1	1	1	3
Discordo parcialmente	5	0	1	6
Discordo plenamente	0	0	0	0
Total	23	14	18	55

Fonte: elaborada pelo autor.

Tabela 14 - Respostas do Pesquisador B para as recomendações feitas (Pergunta 3).

ESSA RECOMENDAÇÃO ESTÁ ADERENTE AO SEU PERFIL? (PESQUISADOR B)				
Escala	Grupo 1	Grupo 2	Grupo 3	Total
Concordo Plenamente	5	4	10	19
Concordo Parcialmente	2	2	1	5
Não concordo, nem discordo	1	2	0	3
Discordo Parcialmente	0	0	1	1
Discordo Plenamente	1	0	0	1
Total	9	8	12	29

Fonte: elaborada pelo autor.

Fazendo-se uma análise geral, onde são consideradas as recomendações dos três grupos, pode-se perceber que o Pesquisador A concorda (plena + parcial) que 83,63% das recomendações são aderentes ao seu perfil, Já o Pesquisador B concorda (plena +parcial) que 82,75% das recomendações são aderentes ao seu perfil.

Na figura 50 são apresentados gráficos nos quais a avaliação dos usuários foi separada para cada um dos grupos. Os gráficos são apresentados na seguinte ordem da esquerda para a direita: grupo 1, grupo 2, grupo 3.

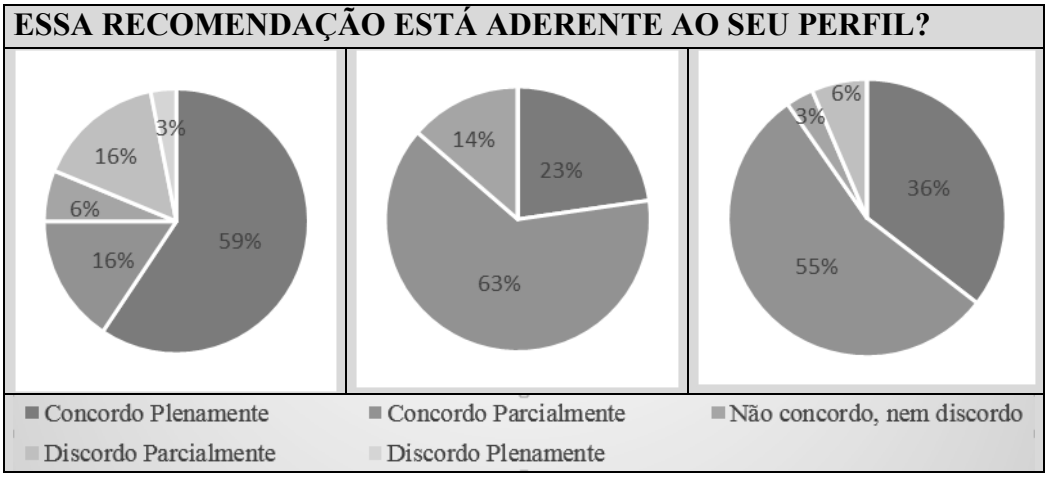


Figura 50 - Avaliação em relação à aderência da recomendação ao perfil.

Em relação à aderência das recomendações ao perfil do usuário, o grupo 1 também foi o que obteve o menor percentual de aceitação, 75%. O Grupo 2 e o Grupo 3 alcançaram 86% e 91% respectivamente.

Em relação a aderência das recomendações ao perfil do usuário, vale observar que, no caso do grupo 1 o Pesquisador A apontou que concorda plenamente com 61% das recomendações, enquanto que para o grupo 2 e grupo 3 o percentual de aceitação plena foi 7% e 5%, respectivamente, sendo que a aceitação parcial atingiu 86% e 83%, respectivamente. Não se pode afirmar com exatidão que o idioma tenha sido o fator determinante para esta

avaliação, no entanto pode-se perceber, no momento da realização da prova de conceito, que o Pesquisador A, que indicou no seu perfil ter conhecimento nos idiomas português e espanhol, avaliou com parcialmente aderente a grande maioria dos links, extraídos da DBPEDIA e os vídeos disponíveis no Youtube, que possuem como idioma principal o inglês. Ao passo que os recursos educacionais extraídos do repositório do BROAD-RSI, que tem como idioma principal o português, foram avaliados como plenamente coerentes.

6.1.5.2 Evidências observadas na terceira PoC

A realização da terceira prova de conceito, cujo foco foram as recomendações feitas aos usuários pelo BROAD-RSI, trouxe indicativos de que a segunda hipótese apresentada nesta dissertação pode ser confirmada. Através dela foi possível gerar recomendações educacionais relevantes aos usuários, usando características de perfil e de contexto extraídas das suas interações em redes sociais, sendo que essas recomendações puderam ser apresentadas aos usuários através de um aplicativo dentro da rede social.

Acredita-se que a decisão de considerar somente os interesses que os usuários concordaram plena e parcialmente serem seus reais interesses educacionais influenciou a avaliação positiva em relação à recomendação. Isso fortalece o indício de que é importante a validação dos usuários na etapa de enriquecimento.

Na realização desta prova de conceito foi possível observar que as recomendações de recursos educacionais oriundos da Web, seja de links disponíveis no *dataset* da DBPEDIA, seja de vídeos disponíveis em um canal de vídeo, foram bem avaliadas pelos usuários. Essa avaliação pode indicar ser interessante diversificar as fontes de busca de recursos educacionais, não se restringindo a repositórios de objetos de aprendizagem que, conforme explicados no capítulo 4, apresentam dificuldades para serem mantidos e atualizados com frequência, além de muitas vezes serem restritos a algumas instituições ou a uma área de conhecimento.

Um exemplo disso foi percebido durante a realização desta terceira prova de conceito, quando foi possível averiguar a possibilidade de realizar recomendações do Grupo 2 e do Grupo 3 para os interesses do Pesquisador A que não estavam relacionados à área de Tecnologia, como, por exemplo, História do Mundo e História do Brasil, ao passo que não foi possível encontrar recomendações do Grupo 1 para esses interesses, já que o repositório usado para o Grupo 1 possui apenas recursos educacionais específicos para a área de Computação.

A ampla avaliação da influência de cada uma das características de perfil na

avaliação das recomendações exigiria uma série de análises mais minuciosas. Em virtude da complexidade e do tempo necessário para a realização deste trabalho, não foi possível realizá-las no escopo desta dissertação.

Através da análise feita para a terceira questão desta prova de conceito pode-se verificar a viabilidade técnica de oferecer recomendações educacionais aderentes ao perfil dos usuários, considerando-se informações de perfil e de contexto extraídas das interações em redes sociais.

Além de possibilitar a verificação da viabilidade técnica da proposta, e dar os primeiros indícios acerca das hipóteses que estão sendo avaliadas nesta dissertação, a realização das três provas de conceito auxiliaram a formulação do questionário usado como instrumento de avaliação no estudo de caso. A partir do uso do protótipo, durante a realização das provas de conceito, foi possível ajustar alguns parâmetros para a avaliação, solucionar problemas técnicos e perceber algumas questões que poderiam ser melhor avaliadas em um estudo de caso.

Assim sendo, os resultados observados nas três provas de conceito apresentadas motivaram o planejamento e execução de um estudo de caso visando aumentar a capacidade de observação e a confiança na viabilidade da proposta do BROAD-RSI.

6.2 ESTUDO DE CASO

Segundo WOHLIN *et al.* (2012), dependendo do propósito da avaliação, se está se avaliando uma técnica, um método ou uma ferramenta, e dependendo das condições para uma investigação empírica, existem três principais tipos de estratégias de investigação: *surveys*, estudo de caso e experimento. Um *survey* é um sistema para coletar informações de ou sobre pessoas para descrever, comparar ou explicar seus conhecimentos, atitudes e comportamento. Já um estudo de caso no contexto de engenharia de software é uma investigação empírica, que se baseia em diferentes fontes de evidências, usada quando o objeto de estudo é um fenômeno contemporâneo difícil de ser estudado de forma isolada. O experimento geralmente é realizado em laboratório e oferece maior nível de controle das variáveis envolvidas, manipulando uma ou algumas variáveis e mantendo outras fixas medindo o efeito do resultado.

Segundo YIN (2001) o estudo de caso deve ser usado quando o pesquisador tem pouco controle sobre os acontecimentos, em contextos reais e sem controle total sobre as variáveis.

As principais características do estudo de caso são (RUNESON *et al.*, 2012), (apud WOHLIN *et al*., 2012):

- é um tipo de estudo flexível, que lida com as características complexas e dinâmicas de fenômenos do mundo real;
- suas conclusões são baseadas em uma clara cadeia de provas, seja através de uma análise qualitativa ou quantitativa , coletadas de várias fontes diferentes, de forma planejada e consistente;
- o conhecimento existente pode ser baseado em uma teoria previamente estabelecida, ou através da construção de uma teoria.

De acordo com (WOHLIN *et al.*, 2012), a vantagem de um estudo de caso é que eles são mais fáceis de planejar e mais realistas, mas a desvantagem é que os resultados dificilmente podem ser generalizados e são mais difíceis de interpretar.

No caso desta pesquisa a melhor forma de avaliação é um estudo de caso por se tratar de uma pesquisa empírica realizada em um contexto real, que deseja investigar a aceitação por parte dos usuários do processo de recomendação proposto através do BROAD-RSI e obter retorno dos usuários sobre críticas e sugestões.

6.2.1 **Planejamento do estudo de caso**

O estudo de caso conduzido nesta pesquisa tem caráter exploratório-descritivo. As atividades exploratórias foram usadas para observar alguns aspectos que ainda não estão claramente definidos, ao passo que as atividades descritivas foram aplicadas nos fenômenos para os quais já havia uma hipótese definida, buscando através de tais atividades confirmar ou negar as hipóteses apresentadas neste trabalho.

A pesquisa descritiva foi empregada na observação dos seguintes fenômenos:

- Aceitação do usuário em relação às características de perfil e de contexto identificadas através de suas redes sociais;
- Aceitação dos usuários em relação às recomendações educacionais feitas de acordo com as características de perfil e de contexto extraídas;
- Qual fonte de recursos educacionais (repositório do BROAD-RSI, recursos extraídos de conjuntos de dados ligados, recomendação de vídeos do Youtube) usada gerou recomendações mais aceitas pelos usuários.

Já a pesquisa exploratória foi aplicada na observação das seguintes questões:

- Como os usuários analisam o uso de redes sociais para envio de recomendações educacionais;
- Quais as redes sociais mais utilizadas pelos usuários;
- Quanto tempo os usuários interagem nas redes sociais;
- Quais redes sociais os usuários preferem para receber recomendações educacionais.

6.2.1.1 Objetivos do estudo de caso

Os objetivos deste estudo de caso também foram definidos de acordo com a abordagem *Goal/Question/Metric* (GQM) (BASILI, 1992). Este estudo de caso possui dois objetivos principais relacionados à extração das informações a partir da interação dos usuários em suas redes sociais e às recomendações oferecidas ao usuário a partir das suas características de perfil e de contexto, a citar:

i. ***Analisar a*** *extração de características de perfil e de contexto do usuário* ***para o propósito de*** *avaliar a extração de informações sobre o perfil e os interesses educacionais do usuário* ***com relação à*** *precisão da extração* ***do ponto de vista do*** *usuário no contexto de usuários da rede social Facebook, que possuem interesses relacionados à área de Computação e expressam tais interesses através das suas interações nesta rede social.*

ii. ***Analisar as*** *recomendações educacionais através do BROAD-RSI* ***para o propósito de*** *avaliar as recomendações* ***com relação à*** *relevância das recomendações* ***do ponto de vista do*** *usuário no contexto de recomendações geradas a partir do BROAD-RSI para usuários da rede social Facebook, que possuem interesses relacionados à área de tecnologia e expressam tais interesses através das suas interações na rede social.*

6.2.2 **Seleção dos indivíduos**

O estudo de caso foi realizado com 14 participantes voluntários, estando eles distribuídos entre estudantes do mestrado em Ciência da Computação da Universidade Federal de Juiz de Fora e profissionais da área de Computação (analistas de sistema e professores), conforme mostrado na figura 51. Um requisito para a seleção dos indivíduos foi possuir uma conta na rede social Facebook e expressar interesses através das suas interações

nessa rede social.

Figura 51 - Distribuição dos participantes do estudo de caso por atividade profissional.

Vale ressaltar que a escolha dos indivíduos levou em consideração a área da Computação, por que o repositório de recursos educacionais usado como fonte para as recomendações contém catalogados recursos específicos para essa área.

6.2.3 **Preparação e coleta dos dados**

Existem diferentes fontes de informação que podem ser usadas em um estudo de caso. De acordo com LETHBRIDGE *et al*. (2005) existem três principais técnicas de coleta de dados em um estudo de caso, a citar: métodos diretos, onde o pesquisador tem um contato direto com os indivíduos participantes e a coleta é feita em tempo real (entrevistas, questionários, etc); métodos indiretos, onde os pesquisadores coletam indiretamente os dados a partir das interações dos indivíduos durante a coleta de dados (diários de trabalho, observação através de vídeos e áudio, logs do sistema); e independente, onde o pesquisador realiza a análise de artefatos de trabalho (análise de documentos).

De acordo com WOHLIN *et al.* (2012), as entrevistas e questionários podem ser divididas em não-estruturadas, semiestruturadas e totalmente estruturadas. As principais características de cada tipo de entrevista podem ser vistas na tabela 15.

Tabela 15 - Tipos de entrevistas.

	Não-estruturadas	Semiestruturadas	Totalmente estruturadas
Foco	Análise qualitativa a respeito da experiência dos indivíduos em relação a um fenômeno.	Análise qualitativa e quantitativa a respeito da experiência dos indivíduos em relação a um fenômeno.	Pesquisadores buscam encontrar relações entre dois fenômenos.

Questões	Um roteiro de entrevista englobando as áreas especificadas no foco da pesquisa.	Mistura entre questões abertas e fechadas.	Questões fechadas
Objetivo	Exploratória	Descritiva e exploratória	Descritiva e exploratória

Fonte: WOHLIN *et al*., 2012.

Os dados usados para a realização desta pesquisa foram coletados usando método direto, através de entrevistas semiestruturadas, já que o foco é realizar análises quantitativas e qualitativas. Perguntas fechadas foram usadas na avaliação feita pelos usuários para cada item de perfil, cada interesse extraído e inferido, para cada recomendação gerada pelo BROAD-RSI e, ainda, após o uso do BROAD-RSI para verificar alguns pontos da sua experiência. As perguntas discursivas tiveram o objetivo de capturar possíveis informações sobre o estudo de caso não previstas nas perguntas de múltipla escolha. Além de buscar a opinião dos usuários sobre a ferramenta e as suas funcionalidades, também foi esperado obter sugestões de melhoria para o processo de recomendação.

Para realizar a coleta dos dados foi implementado no protótipo do BROAD-RSI, uma funcionalidade onde os usuários podem avaliar as características de perfil, os interesses e as recomendações, conforme mostrado na figura 52. As opções de resposta seguiram a escala de Likert, conforme detalhado anteriormente.

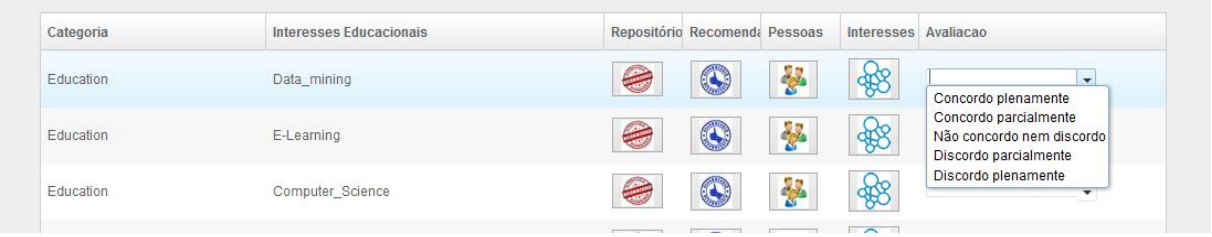


Figura 52 – Tela para o usuário avaliar as recomendações dentro do BROAD-RSI.

A entrevista aplicada após o uso do BROAD-RSI foi feita através do Google Docs[35], usando a ferramenta de criação de formulários[36].

Os dados capturados através do BROAD-RSI foram armazenados no próprio banco de dados do BROAD-RSI, desenvolvido usando MYSQL[37], já os dados capturados através de formulário foram armazenados em uma tabela, também no Google Docs. Essa tabela está associada ao formulário de avaliação e recebe automaticamente os dados a partir

[35] http://www.google.com/docs/about/

[36] https://support.google.com/docs/answer/87809?hl=pt-BR

[37] http://www.mysql.com/

das respostas dos usuários.

6.2.4 **Métricas**

Em SHANI & GUNAWARDANA (2011) é apresentada uma discussão sobre os métodos de avaliação de recomendações, sendo eles resumidos em três principais métricas mais comumente utilizadas em Sistemas de Recomendação: *Precision, Recall, False Positive Rate*. Tais métricas são calculadas a partir dos possíveis resultados das recomendações apresentados na tabela 16:

Tabela 16 –Possibilidades de resultados de recomendações.

	Itens recomendados	Itens não-recomendados
Preferidos	Verdadeiro-Positivo (VP)	Falso-Negativo (FN)
Não Preferidos	Falso-Positivo (FP)	Verdadeiro-Negativo (VN)

Fonte: (SHANI; GUNAWARDANA, 2011).

A partir do número de ocorrências de cada item apresentado, pode-se calcular as seguintes métricas.

Tabela 17 - Métricas para avaliar Sistemas de Recomendação.

Métrica	Equação	Interpretação
Precision	$\frac{\text{Nº de VP}}{\text{Nº de VP} + \text{Nº de FP}}$	Mede a probabilidade de um item recomendado ser relevante.
Recall	$\frac{\text{Nº de VP}}{\text{Nº de VP} + \text{Nº de FN}}$	Mede a probabilidade de um item relevante ser recomendado.
False Positive Rate	$\frac{\text{Nº de FP}}{\text{Nº de FP} + \text{Nº de VN}}$	Mede a probabilidade de um item ser mal recomendado.
F-measure (F1)	$2x\frac{\text{Precision} + \text{Recall}}{\text{Precision x Recall}}$	Combina as características do *Precision* e do *Recall* numa métrica conjunta

Fonte: (SHANI; GUNAWARDANA, 2011).

No caso da avaliação feita neste trabalho foi usada somente a métrica *Precision*, porque não é possível calcular o número de itens preferidos não recomendados (FN), tampouco os itens não preferidos não recomendados (VN). Isso se deve ao fato de usarmos recursos extraídos de um conjunto de dados ligados e de um site de catalogação de vídeos, onde não é possível calcular a totalidade de itens possíveis para um usuário.

Além disso, deseja-se neste trabalho avaliar a aceitação dos itens que foram recomendados, buscando evidências se esses itens são relevantes aos usuários. Sendo assim, a métrica *Precision* atende ao que se deseja avaliar neste caso. Essa métrica não representa um valor percentual, segundo SHANI & GUNAWARDANA (2011), quanto mais próximo de 1 o

resultado da métrica *Precision*, melhor o sistema de recomendação.

Segundo ÁLVAREZ (2007), a métrica *Precision* também é usada com frequência para avaliar Sistemas de Extração de Informação, sendo definida como a quantidade de informações corretamente extraídas sobre todas as informações extraídas. Sendo assim, a mesma métrica será usada para avaliar a extração de características de perfil, os interesses extraídos e os interesses descobertos pela etapa de enriquecimento.

A métrica *Recall* também poderia ser usada para Sistemas de Extração de Informação, no entanto, no escopo deste estudo torna-se inviável por se tratar de informações extraídas de redes sociais, cujo conjunto total de informações não é bem definido, tampouco bem quantificado.

6.2.5 **Análise quantitativa dos dados**

Após coletados os dados, através das entrevistas e do questionário de avaliação dos itens, foi feito um levantamento dos dados capturados para os seguintes itens: (1) avaliação de características de perfil e de interesses extraídos ou descobertos pela etapa de enriquecimento e (2) avaliação das recomendações.

6.2.5.1 Características de perfil, interesses extraídos e interesses descobertos

Os participantes do estudo de caso foram convidados a avaliar os itens de perfil extraídos através das suas interações nas redes sociais. Os itens avaliados foram idiomas falados, preferências por mídia, grau de formação, dia e horário de maior acesso. Busca-se através das análises feitas a seguir obter indícios a respeito da seguinte hipótese apresentada neste trabalho: se extrairmos dados, gerados espontaneamente, a partir da interação dos usuários em redes sociais, então será possível identificar características de perfil e de contexto do usuário (H1). Aplicando a métrica *Precision* para cada um dos elementos tem-se os valores mostrados na tabela 18.

Tabela 18 - Avaliação dos usuários sobre as características de perfil e de contexto extraídos.

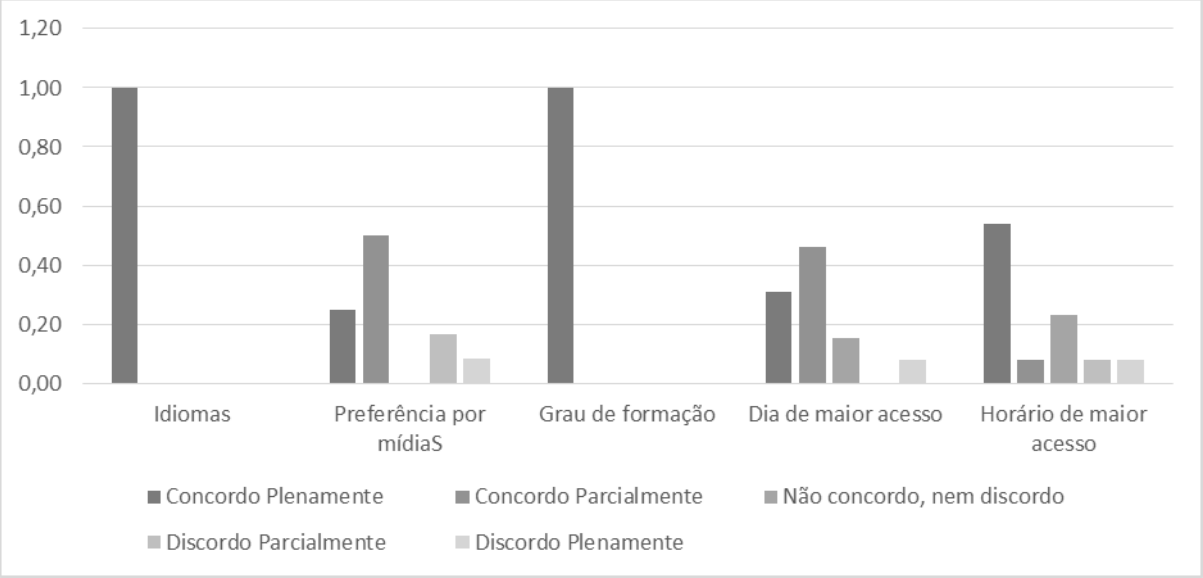
	Idiomas	Preferência por mídias	Grau de formação	Dia de maior acesso	Horário de maior acesso
Concordo Plenamente	1,00	0,25	1,00	0,31	0,54
Concordo Parcialmente	0,00	0,50	0,00	0,46	0,08
Não concordo, nem discordo	0,00	0,00	0,00	0,15	0,23

Discordo Parcialmente	0,00	0,17	0,00	0,00	0,08
Discordo Plenamente	0,00	0,08	0,00	0,08	0,08

Fonte: elaborada pelo autor.

O quantitativo com a respostas dos usuários é apresentado nos gráficos 5.

Gráfico 5 - Avaliação dos usuários sobre as características de perfil extraídas.



Fonte: elaborada pelo autor.

Os participantes do estudo de caso avaliaram também cada interesse educacional extraído, respondendo à seguinte questão: este tema de interesse identificado através das suas interações na rede social Facebook representa de fato um interesse educacional para você? Foi explicado aos participantes que interesse educacional deve ser interpretado como qualquer assunto para o qual ele tem interesse em obter conhecimento, em aprender algo.

Os 14 participantes avaliaram 255 interesses extraídos, ressaltando que há repetição de interesses somente para usuários distintos. Na tabela 19 são mostrados os resultados das avaliações feitas por eles, já com aplicação da métrica *Precision.*

Tabela 19 - Avaliação dos participantes sobre interesses extraídos.

ESCALA	**INTERESSES EXTRAÍDOS**
Concordo Plenamente	0,47
Concordo Parcialmente	0,22
Não concordo, nem discordo	0,10
Discordo Parcialmente	0,04
Discordo Plenamente	0,18

Fonte: elaborada pelo autor.

Os participantes do estudo de caso também foram convidados a avaliar cada um dos interesses enriquecidos através de técnicas de Extração de Informação e ontologias, respondendo à seguinte questão: os interesses descobertos representam um dos seus reais

interesses educacionais?

Na tabela 20 são apresentados os resultados da avaliação dos usuários, seguindo a métrica *Precision*. Na segunda coluna são mostrados os resultados para os interesses gerais, ou seja, as grandes áreas de interesse, extraídas a partir da análise de todas as páginas. Já na terceira coluna são mostrados os resultados para os interesses especificamente educacionais relacionados aos interesses extraídos. Foram analisados 33 interesses gerais e 519 interesses educacionais, sendo 224 enriquecidos a partir de ontologias e 295 através de técnicas de Extração de Informação.

Tabela 20 - Avaliação dos participantes sobre interesses descobertos.

	INTERESSES GERAIS DESCOBERTOS	**INTERESSES EDUCACIONAIS DESCOBERTOS**
Concordo Plenamente	0,61	0,46
Concordo Parcialmente	0,15	0,17
Não concordo, nem discordo	0,03	0,22
Discordo Parcialmente	0,09	0,04
Discordo Plenamente	0,12	0,11

Fonte: elaborada pelo autor.

Foi feita também uma análise a respeito das duas abordagens usadas no enriquecimento: técnicas de Extração de Informação e usando ontologias. Os resultados são mostrados na tabela 21.

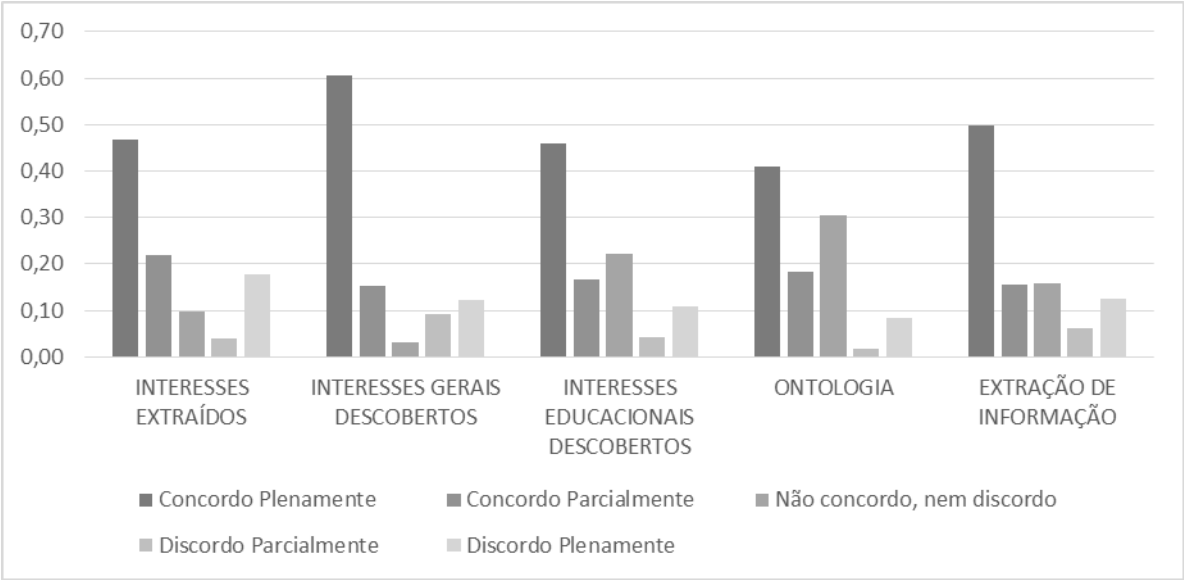
Tabela 21 - Avaliação dos participantes sobre interesses descobertos através das duas abordagens usadas.

	ENRIQUECIMENTO ATRAVÉS DE ONTOLOGIAS	**ENRIQUECIMENTO ATRAVÉS DE TÉCNICAS DE EXTRAÇÃO DE INFORMAÇÃO**
Concordo Plenamente	0,41	0,50
Concordo Parcialmente	0,18	0,16
Não concordo, nem discordo	0,30	0,16
Discordo Parcialmente	0,02	0,06
Discordo Plenamente	0,08	0,13

Fonte: elaborada pelo autor.

No gráfico 6 são mostrados todos os resultados das avaliações a respeito dos interesses.

Gráfico 6 – Avaliação dos interesses.



Fonte: elaborado pelo autor.

6.2.5.2 RECOMENDAÇÕES

A avaliação das recomendações foi feita solicitando aos participantes do estudo de caso que avaliassem cada uma das recomendações em relação à sua relevância, respondendo à seguinte pergunta: considerando que você está buscando recursos educacionais a respeito deste tema, esta recomendação é relevante para você? Busca-se através das análises feitas a seguir obter indícios a respeito da seguinte hipótese apresentada neste trabalho: se usarmos as características de perfil e de contexto extraídas da interação do usuário em redes sociais, então será possível realizar recomendações educacionais relevantes a esses usuários (H2).

O repositório local de recursos educacionais usado pelo BROAD-RSI tinha, no momento da avaliação, 60 recursos educacionais relacionados às áreas de engenharia de software, redes de computadores, computação gráfica e banco de dados.

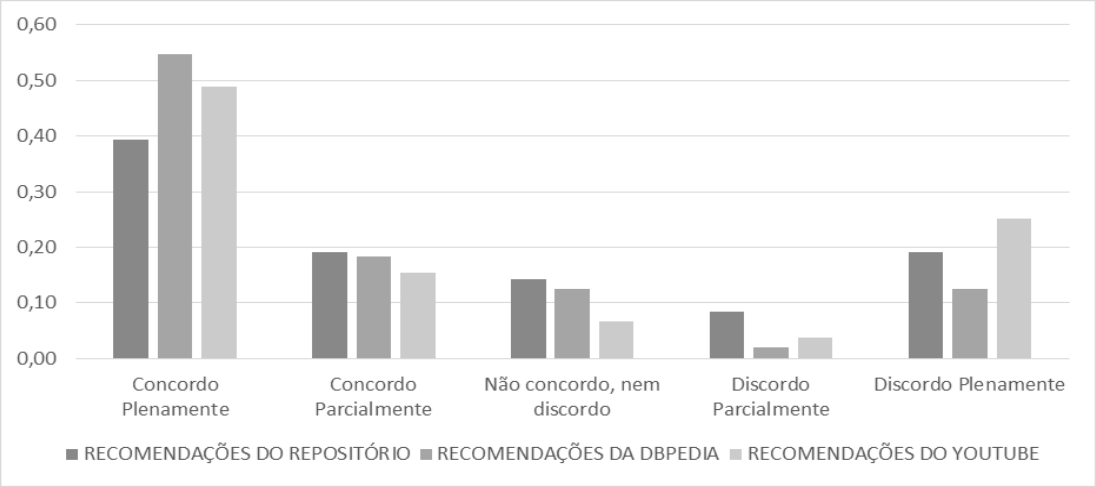
O resultado das avaliações das recomendações, aplicando a métrica *Precision*, é apresentada na tabela 22 e no gráfico 7, separado por cada um dos grupos de recomendação, conforme explicado anteriormente na PoC 3. Na última coluna da tabela é apresentado o resultado da avaliação sem diferenciar os grupos. Foram analisadas 427 recomendações, sendo 84 do Grupo 1, 104 do Grupo 2, 239 do Grupo 3.

Tabela 22 - Resultados da avaliação das recomendações.

	GRUPO 1	GRUPO 2	GRUPO 3	AVALIAÇÃO GLOBAL
Concordo Plenamente	0,39	0,55	0,49	0,48
Concordo Parcialmente	0,19	0,18	0,15	0,17
Não concordo, nem discordo	0,14	0,13	0,07	0,10
Discordo Parcialmente	0,08	0,02	0,04	0,04
Discordo Plenamente	0,19	0,13	0,25	0,21

Fonte: elaborado pelo autor.

Gráfico 7 - Avaliação das recomendações.



Fonte: elaborado pelo autor.

6.2.5.3 DEMAIS ANÁLISES QUANTITATIVAS

Outras duas perguntas foram feitas para os participantes com o propósito de averiguar a capacidade de alcance das recomendações a outros usuários e se as recomendações propostas pelo BROAD-RSI foram coerentes com as características de perfil e contexto do usuário. São elas:

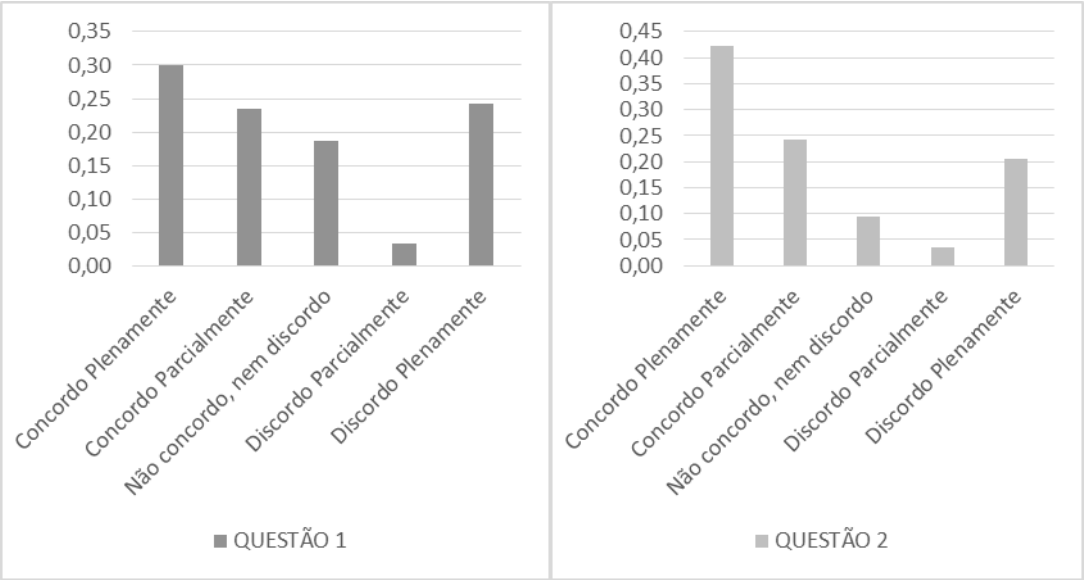
- Questão 1: Você compartilharia essa recomendação na sua rede social?
- Questão 2: Essa recomendação está aderente ao seu perfil?

Tabela 23 - Resultado das avaliações para a questão 1 e questão 2.

	QUESTÃO 1	QUESTÃO 2
Concordo Plenamente	0,30	0,42
Concordo Parcialmente	0,24	0,24
Não concordo, nem discordo	0,19	0,09
Discordo Parcialmente	0,03	0,04
Discordo Plenamente	0,24	0,21

Fonte: elaborado pelo autor.

Gráfico 8 – Resultados das avaliações para a questão 1 e questão 2.



Fonte: elaborado pelo autor.

6.2.5.4 ANÁLISE DOS RESULTADOS

6.2.5.4.1 Características de perfil, interesses extraídos e interesses descobertos

Analisando os resultados obtidos com a realização do estudo de caso, foi possível perceber que a extração das características de perfil dos usuários foi bem avaliada pelos participantes, sendo que o horário de maior acesso foi a característica que obteve menor aceitação, atingindo 62% de aceitação (parcial + total). Essa característica apresentou 23% de neutralidade por parte dos usuários, sendo percebido durante o estudo de caso, através de um questionamento feito ao usuário sobre a baixa avaliação do item, que boa parte dos usuários não conseguiu analisar esse item por não terem clara percepção do horário que interagem no Facebook.

Em relação aos interesses extraídos também houve uma boa aceitação chegando a atingir o índice de 0,69 contra 0,22 de rejeição, sendo que quanto mais próximo de 1 o resultado da métrica *Precision*, melhor o sistema de recomendação. Esse dado traz indícios de que a seleção das categorias das páginas foi satisfatória para filtrar os interesses educacionais dos usuários.

Os interesses gerais descobertos obtiveram índice de 0,75 de aprovação contra 0,21 de rejeição, enquanto os interesses educacionais descobertos ficaram com um índice de aprovação de 0,63 contra 0,15 de rejeição, sendo que nesse caso o índice de neutralidade ficou

em 0,22. As aceitações observadas no estudo de caso foram superiores àqueles observadas na PoC2, no entanto, vale considerar a possibilidade de solicitar confirmação dos usuários em relação aos interesses, buscando diminuir o número de recomendações sobre assuntos que não interessam ao usuário, fazendo com eles percam o interesse no uso do BROAD-RSI. Ainda sobre esse item, foi possível observar um comportamento semelhante ao observado na PoC2, em relação às avaliações neutras, ou seja, quando o usuário desconhecia um determinado tema sugerido, ele manteve-se neutro, mesmo estando esse tema relacionados diretamente a um assunto do seu interesse.

Se comparadas as duas abordagens de enriquecimento usadas nesta proposta, houve uma mudança de tendência dos resultados do estudo de caso em relação à PoC2. Verificou-se no estudo de caso uma aceitação de 0,66 nos itens extraídos através das técnicas de extração da informação, enquanto o enriquecimento por ontologias atingiu o índice de aceitação de 0,59. Foi possível perceber durante a aplicação do estudo de caso que alguns usuários curtiram páginas que abordam assuntos amplos, por exemplo Sistemas de Informação, no entanto possuíam interesses específicos, como por exemplo Desenvolvimento para Web. O enriquecimento por ontologia trouxe ao usuário todos os temas relacionados a Sistemas de Informação e a maior parte deles foi rejeitado pelo usuário, enquanto que as técnicas de Extração de Informação trouxeram apenas interesses possíveis de serem extraídos do texto da página de Sistemas de Informação, geralmente termos também bem gerais, onde o usuário avaliava como aceito. Pelos resultados obtidos não é possível afirmar qual abordagem é mais eficiente, reforçando a estratégia de uso complementar de ambas.

De forma geral, foi possível perceber que existem informações relevantes nas redes que podem auxiliar na tarefa de identificar características de perfil e de contexto dos usuários, reforçando os indícios de que a primeira hipótese apresentada nesta dissertação pode ser confirmada.

6.2.5.4.2 Recomendações

As recomendações oferecidas aos usuários obtiveram, no geral, 0,65 de aprovação e 0,25 de rejeição em relação à relevância que as mesmas teriam caso o usuário tivesse buscado recursos educacionais dentro de um determinado tema. Se separadas por grupos, os índices de aprovação para o Grupo 1, Grupo 2 e Grupo 3 foram 0,58, 0,73, 0,65, respectivamente. Acredita-se que o baixo índice de aceitação do repositório deve-se à

quantidade de recursos educacionais disponíveis para recomendação, pois existiam poucos recursos para atender os diferentes interesses extraídos e enriquecidos.

Em relação à aderência das recomendações ao perfil do usuário, foi obtido um índice de 0,66 de aceitação e 0,25 de rejeição.

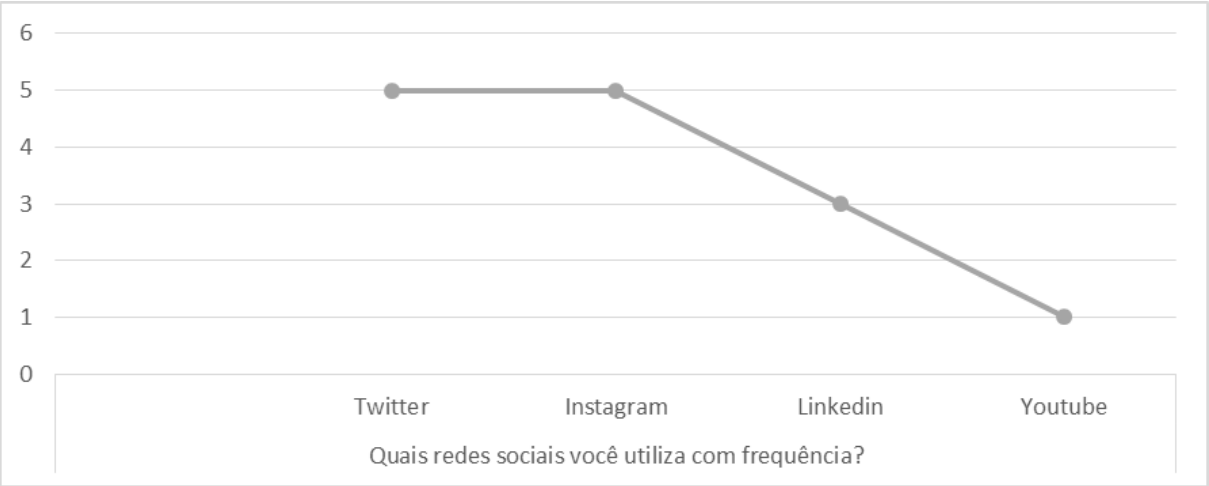
Considerando os índices de aceitação tanto da relevância, quanto da aderência da recomendação há indícios que a segunda hipótese poderá ser aceita, já que a partir das características de perfil e de contexto extraídas da interação do usuário em redes sociais foram geradas recomendações educacionais a esses usuários e eles avaliaram a maior parte delas como relevantes e aderentes.

6.2.6 **Análise qualitativa dos dados**

A partir das sugestões e comentários feitos no estudo de caso realizado com 14 participantes foi possível ter a percepção de alguns pontos importantes para a evolução desta pesquisa.

Os usuários foram questionados sobre o hábito de uso das redes sociais e cerca de 55% indicou passar de 4 a 6 horas navegando e/ou interagindo nas redes sociais, enquanto que o restante indicou passar de 2 a 4 horas. Esses números reforçam a importância de usar o tempo que os usuários passam interagindo nas suas redes sociais para fornecer a eles recomendações educacionais que possam contribuir com a sua formação intelectual. Além do Facebook, alguns usuários apontaram usar as redes sociais Twitter, Instagram, Linkedin e Youtube, conforme mostrado no gráfico 9.

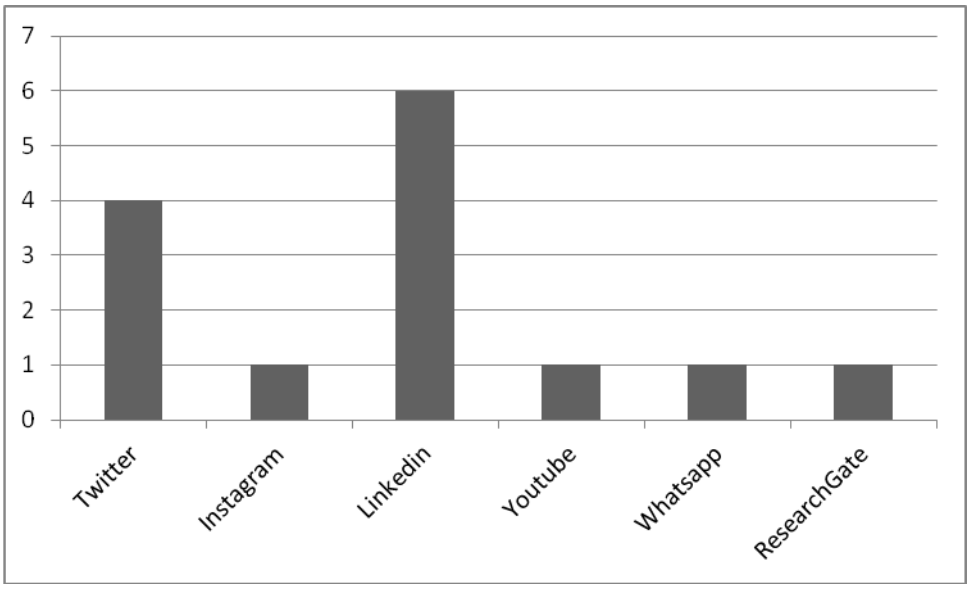
Gráfico 9 - Uso de redes sociais.



Fonte: elaborado pelo autor.

A sugestão mais comum apresentada pelos usuários é de expansão da proposta para outras redes sociais. No gráfico 10, elaborado a partir das sugestões dos usuários, é possível perceber que as duas redes sociais que se destacaram foram o Twitter e o Linkedin.

Gráfico 10 – Redes sociais nas quais os usuários gostariam de receber recomendações.



Fonte: elaborado pelo autor.

Alguns usuários apontaram que seria interessante receber recomendações na rede social Twitter pela frequência que eles a utilizam, já o Linkedin por ser uma rede social mais voltada para o campo profissional, onde teriam diversos interesses educacionais e profissionais disponíveis nos seus perfis, conforme mostra um dos depoimentos coletados:

> *"Acredito que a utilização de outras redes sociais como Linkedin ou ResearchGate, em conjunto com Facebook, podem aprimorar a ferramenta, uma vez que são redes voltadas para trabalho, pesquisa e o meio acadêmico.*
>
> *Caso o Lattes permita explorar suas bases de dados, por intermédio de uma API, seria interessante também utiliza-la para obter informações relevantes para recomendações, uma vez que a plataforma armazena informações precisas das áreas de interesse dos pesquisadores."*

Conforme explicado anteriormente, há possibilidade de expandir a proposta aqui apresentada para outras redes sociais, sendo necessárias algumas adaptações no uso das APIs de cada rede social e ainda na forma de apresentar a informação nas diferentes redes sociais.

Os participantes mostraram aprovar o envio de recomendações para as suas redes sociais, sendo que cerca de 90% apontou concordar plenamente quando questionado se acha

interessante receber recomendações no Facebook e também em, aproximadamente, 90% dos casos, os participantes disseram desejar continuar utilizando o BROAD-RSI, conforme pode ser percebido em alguns depoimentos:

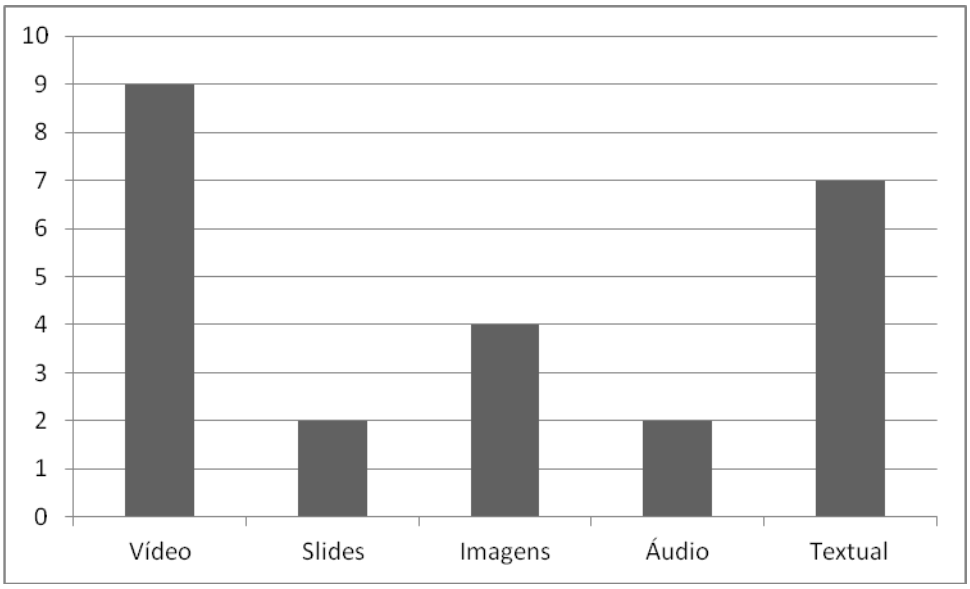
> *"Ele recomendou muitos assuntos afins, aumentando o leque de recomendações. E as recomendações me interessaram."*
>
> *"Os conteúdos sugeridos ficaram bastante relacionados com os temas de interesse, considerei a proposta muito relevante para a área educacional. A utilização do Facebook foi uma grande ideia, pois é uma rede que pode ser segmentada em interesses e amizades. Além de já prover uma forma de identificar os assuntos (curtir)."*

A limitação dos repositórios em relação ao número de recursos educacionais disponíveis e em relação à especificidade dos temas também foi notada e apontada pelos usuários:

> *"O conteúdo do repositório limitou uma recomendação mais condizentes com alguns dos interesses. Foi possível ver que a recomendação funciona, porém quando eram oriundas do repositório percebeu-se a falta de materiais."*
>
> *"No caso do repositório ele me recomendou muitos assuntos que são da área de computação, mas que não estavam dentro do tema que eu estava pesquisando."*
>
> *"Os resultados do repositório foram muito específicos, quando os temas eram mais amplos as recomendações não eram boas ou não era apresentada nenhuma recomendação."*

Os participantes apontaram que os recursos educacionais disponíveis em vídeos e em material textual são seus preferidos para estudar, conforme mostrado no gráfico 10.

Gráfico 11 - Tipo de mídia preferida pelo usuário.



Fonte: elaborado pelo autor.

Todos os usuários apontaram ter dois ou mais tipos de mídia como preferência para estudar. Esse fato, unido ao baixo índice de aceitação plena, 0,25, notado no estudo de caso para esse item, pode indicar que a extração da mídia preferida pelo usuário, através das mídias curtidas e compartilhadas, conforme foi realizada nesta proposta, provavelmente, não representa a totalidade das preferências dos usuários, sendo necessário, por exemplo, desenvolver um ranking dos tipos de mídias mais utilizadas e avaliar a precisão desta nova maneira de realizar a extração. Há ainda a possibilidade de usar o tipo de mídia de cada recomendação bem avaliada pelo usuário para criar um parâmetro de mídias preferidas.

6.2.7 **Ameaças a validade**

As ameaças à validade desse estudo serão tratadas em relação à: (1) validade do constructo, que é confiança que se tem de que as medidas que estão sendo estudadas realmente representam aquilo que o pesquisador tem em mente e o que é investigado de acordo com as questões de pesquisa; (2) validade interna, que é a confiança que se tem de que o efeito observado é realmente devido à manipulação feita, e não a outros fatores; (3) à validade externa que é a confiança que se tem que o efeito observável é generalizável (WOHLIN *et al.*, 2012).

Entende-se que uma ameaça em relação ao constructo é aplicação do questionário tendo uma situação hipotética, onde o usuário foi estimulado a analisar as recomendações supondo que ele estava, naquele momento, necessitando de uma recomendação para aquele

tema. Sabe-se que uma análise mais minuciosa precisaria ser feita para reconhecer a utilidade de um recurso educacional para o usuário ao longo do tempo, uma vez que fatores externos, tais como o uso do recurso, a atividade que o usuário está realizando no momento da recomendação, a qualidade do recurso, entre outros, podem exercer influência no quão relevante uma recomendação é para o usuário.

Em relação à validade interna deste estudo, uma ameaça é a influência que a qualidade dos recursos educacionais exerce sobre o interesse do usuário, ou seja, é possível que a qualidade do material desperte maior ou menor interesse sobre o usuário, fazendo com que ele avalie a recomendação sob essa ótica. Outra ameaça é a catalogação do recurso educacional, já que se a catalogação não tiver sido feita de modo eficiente pode prejudicar a recuperação do recurso ou o cálculo de aderência ao perfil do usuário. Em relação a essas duas ameaças foi adotado o critério de usar apenas recursos produzidos e catalogados por professores com domínio nas suas áreas de atuação, sendo que tais materiais já são usados com frequência em disciplinas do curso de Ciência da Computação. No entanto, o mesmo não pode ser feito para os recursos originados da DBPEDIA e do Youtube.

Em relação à validade externa, as principais ameaças que poderiam impedir a generalização dos resultados obtidos para qualquer usuário da rede social Facebook são a restrição da área dos recursos educacionais catalogados, conforme já mencionado anteriormente e a possibilidade de que os usuários não tenham na sua rede social nenhum tipo de interesse educacional que possa ser capturado. Em relação à primeira ameaça foram usadas outras fontes de recursos educacionais para buscar atender todos os diferentes tipos de interesses e também foi feita uma restrição do público alvo, buscando somente pessoas com interesse em áreas da Computação. Porém, ainda deve-se considerar a necessidade de uma ampliação das áreas de interesse dos recursos catalogados no repositório do BROAD ou o uso de outros repositórios para gerar recomendações. Em relação à segunda ameaça, o contexto de aplicação do estudo de caso já foi limitado a participantes que expressam seus interesses educacionais através das suas interações na rede social Facebook, já que esse é um requisito básico para que o modelo proposto seja eficaz.

7 CONSIDERAÇÕES FINAIS

Este capítulo apresenta as considerações finais sobre esta dissertação, argumenta a importância da arquitetura do BROAD-RSI, as principais contribuições deste trabalho, suas limitações e os trabalhos futuros.

7.1 CONCLUSÕES

Esta pesquisa apresentou uma proposta de arquitetura, nomeada BROAD-RSI, capaz de extrair e explorar as informações disponíveis na rede social Facebook para identificar características de perfil e de contexto dos usuários e, a partir dessas características, fazer recomendações de recursos educacionais individualizadas e personalizadas aos usuários.

Os resultados obtidos a partir da realização das provas de conceito e do estudo de caso mostraram a viabilidade de aplicação da proposta apresentada e trouxeram indícios de que as duas hipóteses apresentadas neste trabalho.

Os objetivos propostos neste trabalho foram atendidos através da arquitetura do BROAD-RSI que engloba a extração de características de perfil e de contexto e a recomendação de recursos educacionais de acordo com essas características. A implementação da proposta mostrou ser viável o envio de recomendações educacionais aos usuários nas suas redes sociais, usando assim todos os recursos disponíveis nesses ambientes e explorando o tempo que os usuários usam na interação em redes sociais.

Algumas conclusões podem ser alcançadas após o desenvolvimento desta pesquisa, a citar:

- A descoberta de interesses implícitos, relacionados àqueles expressos através das interações dos usuários nas redes sociais, permite que sejam geradas recomendações a respeito de temas correlatos, ampliando a possiblidade do usuário ter acesso às informações que possam ser relevantes ao seu aprendizado.
- O uso de dados ligados e de um repositório de vídeos no processo de recomendação pode ampliar as possibilidades de recomendações oferecidas aos usuários, não ficando restritos a um repositório de objetos de aprendizagem com temas específicos.
- O envio de recomendações educacionais aos usuários dentro das suas redes sociais permite que eles utilizem os recursos de interação desses ambientes,

compartilhem suas experiências a respeito de um recurso educacional recomendado e ainda utilizem o longo tempo e a facilidade de uso da rede social para receber recomendações educacionais.

- A recomendação de pessoas que possuem interesses comuns pode contribuir para troca de informações entre os usuários do BROAD-RSI formando assim redes de interesse a respeito de um determinado tema.

A proposta apresentada neste trabalho é bem ampla, contemplando diferentes áreas de conhecimento e, por isso, entende-se que a mesma pode ser refinada e melhorada em cada uma das etapas, buscando aumentar os índices de aceitação tanto da extração das características de perfil e contexto, quanto das recomendações.

7.2 CONTRIBUIÇÕES

O sistema de recomendação apresentado neste trabalho traz aspectos inovadores no sentido de extrair e considerar informações geradas espontaneamente nas redes sociais, mais especificamente no Facebook, bem como gerar recomendações individuais, especificamente educacionais, coerentes com os interesses e as preferências identificados a partir da análise das informações oriundas desse sistema.

O desenvolvimento da arquitetura do BROAD-RSI, do protótipo, das provas de conceito e do estudo de caso mostraram a viabilidade técnica e a aceitação dos usuários em relação à extração das informações, seu enriquecimento e das recomendações feitas.

A extração de características de perfil e de contexto a partir de uma rede social não obriga os usuários a manterem seus dados atualizados, tampouco os obriga a ter que declarar constantemente seus interesses para que recomendações educacionais sejam geradas.

A geração automática de recomendações, considerando as características individuais dos usuários pode auxiliá-los na tarefa de encontrar recursos educacionais relevantes e compatíveis com o seu perfil, podendo assim despertar maior interesse a respeito de um determinado conteúdo.

A extração e recomendação de recursos educacionais através de fontes variadas, além do repositório de objetos de aprendizagem, tais como conjuntos de dados ligados e o site de vídeos Youtube contribui para o campo de pesquisa em recomendação de recursos educacionais, mostrando a possibilidade de explorar essas fontes, que possuem atualização colaborativa e dispõe de recursos que podem ser úteis para diferentes usuários nos mais variados temas.

O envio da recomendação diretamente na rede social do usuário contribui para a facilidade de acesso aos recursos educacionais e permite que o usuário use todos os recursos de interação fornecidos por esses ambientes.

Outra contribuição é a possibilidade da recomendação de pessoas, possibilitando a formação de uma rede de interesses em torno de um determinado tema. O cálculo de similaridade entre os interesses das pessoas aproxima de um determinado usuários todas as pessoas com as quais ele possui interesses comuns, informando a ele quais dentre elas têm maior similaridade, indicando assim bons parceiros para estudo e pesquisa, por exemplo.

7.3 LIMITAÇÕES

A API usada neste trabalho como apoio na extração de informações, a AlchemyAPI, apresentou bons resultados, porém possui limitações, principalmente em relação aos idiomas dos textos analisados. Alguns métodos não estão disponíveis para todos os idiomas, inclusive para o Português. Por isso, observou-se que os resultados apresentados são mais satisfatórios quando os textos estão no idioma Inglês. O desenvolvimento de ontologias em português ameniza este problema, porém o problema da linguagem continua sendo considerado como uma limitação.

O número de recursos educacionais disponíveis no repositório do BROAD-RSI é uma limitação para a geração de recomendações relevantes para os usuários. Acredita-se que se o repositório dispusesse de mais recursos e se esses abordassem um conjunto maior de temas, as recomendações seriam melhores avaliadas pelos usuários.

Houve uma sugestão em relação ao uso de grupos dos usuários no Facebook, que não foram considerados na recomendação, conforme mostra depoimento abaixo:

> *"Além das páginas curtidas, poderia ser pesquisado os grupos que participamos, pois eu e outras pessoas no meu Facebook, por exemplo, participamos de vários grupos de interesse educacional, onde funciona como fóruns para troca de experiências daquilo que temos interesse em conhecer ou especializar."*

Os grupos dos quais os usuários participam foram considerados na extração, são representados semanticamente através da ontologia SIOC, e submetidos à extração de entidades, conceitos e palavras-chaves para a extração dos interesses. No entanto, a recomendação não considera os grupos de usuários, uma vez que não foi possível identificar ou inferir uma categorização dos grupos para então determinar aqueles que tratam de assuntos

educacionais. Sendo assim, essa é uma limitação deste trabalho que deve ser explorada em pesquisas futuras, buscando soluções para categorizar os temas discutidos nos grupos ou adaptar a abordagem para uma recomendação para grupos e não individualmente para participantes de grupos.

O presente trabalho se limitou a buscar recursos educacionais que contivessem termos exatos aos interesses extraídos e enriquecidos, não tendo explorado todas as possibilidades das técnicas de recuperação de informação (MANNING; RAGHAVAN, 2009), como por exemplo buscando por partes do nome ou a combinação entre diferentes interesses. O uso do Lucene na recuperação dos recursos educacionais no repositório do BROAD é um passo inicial para explorar melhor as possibilidades de melhoria na recuperação dos documentos, já que através desta API existem inúmeras possibilidades a serem exploradas nesse sentido.

7.4 TRABALHOS FUTUROS

A pesquisa, no sentido de criar um ambiente educacional de recomendação usando informações extraídas de redes sociais, pode ser ampliada sobre diversos aspectos, tais como: (1) extração de informações de perfil e de contexto do usuário; (2) uso das redes sociais para Extração de Informação e envio de recomendações; (3) abordagens de recomendação; (4) comparação entre propostas de recomendação sensíveis ao contexto.

Em relação à Extração de Informação seria importante ampliar o conjunto de informações usadas na extração, seja através da inclusão de postagens individuais em redes sociais, postagens em grupo e compartilhamentos; seja através da inclusão, junto com as redes sociais, de outros sistemas que possam complementar essa extração, tais como ambientes virtuais de aprendizagem, sistemas acadêmicos e sistemas de apoio a MOOCS. No caso de inclusão de outros sistemas seria muito interessante também avaliar a qualidade do perfil extraído nos diferentes sistemas.

A Extração de Informação pode ser ampliada também através da construção de novas ontologias, que englobem diferentes áreas do conhecimento. Cabe ressaltar a importância de ampliar toda a proposta de extração para diferentes idiomas, já que as ontologias e as Técnicas de Extração de Informação usadas mostraram resultados satisfatórios para os idiomas inglês e português, mas não foram ampliadas para os demais.

O uso de dados ligados nesta proposta poderia contribuir também para o enriquecimento dos interesses. Algumas experiências nesse sentido foram realizadas ao longo

do desenvolvimento desta proposta, mostrando ser uma alternativa que apresenta bons resultados. No entanto, tendo em vista o uso da AlchemyAPI, que em conjunto com as ontologias de domínio atendeu ao problema que estava sendo proposto, e ainda, a necessidade de delimitar o escopo do trabalho, não foi possível explorar completamente esta possiblidade.

A ampliação da proposta apresentada nesta dissertação para outras redes sociais também é uma sugestão de trabalhos futuros. A proposta foi projetada para sofrer poucas adaptações em caso de expansões para outras redes, sendo necessário, no entanto, considerar as peculiaridades das outras redes sociais, como por exemplo, as permissões e as formas de extração de informações de cada ambiente.

O uso de agentes de software integrados a essa proposta também seria um trabalho interessante, já que eles permitiriam a automatização e integração de várias atividades entre o BROAD-RSI e os sistemas usados para a Extração de Informação e envio de recomendação. Seria um trabalho motivador, por exemplo, integrar todas as redes sociais usadas pelos usuários com o sistema de recomendação, através de agentes de software que poderiam distribuir as recomendações educacionais geradas pelo BROAD-RSI para todas as redes sociais, em situações e momentos diferentes.

Por fim, a própria abordagem de recomendação usada neste trabalho pode ser refinada e melhorada, considerando, por exemplo, as recomendações feitas dentro da rede de pessoas que possuem interesses comuns, às avaliações das recomendações realizadas pelos usuários do BROAD-RSI, a interação do usuário dentro do BROAD-RSI, através de logs ou mineração de dados, por exemplo. A própria abordagem usada neste trabalho pode ser refinada, buscando melhorias nos cálculos de aderência de um recurso educacional ao perfil e ao contexto do usuário. Há melhorias a serem feitas também em relação à recuperação dos documentos que são recomendados aos usuários.

A categorização eficaz dos grupos nos quais o usuário participa nas redes sociais e a consideração dos interesses extraídos deles na recomendação também é um possível trabalho futuro, bem como a adaptação da proposta para a recomendação para grupos.

REFERÊNCIAS BIBLIOGRÁFICAS

ADOMAVICIUS, Gediminas; TUZHILIN, Alexander. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **IEEE Transactions on Knowledge and Data Engineering** v. 17, n. 6, p. 734–749 , 2005.

ALHARBI, Mafawez T; PLATT, Amelia; AI-BAYATTI, Ali H. Context-aware Personal Learning Environment. **The 7th International Conference for Internet Technology and Secured Transactions (ICITST-2012)** p. 692–697 , 2012.

ALLEN, Mark. *RestFB - A Lightweight Java Facebook Graph API and Old REST API Client*. Disponível em: <http://restfb.com/>.

ÁLVAREZ, Alberto Cáceres. **Extração de informação de artigos científicos: uma abordagem baseada em indução de regras de etiquetagem**. USP - Universidade de São Paulo, 2007. Disponível em: <http://lakh.unm.edu/handle/10229/35211>.

ANTONIOU, Grigoris; HARMELEN, Frank Van. **A Semantic Web Primer**. 2º. ed. [S.l.]: MIT press, 2008. p. 264. 9780262012423.

AQUIN, Mathieu D. Linked Data for Open and Distance Learning. n. June , 2012.

ASSOCIATION FOR COMPUTING MACHINERY. *The ACM Computing Classification System*. Disponível em: <http://dl.acm.org/ccs_flat.cfm>.

AUER, Sören *et al.* DBpedia: A Nucleus for a Web of Open Data. **The Semantic Web** v. 4825, p. 722–735 , 2007.

BASILI, Victor R. *Software modeling and measurement: the Goal/Question/Metric paradigm* .**Quality**. [S.l: s.n.]. Disponível em: <http://portal.acm.org/citation.cfm?id=137076>. , 1992

BAZIRE, Mary; BRÉZILLON, Patrick. Understanding Context Before Using It. **5th International and Interdisciplinary Conference** v. LNAI 3554, p. 29–40 , 2005.

BERNERS-LEE, Tim. *Linked Data*. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The Semantic Web. **Scientific American** v. 284, n. 5, p. 34–43 , 2001.

BITTENCOURT, Ig Ibert *et al.* Research Directions on Semantic Web and Education. **Interdisciplinary Studies in Computer Science** v. 19, n. June, p. 59–66 , 2008.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked Data - The Story So Far. **International Journal on Semantic Web and Information Systems (IJSWIS)** v. 5, n. 3, p. 1–22 , 2009.

BOYD, Danah M.; ELLISON, Nicole B. Social Network Sites: Definition, History, and Scholarship. **Journal of Computer-Mediated Communication** v. 13, n. 1, p. 210–230 , 17

out. 2007. Disponível em: <http://doi.wiley.com/10.1111/j.1083-6101.2007.00393.x>. Acesso em: 9 jul. 2014.

BRESLIN, John *et al.* SIOC: an approach to connect web-based communities. **International Journal of Web Based Communities - IJWBC** v. 2, n. 2, p. 133–142 , 2006.

BRICKLEY, Dan; MILLER, Libby. *FOAF Vocabulary Specification 0.99*. Disponível em: <http://xmlns.com/foaf/spec/#term_Document>. Acesso em: 7 jan. 2015.

BURKE, Robin. Hybrid Recommender Systems: Survey and Experiments. **User modeling and user-adapted interaction** v. 12, n. 4, p. 331–370 , 2002.

CAMPOS, Fernanda *et al.* Rede de Ontologias: apoio semântico a linha de produtos de objetos de aprendizagem. **Anais do Simpósio Brasileiro de Informática na Educação** v. 23, n. Sbie, p. 26–30 , 2012.

CASAGRANDE, Marcus Filipi Rosso; KOZIMA, Gustavo; WILLRICH, Roberto. Técnica de Recomendação Baseada em Metadados para Repositórios Digitais Voltados ao Ensino. **XXIV Simpósio Brasileiro de Informática na Educação (SBIE 2013)** v. 24, p. 677–686 , 22 nov. 2013.

CASTELLS, Manuel. Network Society: A Cross-cultural Perspective. **Art Journal** v. 59, p. 5 , 2004.1843765055 (hbk.).

CHOWDHURY, Gobinda G. Natural language processing. **Annual review of information science and technology** v. 37, p. 51–89 , 2003.8789809289.

COMITÊ GESTOR DA INTERNET NO BRASIL – CGI.BR. **Pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras : TIC Educação 2013**. 1. ed. São Paulo: [s.n.], 2014. p. 1–507. 978-1-4799-1056-4.

CORPORATION. Oracle. **The Java EE 6 Tutorial**. **The Java EE 6 Tutorial** [S.l: s.n.], 2013. Disponível em: <http://docs.oracle.com/javaee/6/tutorial/doc/gijqy.html>.

COSTA, Evandro; AGUIAR, Janderson; MAGALHÃES, Jonathas. Sistemas de Recomendação de Recursos Educacionais: conceitos, técnicas e aplicações. **Anais da Jornada de Atualização em Informática na Educação** n. Cbie, p. 57–78 , 19 dez. 2013.

CYGANIAK, Richard; JENTZSCH, Anja. *The Linking Open Data cloud diagram*. Disponível em: <http://lod-cloud.net/>. Acesso em: 1 dez. 2014.

DE MELLO, Rafael M. *et al.* Checklist-based inspection technique for feature models review. **Proceedings - 2012 6th Brazilian Symposium on Software Components, Architectures and Reuse, SBCARS 2012** p. 140–149 , 2012.9780769549170.

DEVEDZ, Vladan. Education and the Semantic Web. **International Journal of Artificial Intelligence in Education** v. 14, p. 39–65 , 2004.

DEY, Anind K; ABOWD, Gregory D. Towards a Better Understanding of Context and Context-Awareness. **Handheld and ubiquitous computing.** p. 304–307 , 1999.

DUVAL, Erik; HODGINS, Wayne. A LOM Research Agenda. p. 1–17 , 2003. Disponível em: <http://www.www2003.org/cdrom/papers/alternate/P659/p659-duval.pdf>.

DWIVEDI, Pragya; BHARADWAJ, Kamal K. E-Learning Recommender System for Learners in Online Social Networks Through Association Retrieval. **Proceedings of the CUBE International Information Technology Conference on - CUBE '12** p. 676 , 2012. Disponível em: <http://dl.acm.org/citation.cfm?doid=2381716.2381846>.9781450311854.

FERNANDES, Gustavo. **CAPTURA DE PERFIS DINÂMICOS DE USUÁRIOS A PARTIR DA ANÁLISE DE MENSAGENS EM REDES SOCIAIS**. Universidade Federal do Estado do Rio de Janeiro, 2013.

FERREIRA, Luís Gustavo Araujo; BARBOSA, Jorge Luis Victória; GLUZ, João Carlos. Um Modelo de Recomendação Ubíqua de Conteúdo para Grupos de Aprendizes. **XXIV Simpósio Brasileiro de Informática na Educação (SBIE 2013)** v. 24, n. Cbie, p. 697–706 , 22 nov. 2013.

FINKENZELLER, Klaus. **RFID handbook: radio-frequency identification fundamentals and applications**. New York: [s.n.], 1999. p. 151–158.

FOUNDATION, The Apache Software. *Apache Jena*. Disponível em: <https://jena.apache.org/index.html>. Acesso em: 7 jan. 2015.

FRADE, Rodrigo Valença C. *et al.* Um Ambiente Virtual 3D Multiagente com Recomendação Personalizada de Objetos de Aprendizagem. **XXV Simpósio Brasileiro de Informática na Educação (SBIE 2013)** v. 25, n. Cbie, p. 1068–1077 , 2014.

FRITZEN, Eduardo; SIQUEIRA, Sean W.M.; DE ANDRADE, Leila C.V. Recuperação Contextual de Informação na Web para Apoiar Aprendizagem Colaborativa em Redes Sociais. **XXIII Anais do Simpósio Brasileiro de Informática na Educação** v. 23, n. Simpósio Brasileiro de Informática na Educação, p. 26–30 , 2012.

GALLEGO, Daniel *et al.* Incorporating Proactivity to Context-Aware Recommender Systems for E-Learning. 2013, [S.l: s.n.], 2013. p.1–6. 9781479904624. .

GASPARINI, Isabela *et al.* Modelo do usuário sensível ao contexto cultural em um sistema e-learning adaptativo Culturally-aware user model in an adaptive e-learning system. **Informática na Educação: teoria e prática** v. 14, p. 123–135 , 2011.

GREER, Jim *et al.* The Intelligent Helpdesk: Supporting Peer-Help in a University Course. 1998, [S.l.]: Springer-Verlag, 1998. p.494–503.

GRUBER, Thomas R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. **International journal of human-computer studies** v. 43, n. 5, p. 907–928 , 1995.

GUARINO, Nicola. Understanding, Building, And Using Ontologies. **International Journal of Human-Computer Studies** v. 46, n. 2, p. 293–310 , 1988.

GUARINO, Nicola; GIARETTA, Pierdaniele. Ontologies and Knowledge Bases. **Towards very large knowledge bases: knowledge building & knowledge sharing** v. 25 , 1995.

HEATH, Tom *et al.* Assembling and Applying an Education Graph based on Learning Resources in Universities. 2012, [S.l: s.n.], 2012.

HODGINS, H Wayne. The Future of Learning Objects. , 2004.

IBOPE. **Brasileiros caem na rede social**. [S.l: s.n.], 2011. Disponível em: <http://www.ibope.com.br/pt-br/noticias/Paginas/Brasileiros caem na rede social.aspx>.

JIANG, Jing. INFORMATION EXTRACTION FROM TEXT. In: AGGARWAL, Charu C.; ZHAI, ChengXiang (Orgs.). . **Mining text data**. Boston, MA: Springer US, 2012. p. 11–41. 978-1-4614-3222-7.

KARAMPIPERIS, Pythagoras; SAMPSON, Demetrios. Adaptive Learning Resources Sequencing in Educational Hypermedia Systems. **Educational Technology & Society** v. 8, p. 128–147 , 2005.

KITCHENHAM. Barbara. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. **Technical report, EBSE Technical Report EBSE** [S.l: s.n.], 2007.1595933751.

KLANJA-MILICEVIC, Aleksandra *et al.* Personalisation of Programming Tutoring System Using Tag-Based Recommender Systems. **2012 IEEE 12th International Conference on Advanced Learning Technologies** p. 666–667 , jul. 2012.978-1-4673-1642-2.

KLASNJA-MILICEVIC, Aleksandra *et al.* Integration of recommendations and adaptive hypermedia into java tutoring system. **Computer Science and Information Systems** v. 8, n. 1, p. 211–224 , 2011. Disponível em: <http://www.doiserbia.nb.rs/Article.aspx?ID=1820-02141000021K>. Acesso em: 11 nov. 2014.

KOVACIC, Tomaz. **Evaluating Web Content Extraction Algorithms**. UNIVERSITY OF LJUBLJANA, 2012.

LEHMAN, Rosemary. Learning object repositories. New directions for adult and continuing education. **Embroidery** v. 56, n. 113, p. 57–66 , 2007. Disponível em: <www.interscience.wiley.com>.1092-1095 (Print)r1092-1095 (Linking).

LETHBRIDGE, Timothy C.; SIM, Susan Elliott; SINGER, Janice. Studying software engineers: Data collection techniques for software field studies. **Empirical Software Engineering** v. 10, p. 311–341 , 2005.1382-3256.

MANNING, Christopher D.; RAGHAVAN, Prabhakar. **An Introduction to Information Retrieval**. [S.l.]: Cambridge University Press, 2009. p. 1. 0521865719.

MANOUSELIS, Nikos *et al.* Recommender Systems in Technology Enhanced Learning. **Recommender systems handbook**. [S.l.]: Springer US, 2011. p. 387–415.

MCIVER, J. P; CARMINES, E. G. **Unidimensional scaling**. [S.l.]: Thousand Oaks, CA: Sage, 1981.

MCWHORTER, Rochell R; LINDHJEM, Krystle a. Virtual Learning Environments. p. 15–17 , 2013.

MOENS, Marie-francine. Information Extraction: Algorithms and Prospects in a Retrieval Context. , 2006.9781402049873.

NERY, Thiago *et al.* BROAD Project: Semantic Search and Application of Learning Objects. **IEEE Technology and Engineering Education (ITEE)** v. 7, n. 3 , 2012.

OHLER, Jason. The Semantic Web in Education. n. 4, p. 7–9 , 2008.

PAPPANO, Laura. The Year of the MOOC. **The New York Times** p. 1–7 , 2012.

PARK, Seung-Taek; CHU, Wei. Pairwise preference regression for cold-start recommendation. **Proceedings of the third ACM conference on Recommender systems - RecSys '09** p. 21 , 2009.9781605584355.

PEREIRA, Crystiam Kelle; CAMPOS, Fernanda; BRAGA, Regina; *et al.* Elementos de Contexto em Sistemas de Recomendação no Domínio Educacional: um Mapeamento Sistemático. **XIX Conferência Internacional sobre Informática na Educação** v. 19, p. 10 , 2014. Disponível em: <http://www.tise.cl/volumen10/TISE2014/tise2014_submission_57.pdf>.

PEREIRA, Crystiam Kelle; CAMPOS, Fernanda; STRÖELE, Victor; *et al.* Extração de Características de Perfil e de Contexto em Redes Sociais para Recomendação de Recursos Educacionais. **XXV Simpósio Brasileiro de Informática na Educação (SBIE 2014)** v. 25, n. Cbie, p. 506–515 , 2014a.

PEREIRA, Crystiam Kelle; CAMPOS, Fernanda; STRÖELE, Victor; *et al.* Extracting User Profile and Context from Social Networks and Virtual Learning Environments to Recommend Educational Resources. **Proceedings of the 13th International Conference WWW/Internet 2014** p. 171–178 , 2014b.

PISKORSKI, Jakub; YANGARBER, Roman. Multi-source, Multilingual Information Extraction and Summarization. Theory and Applications of Natural Language Processing. p. 23–50 , 2013. Disponível em: <http://link.springer.com/10.1007/978-3-642-28569-1>. Acesso em: 12 dez. 2014.978-3-642-28568-4.

RECKER, Mimi M.; WILEY, David a. A Non-authoritative Educational Metadata Ontology for Filtering and Recommending Learning Objects. **Interactive Learning Environments** v. 9, n. 3, p. 255–271 , 1 dez. 2001. Disponível em: <http://www.tandfonline.com/doi/abs/10.1076/ilee.9.3.255.3568>.

REIS, Gustavo; BARRÉRE, Eduardo. Recomendação Colaborativa de Conteúdos Educacionais para Dispositivos Portáteis. **XXV Simpósio Brasileiro de Informática na Educação (SBIE 2013)** v. 25, p. 934–943 , 2014.

REZENDE, Paulo *et al.* Broad-rs: uma arquitetura para recomendação de objetos de aprendizagem sensível ao contexto usando agentes. **Anais do X Congresso Brasileiro de Ensino Superior a Distância** p. 11–13 , 2013.

RIBEIRO, Francisco Adelton Alves; FONSECA, Luis Carlos Costa; FREITAS, Miguel De Sousa. Recomendando Objetos de Aprendizagem a partir das hashtags postadas no Moodle. **XXIV Simpósio Brasileiro de Informática na Educação (SBIE 2013)** v. 25, p. 82–91 , 22 nov. 2013. Disponível em: <http://www.br-ie.org/pub/index.php/sbie/article/view/2486>. Acesso em: 8 jan. 2015.

RUNESON, P *et al.* **Case Study Research in Software Engineering. Guidelines and Example**. Wiley, Hob ed.[S.l: s.n.], 2012.

SARAWAGI, Sunita. Information Extraction. v. 1, n. 3, p. 261–377 , 2008.

SEBASTIANI, Fabrizio. Machine Learning in Automated Text Categorization. **ACM computing surveys (CSUR)** v. 34, n. 1, p. 1–47 , 2002. Disponível em: <http://arxiv.org/abs/cs/0110053>.0360-0300.

SHANI, Guy; GUNAWARDANA, Asela. Evaluating recommendation systems. **Recommender systems handbook** p. 257–298 , 2011. Disponível em: <http://link.springer.com/chapter/10.1007/978-0-387-85820-3_8>.9780387858197.

SUDHANA, Kalla. Madhu; RAJ, V. Cyril; SURESH, R. M. An ontology-based framework for context-aware adaptive e-learning system. **2013 International Conference on Computer Communication and Informatics** p. 1–6 , jan. 2013.978-1-4673-2907-1.

VANOOZZI. Monica; BRIDGESTORK. Laura. **QS Top Un**. [S.l: s.n.], 2013.

VERBERT, Katrien *et al.* Context-Aware Recommender Systems for Learning: A Survey and Future Challenges. v. 5, n. 4, p. 318–335 , 2012a.

VERBERT, Katrien *et al.* Context-Aware Recommender Systems for Learning: A Survey and Future Challenges. v. 5, n. 4, p. 318–335 , 2012b.

VICARI, Rosa Maria *et al.* The OBAA Proposal for Learning Objects Supported by Agents. **Syntax And Semantics** n. Aamas, p. 9–12 , 2010.

VIEIRA. Vaninha *et al*. **Uso e Representação de Contexto em Sistemas Computacionais**. [S.l: s.n.], 2006.

VIEIRA, Vaninha; TEDESCO, Patricia; SALGADO, Ana Carolina. Modelos e Processos para o Desenvolvimento de Sistemas Sensíveis ao Contexto. 2009, [S.l: s.n.], 2009. p.381–431.

WANG, Shu-Lin; WU, Chun-Yi. Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. **Expert Systems with Applications** v. 38, n. 9, p. 10831–10838 , set. 2011.

WOHLIN, C *et al.* **Experimentation in Software Engineering**. [S.l.]: Springer Berlin Heidelberg, 2012.

YIN, Robert K. **Estudo De Caso - Planejamento E Metodos**. [S.l: s.n.], 2001. p. 206. 8573078529.

YU, Liyang. **A developer's guide to the semantic web**. [S.l: s.n.], 2011. p. I–XIX.

ZHENG, Yanlin; LI, Luyi; ZHENG, Fanglin. Context-Awareness Support for Content Recommendation in E-learning Environments. **2009 International Conference on Information Management, Innovation Management and Industrial Engineering** p. 514–517 , 2009.978-0-7695-3876-1.