

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA
COMPUTAÇÃO**

Ronaldo Lima Rocha Campos

**Modelo de Busca Inteligente e Recomendação de
Objetos de Aprendizagem em Repositórios
Heterogêneos**

Dissertação submetida à Universidade
Federal de Santa Catarina para a
obtenção do Grau de Mestre em
Ciências da Computação.

Orientador: Ricardo Azambuja Silveira
Co-orientadora: Rafaela Lunardi
Comarella

Florianópolis
2013

Catálogo na fonte elaborada pela biblioteca da
Universidade Federal de Santa Catarina

Campos, Ronaldo Lima Rocha

Modelo de busca inteligente e recomendação de objetos de aprendizagem em repositórios heterogêneos / Ronaldo Lima Rocha Campos ; orientador, Ricardo Azambuja Silveira ; co-orientadora, Rafaela Lunardi Comarella. - Florianópolis, SC, 2013.

87 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Ciência da Computação.

Inclui referências

1. Ciência da Computação. 2. Recuperação da Informação. 3. Repositórios. 4. Sistemas Multiagentes . 5. Ontologia. I. Silveira, Ricardo Azambuja. II. Comarella, Rafaela Lunardi. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Computação. IV. Título.

Ronaldo Lima Rocha Campos

Modelo de Busca Inteligente e Recomendação de Objetos de Aprendizagem em Repositórios Heterogêneos

Esta Dissertação foi julgada adequada para obtenção do Título de Mestre em Ciências da Computação e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciências da Computação.

Florianópolis, 22 de fevereiro de 2013.

Prof. Dr. Ronaldo dos Santos Mello
Coordenador do Curso

Banca Examinadora:

Orientador Prof., Dr. Ricardo Azambuja Silveira,
Universidade Federal de Santa Catarina

Co-Orientadora MSc. Rafaela Lunardi Comarella,
Universidade Federal de Santa Catarina

Profa, Dra. Anarosa Alves Franco Brandao,
Universidade de São Paulo

Prof., Dr. Roberto Willrich,
Universidade Federal de Santa Catarina

Profa., Dra. Silvia Modesto Nassar,
Universidade Federal de Santa Catarina

À minha família.

AGRADECIMENTOS

Agradeço a todos que de alguma forma contribuíram para que eu chegasse até aqui. Em especial aos membros da banca por suas considerações e ao meu orientador Ricardo Azambuja Silveira e minha co-orientadora Rafaela Lunardi Comarella por me mostrarem os caminhos para a realização do presente trabalho.

RESUMO

Este trabalho propõe um modelo de sistema de busca inteligente e de recomendação para objetos de aprendizagem baseado em sistemas multiagentes. Para tal, o modelo é capaz de indexar e recuperar objetos de aprendizagem, armazenados em repositórios diferentes e heterogêneos. Ou seja, o mecanismo de busca inteligente abrange vários repositórios de objetos de aprendizagem, e os objetos podem ser descritos pelo uso de diferentes padrões de metadados. O modelo proposto utiliza ontologias para descrever o domínio de conhecimento para, em um primeiro momento, expandir a busca utilizando sinônimos ou outros termos relacionados e, em um segundo momento, a ontologia é utilizada para restringir os resultados apresentados, dentro do domínio de conhecimento do usuário, levando em conta as especificidades da área. Com o objetivo de melhorar a precisão, cobertura e recomendação de um objeto de aprendizagem, o sistema prevê um modelo de recomendação de objetos de aprendizagem utilizando informações e estatísticas dos usuários, ou seja, seu perfil, para realizar a ordenação dos resultados. Tal modelo foi implementado sobre o *framework* JADE e testado sobre o índice de duas bases de dados, com termos da área médica. Os resultados mostram uma boa cobertura e recuperação para objetos, por fazer uso dos termos expandidos para seus sinônimos e os resultados do perfil mostram uma melhora na ordem e recomendação dos resultados.

Palavras-chave: Recuperação da Informação, Repositórios, Ontologia, Sistemas Multiagentes.

ABSTRACT

This paper proposes a model of intelligent search engine and recommendation for learning objects based on multiagent systems. To this end, the model is able to index and retrieve learning objects stored in different and heterogeneous repositories. The intelligent search engine covers several repositories of learning objects, and the objects can be described in different metadata standards. The proposed model uses ontologies to describe the domain knowledge for expand the search terms, using synonyms or related terms and in a second step, the ontology is used to restrict the results within the domain of knowledge taking into account the specificities of the area. Aiming to improve accuracy, coverage and recommendation of learning objects the system provides a model for recommendation using information and statistics of the users your profile, and perform the ordering of results. This model was implemented on the JADE framework and tested on the content of two databases with use of medical terms. The results show good coverage and recovery for objects, by making use of expanded terms on their synonyms, and the profile results show an improvement in the results order and recommendation.

Keywords: Multiagent System, Recommender System, Ontology, Learning Object, Information Retrieval.

LISTA DE FIGURAS

Figura 1: Arquitetura do sistema multiagente	49
Figura 2: Fluxo da informação.....	50
Figura 3: Diagrama de agentes.	52
Figura 4: Diagrama de interação do Agente Buscador.....	54
Figura 5: Diagrama de interação do Agente Perfil - primeira etapa.....	57
Figura 6: Diagrama de interação do Agente Perfil - segunda etapa.	58
Figura 7: Modelo de ontologia de domínio.....	60
Figura 8: Diagrama de interação do Agente indexador.....	63
Figura 9: Representação gráfica do vocabulário DECS.....	68
Figura 10: Representação gráfica do vocabulário DECS - Visão detalhada	68
Figura 11: Representação gráfica dos resultados obtidos para Acidente Cerebral	73

LISTA DE QUADROS

Quadro 1: Mapeamento semântico de metadados.....	45
Quadro 2: Exemplo de Leilão	66
Quadro 3: metadados padrão para perfil especializado.....	69
Quadro 4: expansão dos termos.	72
Quadro 5: comparação entre os resultados do SMA e ARES para o termo Doenças Respiratórias.	74
Quadro 6: comparação entre os resultados do SMA e LUME para o termo Doenças Respiratórias.	75
Quadro 7: comparação entre os resultados do LUME e SMA para o termo Doenças Respiratórias na busca simples do LUME.	76
Quadro 8: Comparação entre os resultados do SMA e ARES para o termo Dengue.	77

LISTA DE TABELAS

Tabela 1: Total de objetos recuperados por repositório.....	71
Tabela 2: Comparativo entre a busca sem perfil no SMA.....	78
Tabela 3: Comparativo entre a busca com perfil no SMA.	78

LISTA DE ABREVIATURAS E SIGLAS

AVA – Ambiente Virtual de Aprendizagem
BIOE - Banco Internacional de Objetos Educacionais referência
CORDRA - *Content Object Repository Discovery and Registration/Resolution Architecture*
FEB - Federação Educa Brasil
LDAP - *Lightweight Directory Access Protocol*
LOM - *Learning Object Metadata*
OA - Objetos de Aprendizagem
OAI-ORE - *Open Archives Initiative Object Reuse and Exchange*
OAI-PMH - *Open Archives Initiative Protocol for Metadata Harvesting*
OBAA - Objetos de Aprendizagem Baseados em Agentes
RI - Repositório Institucional
ROA - Repositório de Objeto de Aprendizagem
SMA – Sistema Multiagente
UnA-SUS - Universidade Aberta do SUS
XML - *eXtensible Markup Language*
OWL – *Web Ontology Language*
RDF – *Resource Description Framework*

SUMÁRIO

1. INTRODUÇÃO	22
1.1 Objetivos	24
1.1.1 Objetivo Geral	24
1.1.2 Objetivos Específicos	24
1.2 Estrutura da Dissertação	25
2. FUNDAMENTAÇÃO TEÓRICA	27
2.1 Objetos de Aprendizagem	27
2.2 Metadados	28
2.3 Repositórios Digitais	31
2.4 Protocolos de disseminação e coleta	33
2.5 Ontologia	35
2.6 Sistemas Multiagentes	36
2.7 Sistemas de Recomendação	38
2.8 Trabalhos Correlatos	39
3. METODOLOGIA	43
4. O MODELO PROPOSTO	49
4.1 Agente Buscador	54
4.2 Agente Perfil	56
4.2.1 Ontologia	60
4.2.2 Base de Perfis	61
4.3 Agente Indexador	63
4.3.1 Desempenho	66
4.4 Implementação e Validação do Modelo	67
4.4.1 Testes	71
5. CONSIDERAÇÕES FINAIS	80
5.1 Sugestões para Trabalhos Futuros	83
6. REFERÊNCIAS	84

1. INTRODUÇÃO

Os esforços e investimentos provenientes de educadores e instituições de ensino, para elaborar materiais didáticos de qualidade, são consideráveis. Nos moldes atuais de educação a distância, a maior parte do conteúdo e do conhecimento deve estar presente na forma de materiais didáticos digitais. Esses materiais didáticos digitais quando dotados de um propósito e inseridos em um conceito educacional, são chamados de Objetos de Aprendizagem (OA).

Para a IEEE-LTSC (2002), "Um objeto de aprendizagem é definido como qualquer entidade, digital ou não, que pode ser utilizada para o ensino, aprendizagem ou formação". Assim, exemplos de Objetos de Aprendizagem incluem conteúdo multimídia, conteúdo instrucional, objetivos de aprendizagem, softwares, pessoas, organizações ou eventos referenciados durante a aprendizagem apoiada pela tecnologia. Esta é uma definição muito ampla, e não exclui qualquer pessoa, lugar, coisa ou ideia, uma vez que qualquer um destes pode ser referenciado durante o processo. Wiley (2000) restringe a definição de objeto de aprendizagem, para qualquer recurso digital utilizado para apoiar a aprendizagem. Assim, objetos de aprendizagem podem ser desde peças pequenas, tais como imagens, sons, vídeos até partes mais complexas, tais como cursos ou softwares.

A produção e utilização dos Objetos de Aprendizagem estão relacionadas também ao seu armazenamento e disseminação (divulgação). Os repositórios digitais são as ferramentas utilizadas para gerenciar e armazenar esses recursos (objetos de aprendizagem), bem como prover uma série de funcionalidades para permitir que diferentes tipos de objetos sejam armazenados, catalogados, e disponibilizados. Nesse contexto surgem os termos Repositório Institucional (RI) e Repositório de objetos de aprendizagem (ROA).

Os objetos de aprendizagem são produzidos com diversas tecnologias e costumam ser armazenados de diferentes formas, utilizando diferentes tecnologias e descritos por diferentes estruturas de metadados (DOWNES, 2001). Além disso, existe uma variedade de repositórios para atender diferentes demandas, desta forma, torna-se difícil a decisão de implementação e utilização de um determinado repositório. Nem sempre é feita a escolha certa de início, o que leva muitas vezes a adoção de outras opções e cenários (TARRANT *et al*, 2009). Os repositórios ainda provêm mecanismos de busca e recuperação dos Objetos de Aprendizagem neles contidos, que também são dependentes da tecnologia adotada, tanto pelo repositório quanto

pelo tipo de objetos armazenados.

A recuperação de objetos de aprendizagem, que leve em conta um conjunto de repositórios diferentes, é uma operação que não é trivial, devido à existência de diferentes padrões e especificações para a produção e armazenamento, e a grande distribuição de repositórios através da web. Geralmente os mecanismos de busca disponíveis na Internet recuperam a indicação de páginas a partir do seu conteúdo, e não de metadados. Além disso, é clara a falta de ferramentas de busca especializadas em Objetos de Aprendizagem eficientes, que faça a recomendação de objetos sobre um determinado tópico de conhecimento com base nos metadados que o descrevem, e que estejam localizados em diferentes repositórios, o que não permite reutilizar o leque de objetos produzidos (VIAN; SILVEIRA, 2010).

As ferramentas existentes utilizadas para recuperar informações sobre Objetos de Aprendizagem são geralmente baseadas em uma busca sintática. Esse tipo de pesquisa não é a forma mais eficiente para se recuperar Objetos de Aprendizagem, pois, como na web, acaba trazendo uma série de resultados irrelevantes, repetidos ou fora do contexto desejado (ROUSSET; REYNAUD, 2004). Nos repositórios, os Objetos de Aprendizagem estão catalogados com uma série de informações, que envolvem desde o tipo de recurso digital usado até sua área de informação. Estas informações são apresentadas sobre a forma de metadados e podem ser utilizadas em funções semânticas para recuperar, de forma mais precisa, os objetos desejados (LEE; TSAI; WANG, 2008).

Os repositórios podem trazer informações adicionais, tais como estatísticas e informações colaborativas sobre os Objetos de Aprendizagem. Tais informações ampliam as possibilidades de busca e recuperação. A informação pode estar organizada em forma de coleções, de acordo com as áreas de conhecimento ou utilizando sua própria taxonomia. O Repositório da Universidade Aberta do SUS (UnA-SUS, 2013) por exemplo, está segmentado por taxonomia de área temática, o PubMed (2013) está organizado por conjuntos de bases de dados, já o Banco Internacional de Objetos (BIOE, 2013) está organizado sobre níveis de ensino. Percebe-se que, para recuperar Objetos de Aprendizagem em repositórios, onde cada objeto tem seu propósito, e detém uma descrição formal do seu conteúdo, o uso de aspectos semânticos da informação fornece um recurso mais preciso para sua recuperação.

Este contexto mostra que há uma diversidade de padrões e faltam ferramentas inteligentes para facilitar a recuperação de Objetos de Aprendizagem em um cenário heterogêneo de repositórios, limitando,

assim, de forma expressiva a recuperação de objeto de aprendizagem.

Diante da análise dos processos de criação e armazenamento de Objetos de Aprendizagem, concebeu-se a proposta de um modelo de busca e recuperação que permite encontrar objetos de forma federada, tal como *Content Object Repository Discovery and Registration/Resolution Architecture* (CORDRA, 2013) e Federação Educa Brasil (FEB, 2013), em diferentes repositórios heterogêneos, usando diferentes conjuntos de metadados e considerando os aspectos semânticos, e destacando os objetos que atendam a critérios de qualidade, como os propostos por Gil, De La Prieta e Rodríguez (2011).

Isso justifica a proposta de um modelo de busca e recomendação inteligente que facilite a recuperação de objetos de aprendizagem disponíveis em um cenário de repositórios heterogêneos.

O modelo proposto baseia-se no uso de Sistemas Multiagentes, pois os sistemas multiagentes são caracterizados pela capacidade de modelar agentes inteligentes capazes de se adaptar a diferentes cenários, podendo atuar de forma autônoma, de cooperar e se comunicar uns com os outros para atingir um objetivo comum (WOOLDRIDGE, 2002). Utiliza também ontologias para representar domínios de conhecimento, através da qual os agentes atuam, em um primeiro momento, expandindo os termos utilizados como chave de busca e, em um segundo momento, restringindo os resultados apresentados de acordo com o perfil do usuário ou o contexto da busca.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Desenvolver um modelo de sistema inteligente de busca, e recuperação e recomendação para Objetos de Aprendizagem armazenados em diferentes repositórios com heterogeneidade de metadados.

1.1.2 Objetivos Específicos

- Definir uma arquitetura que possibilitem a utilização do perfil do usuário como critério de recomendação.

- Elaborar um mapeamento de correspondência dos campos dos principais padrões de metadados para objetos de aprendizagem.
- Identificar, de forma inteligente, informações sobre os termos da busca, de forma a minimizar a descrição fornecida pelo usuário.
- Definir parâmetros de arquitetura dos agentes para realizar buscas em repositórios com diferentes estruturas e métodos de acesso.
- Avaliar o modelo proposto

1.2 ESTRUTURA DA DISSERTAÇÃO

Este trabalho está organizado em cinco capítulos, de forma a definir um modelo de sistema inteligente de busca, e recuperação e recomendação para Objetos de Aprendizagem armazenados em diferentes repositórios com heterogeneidade de metadados. Esse primeiro capítulo apresenta a contextualização do tema e são descritos os objetivos.

O capítulo dois é apresenta o referencial teórico do trabalho com uma revisão sobre os principais conceitos envolvidos na pesquisa envolvendo os conceitos básicos sobre objetos de aprendizagem e as especificações formais dos metadados. São definidos os conceitos de repositórios digitais e heterogeneidade. Para definição desses conceitos, são apresentadas técnicas para representação de conhecimento, ontologias e protocolos de disseminação e coleta. Ao final do capítulo é apresentado o conceito de sistema multiagente e uma sessão de trabalhos correlatos.

No terceiro capítulo estão descritos os passos desenvolvidos para a execução dessa pesquisa, ou seja, os procedimentos metodológicos adotados.

O quarto capítulo traz o modelo proposto para recuperação de objetos de aprendizagem, mostrando suas características e seus componentes. O modelo segue uma metodologia de descrição de sistemas multiagentes, especificando-se cada agente e suas interações, capacidades e serviços oferecidos. Também são apresentadas as etapas de implementação e desenvolvimento de um protótipo, onde cada componente do modelo é construído e relacionado com a tecnologia que

dá suporte a ela. Ao final do capítulo, são apresentados os procedimentos de validação do modelo com base no protótipo construído e os resultados alcançados.

As considerações finais são apresentadas no quinto capítulo, bem como propostas de trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação teórica dessa pesquisa, que abrange os conceitos básicos relacionados a objetos de aprendizagem, repositórios digitais, bem como as especificações formais de metadados para os OA. Além disso, fundamenta o uso de sistemas multiagentes combinado com técnicas de recuperação da informação para encontrar objetos em ambientes distribuídos e heterogêneos.

Segundo esta revisão da literatura os resultados obtidos apontam para o uso de ontologias de domínio para contextualizar cenários e áreas de conhecimentos específicos, enriquecendo a busca e indexação. E que o uso de técnicas de recomendação melhora a precisão, cobertura e o desempenho das ferramentas de busca.

2.1 OBJETOS DE APRENDIZAGEM

Os objetos de aprendizagem são recursos educacionais que podem ser usados no processo de aprendizagem suportada pela tecnologia (MCGREAL, 2004), descritos formalmente por seus metadados, armazenado em um repositório, e que pode ser combinado com outros objetos de aprendizagem para criar objetos maiores, como aulas e cursos (NASH, 2005). Eles podem ser construídos com base em texto, animações, apresentações, imagens, software, e devem ser descritos por um conjunto de metadados (dados sobre dados) de acordo com alguma especificação formal (DOWNES, 2001).

De acordo com McGreal (2004) certas características dos objetos de aprendizagem podem ser destacadas: acessibilidade, interoperabilidade, adaptabilidade, reusabilidade, durabilidade, recuperação e avaliação. Objetos de aprendizagem podem ser construídos com base em texto, animações, apresentações, imagens, software, e devem ser descritos por um conjunto de metadados (dados sobre dados), de acordo com alguma especificação formal (DOWNES, 2001).

Segundo Gil, De La Prieta e Rodríguez (2011) ao se estudar empiricamente os objetos de aprendizagem por meio dos Repositórios de Objetos de Aprendizagem é possível identificar heurísticas fortes para cortar do sistema objetos irrelevantes e filtrá-los por parâmetros de

qualidade. Para isso, o algoritmo deve considerar tanto aspectos técnicos quanto aspectos semânticos. Em posse disso, definem-se os aspectos mais relevantes:

- quanto aos aspectos técnicos, a integralidade do objeto, ou seja, o objeto deve estar disponível para o acesso, suas informações completas e precisas, significando que sua catalogação deve estar correta e deve atender o mínimo de informações para sua recuperação;
- quanto aos aspectos semânticos estão relacionados a confiabilidade do objeto. Tal confiabilidade está relacionada ao processo de elaboração e concepção do objeto, pois falhas podem surgir nessas etapas. Tais falhas podem ser de dois tipos, sintática, quando há falta de conformidade com o padrão adotado, utilizando, por exemplo, termos em desacordo com vocabulários definidos, ou valores com erros de ortografia, etc.. E falhas semânticas, essas só podem ser detectadas pela inspeção de um especialista, pois estão relacionadas a interpretação dos termos catalogados, como exemplo, atributos ambíguos ou subjetivos.

Quando elaborado o objeto de aprendizagem ainda deve conter um conjunto de características para garantir seu ciclo de vida, dentre elas, reusabilidade, durabilidade, recuperação, entre outras. Para assegurar essas características, os objetos de aprendizagem são armazenados em repositórios digitais de conteúdo.

2.2 METADADOS

O conceito de metadados está relacionado às estruturas de informação que descrevem, sobre diversos aspectos, os próprios recursos e tal conceito é comumente relacionado como sendo os dados dos dados, ou informação sobre a informação (BARGMEYER; GILLMAN, 2011).

Uma das principais justificativas ao uso dos metadados é a facilidade em se recuperar informações de forma relevante. Além disso, auxilia na organização, facilita a interoperabilidade e integração com recursos legados, identificação digital, arquivamento e preservação. Um estudo completo sobre o assunto pode ser visto em *Understanding Metadata* (NISO, 2004).

Nos últimos anos, muitos grupos de pesquisa e organizações respeitáveis têm proposto modelos formais como normas e especificações para objetos de aprendizagem, dentre elas estão:

- Ariadne Foundation (ARIADNE, 2013)
- Aviation Industry Computer-Based Training Committee (AICC, 2013)
- Dublin Core Metadata Initiative (DCMI, 2013)
- Institute of Electrical and Electronics Engineers - Learning Technology Standards Committee (IEEE-LTSC, 2013)
- Instructional Management Systems Global Learning Consortium (IMS, 2012)
- International Organization for Standardization (ISO, 2013)
- Advanced Distributed Learning (ADL, 2013)
- Objetos de Aprendizagem Baseados em Agentes (OBAA, 2013)

Dentre os principais padrões existentes, os de maior relevância e, dessa forma mais adotados, são o Dublin core e LOM e, no contexto brasileiro, o OBAA.

O conjunto de metadados Dublin Core surgiu a partir de decisões em 1995, em um workshop patrocinado pela OCLC e NCSA. O contínuo desenvolvimento do Dublin Core e suas especificações são geridos pela *Dublin Core Metadata Initiative* (DCMI) (DUBLIN CORE, 2013). O objetivo inicial era a criação de alguns campos para descrever recursos na web, por seus autores. Com o grande crescimento dos recursos eletrônicos e a falta de descritores para catalogá-los, o padrão acabou por definir alguns elementos e regras simples para realizar tal catalogação. Inicialmente continha treze campos, estendido depois para quinze (NISO, 2004):

1. Title,
2. Creator,
3. Subject,
4. Description,
5. Publisher,
6. Contributor,
7. Date,
8. Type,
9. Format,
10. Identifier,
11. Source,
12. Language,
13. Relation,
14. Coverage,
15. Rights.

O padrão de metadados Dublin Core foi desenvolvido para ser

simples e conciso, e para descrever conteúdos baseados na web. Porém o padrão tem sido utilizado com outros tipos de documentos, que demandam certa complexidade. Existe toda uma tensão em torno do padrão, pois há os que defendem um conjunto mínimo e simples de elementos, e os adeptos de uma visão estruturada mais refinada e extensa. O conjunto de metadados Dublin Core é padronizado pelas seguintes normas, ISO Standard 15836-2009 e NISO Standard Z39.85 (NISO, 2007).

O conjunto de metadados *Learning Object Metadata* (LOM) foi criado pelo IEEE *Learning Technology Standards Committee* (LTSC) IEEE-LTSC (2013), sendo um padrão voltado ao reuso e descrição de recursos educacionais. O LOM define um conjunto mínimo de atributos para gerenciar, localizar e validar objetos educacionais. Esses atributos são agrupados em oito categorias:

- Geral,
- Ciclo de vida,
- Técnico,
- Educacional,
- Licença,
- Relações,
- Anotações e
- Classificação.

Cada categoria é um conjunto de metadados em hierarquia (NISO, 2007). Por conta dos seus atributos serem agrupados em categorias e cada categoria ser composta por um conjunto de metadados em hierarquia, a representação do padrão é feita em RDF/XML ou Atom/XML. Tal fato facilita a integração com diversos protocolos de disseminação e sistemas de buscas. Devido a suas características, flexibilidade e potencial para tratar os recursos digitais e reuso, o padrão LOM foi incorporado ao modelo de referência *Sharable Content Object Reference Model* (SCORM, 2013).

O OBAA é um padrão Brasileiro de especificação de requisitos técnicos e funcionais para a produção, edição e distribuição de conteúdos digitais interativos, permitindo que os mesmos sejam utilizados em plataformas Web, dispositivos móveis e na televisão digital. Este padrão foi desenvolvido pela Universidade Federal do Rio Grande do Sul (UFRGS) em parceria com a Universidade do Vale dos Sinos (UNISINOS) em resposta a uma chamada dos Ministérios da Educação, Comunicação e Ciência e Tecnologia e utiliza como base o padrão LOM (VICARI *et al*, 2010).

O padrão OBAA é uma extensão do padrão LOM, agregando novos elementos nas categorias *Technical* e *Educational* e acrescentados duas novas categorias *Accessibility* e *SegmentInformationTable*, para desta forma atender às necessidades brasileiras quanto ao acesso dos objetos por pessoas com deficiências e a multimídia e TV digital respectivamente.

O que se percebe é que os modelos seguem um esquema linear de descrição, como no caso do Dublin Core, ou são mais representativos e estão expressos em esquemas com categorias e ligações entre os metadados, como o LOM. Ambos são amplamente aceitos, os esquemas lineares são menos representativos semanticamente, mas são simples de se utilizar e implementar, são utilizados para catalogação de coleções de imagens, bancos de teses e dissertações, bibliotecas virtuais e assim por diante. Já os modelos em categorias são utilizados para catalogar objetos de aprendizagem e materiais didáticos, pois o modelo permite descrever toda a estrutura do objeto.

2.3 REPOSITÓRIOS DIGITAIS

Repositórios digitais de conteúdos são softwares desenvolvidos com a finalidade de armazenar e organizar recursos digitais. Dessa forma provêm mecanismos de busca e recuperação dos conteúdos (DOWNES, 2001; NASH, 2005). Os repositórios dispõem de interfaces para submissão ou catalogação de conteúdo, utilizando um ou mais padrões de metadados, interfaces de disseminação e coleta, de protocolos de comunicação e interfaces de recuperação e busca.

Segundo Lynch (2003) um repositório é um conjunto de serviços que uma instituição oferece aos membros de sua comunidade para a gestão e disseminação de materiais digitais criados pela instituição e membros da sua comunidade. Johnson (2002) afirma que um repositório digital institucional, ou seja, um repositório de informação digital que é uma parte da universidade ou de outro tipo de instituição é um arquivo digital do produto intelectual criado pelo corpo docente, pesquisadores e estudantes de uma instituição e estão acessíveis aos usuários finais dentro e fora da instituição, com pouco ou nenhum obstáculo de acesso.

Heery e Anderson (2005) afirmam que existem características que diferenciam os repositórios digitais dos demais tipos de coleções digitais, tais como: a existência de mecanismos para a disponibilização do

conteúdo pelo proprietário ou por terceiros, a existência de recursos para a gestão do conteúdo, utilizando metadados. Além disso, o repositório deve ser sustentável e confiável e, oferecer um conjunto mínimo de serviços básicos.

Os repositórios digitais têm sido abordados sob dois aspectos. O primeiro diz respeito aos repositórios voltados para o armazenamento, preservação e disseminação da produção intelectual de uma instituição (repositórios institucionais). O segundo diz respeito à produção intelectual de uma disciplina (repositórios temáticos) (COSTA; LEITE, 2006).

Conforme visto na literatura os principais conceitos envolvidos estão relacionados com a disponibilização, arquitetura de gestão, e recuperação do conteúdo, os quais podem ser descritos por um conjunto de características que um repositório deve atender. Dentre essas características, as que são mínimas para garantir um serviço confiável e sustentável são:

- interfaces de submissão e catalogação;
- uso de um padrão de Metadados para catalogar o conteúdo;
- interfaces de disseminação e coleta de objetos;
- mecanismos de recuperação e disponibilização.

Estas características devem estar atreladas a algum serviço que as gerencie o que o torna dependente do tipo de tecnologia utilizada.

Devido aos diferentes requisitos dos repositórios não é possível destacar, no atual estado da arte, apenas um modelo padrão amplamente aceito ou adotado, que poderia ser usado para orientar todas essas políticas. Tem-se dessa forma um cenário heterogêneo. O problema da heterogeneidade tem sido abordado de diversas formas. Segundo Gil, De La Prieta e Rodríguez (2011) ao se criar repositórios que são altamente sustentáveis, eles se tornam também altamente heterogêneos, pois devem lidar com diversos tipos de armazenamento, acesso aos objetos e métodos de consulta. Já para Fabre; Tarouco; e Tamusiunas (2003) heterogeneidade está relacionada de forma mais técnica, destacando diferenças tanto em implementações dos recursos dos repositórios (software) até diferenças no hardware que os sustenta.

Dessa maneira o conceito de heterogeneidade, para os sistemas de buscas, pode ser definido em cima de dois grandes pilares: o acesso aos recursos, e o padrão de metadados utilizado para descrever os objetos.

Nesse trabalho abordamos essas duas áreas no modelo de agentes, mas focando principalmente em técnicas para lidar com os diferentes padrões de metadados. Tal abordagem justifica-se pela dificuldade em se

representar a semântica do Objeto de Aprendizagem em sua recuperação e integração com outros sistemas. Li *et al* (2008) expõem que a adoção de apenas um padrão de metadados para a integração de diversos sistemas é complicada, as demandas são diferentes de um sistema para outro, e que para isso utiliza-se de uma técnica de mapeamento entre metadados chamada de *Crosswalks*.

Li *et al* (2008) também define o conceito *Crosswalks* como sendo um mapeamento de metadados de um conjunto para outro, com o mínimo de perda de informação, essência e significado. No entanto ao levantar as duas principais características do mapeamento, vemos também suas limitações:

- transformações 1-1: Existe um conjunto de metadados comum nos modelos mais populares, que são mapeados com simplicidade. Porém, por causa da complexidade de alguns padrões, não é possível garantir que não haja perda de informação, ou que ela não seja expressa satisfatoriamente;
- transformações n-1: É escolhido um conjunto central de metadados para os quais todos os outros formatos são mapeados, dessa forma reduz-se a complexidade. No entanto a precisão e a semântica são comprometidas.

Diante do exposto, para garantir a interoperabilidade dos metadados, os repositórios estão adotando um ou mais protocolos de disseminação e *Crosswalks*. Esses protocolos sugeriram de esforços para encontrar formas eficazes de replicar tanto a estrutura, quanto os objetos dos repositórios. E acabam definindo uma forma padrão de obter o acesso a esses objetos.

2.4 PROTOCOLOS DE DISSEMINAÇÃO E COLETA

A possibilidade de coletar objetos de outros repositórios e de disponibilizá-los é vital para o funcionamento de um esquema de repositórios federado. Para isso, os repositórios foram adotando formas para realizar essa tarefa, fazendo surgir protocolos de disseminação e coleta conhecidos do inglês *harvesting*. Ou seja, são protocolos que recuperam objetos em outros repositórios, preservando toda sua estrutura (metadados) e também os envia a quem quiser coletá-los.

De mesma forma em que os repositórios diferem em sua tecnologia para lidar com diferentes conceitos e tecnologias, os

protocolos surgiram em diversos contextos para sanar demandas específicas. Dentre os principais, se destacam: o *Lightweight Directory Access Protocol* (LDAP), o *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) e o *Open Archives Initiative Object Reuse and Exchange* (OAI-ORE).

O *Lightweight Directory Access Protocol* (LDAP) é um protocolo aberto que define os padrões de acesso a diretórios que operam por meio do protocolo TCP/IP, permitindo localizar, autenticar e gerenciar usuários e recursos disponíveis na rede. As operações LDAP são baseadas no modelo cliente-servidor, orientado a mensagem, desta forma o cliente constrói uma mensagem LDAP contendo um pedido e envia para o servidor. O servidor processa o pedido e envia o resultado (ou resultados) de volta para o cliente como uma série de mensagens LDAP (HOWES; SMITH; GOOD, 2003). É utilizado no repositório CESTA (CESTA, 2013), onde cada nó é um objeto, descrito por uma série de atributos (metadados) em um sistema hierárquico.

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) é um protocolo desenvolvido pela *Open Archives Initiative* (DUBLIN CORE, 2013). Segundo Li *et al* (2008) a recuperação dos metadados pode ser realizada através da definição de um protocolo comum. OAI-PMH é um protocolo proposto para troca e interoperabilidade de metadados. É composto por duas classes: um provedor de dados e um provedor de serviço. O provedor de dados é responsável por expor, de forma unificada, os metadados estruturados e organizados, em XML. Já o provedor de serviços, a partir de uma lista de formulários (repositórios), coleta os metadados regularmente, agregando-os localmente. Assim, em comparação com a recuperação distribuída, recuperação local traz melhorias nos servidores quanto à velocidade de resposta e à qualidade. O protocolo OAI-PMH fornece ainda um conjunto de metadados para a coleta e disseminação dos metadados, como uma convenção. OAI-PMH é baseado em tecnologia HTTP e seus recursos são descritos em XML.

Open Archives Initiative Object Reuse and Exchange (OAI-ORE) também é um protocolo desenvolvido pela *Open Archives Initiative* (2011). Segundo Tarrant *et al* (2009) o protocolo OAI-ORE é uma evolução do protocolo OAI-PMH. Como o protocolo OAI-PMH foi projetado para ser um método de "descoberta", não há uma interface (API) para importação, ou seja, não há uma maneira de se abstrair um objeto de forma completa, para que não haja perda de informação em uma transposição de plataformas. É nesse cenário que surge o protocolo OAI-ORE, onde há uma interface definida para importação e exportação

de objetos. Segundo as especificações do protocolo OAI-ORE (OAI-ORE, 2013) o modelo ORE foi baseado inteiramente nos recursos web, fazendo uso dos recentes avanços na área da Web-semântica. Mas a grande influencia foi o modelo RDF, onde são usadas triplas para descrever coisas. Essas triplas representam declarações sujeito-predicado-objeto, portanto o modelo é focado nos objetos e nas relações entre objetos. Sua representação é normalmente feita em RDF/XML, para representar os mapas de recursos. Porém outras podem ser usadas, como o Atom/XML que é amplamente divulgado e suportado, e também baseado em tecnologia HTTP.

Tais protocolos, conforme abordados na literatura, surgem para suprir uma demanda de interoperabilidade e intercâmbio de objetos, preservando sua estrutura e semântica. Porém, com os diferentes cenários e requisitos a adoção de um único padrão não supre as necessidades desse intercâmbio, dessa forma há uma séria de padrões, o que reforça o conceito de heterogeneidade nos repositórios.

2.5 ONTOLOGIA

A criação e definição dos conceitos e suas classificações surgem da necessidade do homem traduzir suas ideias de forma que as outras pessoas possam compreender, possibilitando assim, a organização e a fixação do conhecimento através de elementos da linguagem. Sendo que a importância de se definir os conceitos está em considerar que essas definições são pressupostos indispensáveis na argumentação e na comunicação verbal e constituem elementos necessários na construção de sistemas científicos (FACHIN, 2011). Um sistema de busca inteligente pode ser beneficiado pela criação e definição dos conceitos, de forma que a organização da relação entre conceitos pode auxiliar na desambiguação de termos, aumentando a relevância no resultado da busca.

Essa organização pode ser feita utilizando um modelo de representação do conhecimento, tal como as ontologias, que é uma especificação formal e explícita de uma contextualização, ou seja, uma lista finita de termos e os seus relacionamentos.

Gruber (1996) afirma que uma ontologia é uma especificação explícita de uma conceitualização. Em tal ontologia, definições associam nomes de entidades no universo do discurso (por exemplo,

classes, relações, funções ou outros objetos) com textos que descrevem o que os nomes procuram denotar e os axiomas formais que restringem a interpretação e o uso desses termos.

Fachin (2011) diz que se trata da descrição e representação de uma área específica de conhecimento, atendendo a questões específicas da necessidade dos usuários desta área, visando à recuperação eficiente de informações, baseada em semântica e no tratamento do conhecimento.

Ontologias representam um conhecimento que é comum aos usuários de uma determinada área, elas modelam o conhecimento do especialista. Podem ser representadas em sistemas computacionais sobre as formas *Resource Description Framework* (RDF) e *Web Ontology Language* (OWL) e utilizadas para enriquecer a busca de conceitos. Como as ontologias acabam modelando uma área de conhecimento de forma comum a um grupo de usuários, ela traz benefícios, como a definição dos termos de busca, desambiguando, por exemplo, encontrando termos correlatos e sinônimos, enriquecendo a busca. Também podem ser usadas para categorizar objetos (LEE; TSAI; WANG, 2008).

2.6 SISTEMAS MULTIAGENTES

Os sistemas de busca e recuperação de informação podem se beneficiar dos sistemas multiagentes, pois estes são por definição adaptáveis e distribuídos, podendo atuar em conjunto para atingir um objetivo comum, fazendo uso de sua organização na sua arquitetura de sistema para realizar buscas em repositórios federados.

Os agentes podem se comunicar e deliberar (tomar decisões e fazer inferências) sobre qual a melhor forma de se obter a informação desejada, podem também receber estímulos do usuário (agente perfil) e se readaptar ao novo contexto. A arquitetura permite isso, de forma clara e objetiva, e o modelo permite a replicação da técnica de acordo com as características do cenário de uso. O conceito de agentes é empregado em diversas áreas e sob diversas formas. Não existe apenas uma definição, pois o conceito é empregado sobre diversos contextos e, para cada qual, devesse explicitar seu significado. Mesmo quando utilizado em um contexto bem definido, como agentes de software ou agentes inteligentes, ainda assim a palavra "agente" representa um termo genérico.

Dentre os conceitos de sistemas multiagentes existentes na literatura, optou-se pela definição de Wooldridge (2002), que afirma que um sistema multiagente é um sistema composto por vários elementos (agentes) que interagem para realizar alguma ação computacional. Um agente é um sistema computacional capaz de agir de forma autônoma em um ambiente para atingir algum objetivo.

O relacionamento dos agentes com o ambiente ocorre sobre:

- Sua percepção do ambiente, qual o seu estado atual, seu último estado e quais mudanças ocorreram.
- Que decisões e comportamentos adotar para, a partir dessa percepção, atingir os objetivos.
- Que ações devem ser tomadas para que se atinjam tais objetivos.

Quanto ao comportamento dos Agentes, Woodridge (2002) os divide em três:

1. Reativo: para os ambientes dinâmicos, onde não se pode executar sempre uma mesma ação e o programa devem levar em conta possibilidade de falha, um agente apresenta o comportamento reativo se ele mantém uma interação permanente com seu ambiente, e responde às mudanças que nele ocorrem. Portanto o comportamento de reatividade está relacionado a capacidade do agente perceber mudanças no ambiente (estímulo) e executar uma ação a essa mudança.
2. Pro-atividade: o conceito de pro-atividade complementa, de certa forma, o comportamento reativo. O comportamento reativo pode ser visto como uma base de regras, onde cada estímulo desencadeia uma ou mais regras. A pro-atividade é definida então como formas de se atingir os objetivos não apenas impulsionado pelos estímulos, mas de forma a prevê-los ou direciona-los.
3. Sociabilidade: o conceito de sociabilidade está fortemente ligado às capacidades de interação entre os agentes. Como citado em sua definição, os agentes são elementos autônomos e devem interagir para a obtenção de seus objetivos, dessa análise surge a capacidade de se comunicar. Pela amplitude desse comportamento, Wooldridge (2002) ainda aborda três subdivisões, como habilidades do sistema:
 - Cooperação: Os agentes interagem entre si, compartilhando recursos e informações, seja para se atingir um objetivo, ou para se obter o melhor resultado.

- Coordenação: Habilidade de gerir as interdependências entre as atividades, coordenando os recursos e interações ente os agentes.
- Negociação: Capacidade de se chegar a acordos sobre assuntos de interesse comum.

Por tais características, autonomia, pro-atividade e adaptabilidade, os sistemas multiagentes tem sido utilizados em conjunto com as técnicas existentes de busca e recuperação de objetos de aprendizagem, trazendo melhorias e resultados positivos. Exemplos da aplicação de sistemas multiagentes na busca e recuperação da informação podem ser encontrados em Gil, De La Prieta e Rodríguez (2008) e Vian *et al* (2011).

2.7 SISTEMAS DE RECOMENDAÇÃO

Os sistemas de recomendação surgiram com a finalidade de minimizar o problema da sobrecarga de informações. Os sistemas de recomendação têm por objetivo indicar ou receber indicações de produtos e serviços, num processo social (RESNICK; VARIAN, 1997).

Schafer, Konstan e Riedl (2001) define a estrutura de um sistema de recomendação contendo quatro etapas:

- Identificação do usuário: apesar de ser definida como opcional é responsável por gerar perfis de usuários e grupos de usuários.
- Coleta de dados: etapa onde os dados sobre o usuário e produtos são coletados. Pode ser realizada de três formas: implícita, explícita e híbrida.
- Estratégia de recomendação: descreve a técnica (algoritmo) utilizada para relacionar usuários aos produtos e serviços. É formada por três métodos básicos: filtragem colaborativa, relacionamento por conteúdo ou uma abordagem mista.
- Visualização das recomendações: descreve a forma como os produtos e serviços devem ser apresentados aos usuários. Nessa etapa também é feita a coleta de dados, de forma a aprimorar o próprio sistema.

O uso de técnicas de recomendação de conteúdo traz melhores resultados à busca de objetos em repositórios, pois ao fazer uso de recomendação reduz-se o problema da sobrecarga de informação. Isso se justifica, pois os modelos de recomendação trazem melhorias como

relacionar conceitos aos usuários, ordenando assim o conteúdo conforme o perfil do utilizador do sistema, complementando os resultados com objetos visando o nivelamento de conhecimento sobre um conceito, identificando grupos de usuários que prestam grandes contribuições, fazendo uma rede de confiança, e focando em objetos que estão sendo mais acessados listando-os como mais atualizados (PRIMO; LOH, 2006).

2.8 TRABALHOS CORRELATOS

Gil, De La Prieta e Rodríguez (2011), partem do conceito de OA, justificando que seu desenvolvimento e aprimoramento vêm para facilitar a reusabilidade e interoperabilidade dos seus recursos. Os OA são armazenados em repositórios digitais, que devido as diferentes formas de recursos, metadados e empacotamento dos OA devem ser capazes de lidar com uma grande variedade de tecnologias e padrões. Isso os torna também heterogêneos, pois para tratar tais aspectos os repositórios precisam prover uma série de funcionalidades técnicas, protocolos de comunicação, notação de metadados e mecanismos de recuperação.

Partindo desse cenário. Gil, De La Prieta e Rodriguez (2011) propõe uma arquitetura inteligente para recuperação de conteúdos educacionais em ambientes heterogêneos, chamada de *Architecture for Intelligent Recovery of Educational content in Heterogeneous Environments* (AIREH). Tal arquitetura foi desenvolvida sobre o paradigma de sistemas multiagentes, organizados para evitar os aspectos heterogêneos dos objetos e dos repositórios, realizando buscas pelos próprios sistemas de buscas dos repositórios por ele indexado.

Para recuperar os objetos de forma significativa os autores fizeram um levantamento de características relevantes e elencaram parâmetros de qualidade para filtrar e classificar os objetos. Chegaram então, a três características dos objetos fundamentais para uma boa recuperação:

- Completude, o objeto deve trazer, em seus metadados, pelo menos, as informações mais relevantes.
- Confiabilidade, a informação contida nos metadados deve ser válida.
- Precisão da classificação, essa qualidade está relacionada

ao algoritmo de busca e recuperação dos objetos, onde o mesmo deve ser mais eficaz ao levar em conta a semântica do objeto.

Para atender tais características, a arquitetura proposta filtra os objetos retornados, criando um índice único, removendo objetos duplicados e não reutilizáveis. A partir desse índice o sistema constrói um catálogo, com base em duas variáveis. A primeira analisa os objetos, levando em conta aspectos como tamanho, completude dos metadados, palavras-chaves, etc. Realizando uma estimativa de qualidade do objeto. Já a segunda variável faz uma análise por recomendação colaborativa dos utilizadores do sistema, coletando informações como notas prévias dos usuários e revisões. O sistema permite também a retroalimentação dessas características colaborativas, pelo retorno do usuário ao sistema, avaliando o catálogo que a ele foi informado.

Para avaliação do sistema, foi realizada uma bateria de testes com 53 consultas em dois repositórios reais. Os autores concluíram que a inovação trazida pela arquitetura do sistema em sua organização resolve, de forma transparente, o problema da busca em repositórios federados e heterogêneos, facilitando a interação do usuário com o sistema. Concluem também que pelo fato da arquitetura baseada em sistemas multiagentes serem adaptativas, eles se tornam ideias para lidar com a heterogeneidade de repositórios.

Vian *et al* (2011) afirma que dado o grande investimento em produção de conteúdos digitais, a sua reutilização é bem vinda, e por tal fato, vem se disseminando o conceito de Objeto de Aprendizagem. Por ser promissor, o conceito de OA, muitas ideias e especificações surgiram para guiar os aspectos políticos e técnicos de produção e armazenamento dos OA. Como não é possível destacar um único padrão para guiar todas essas políticas, tem-se um cenário heterogêneo, onde há uma gama de padrões técnicos e políticos tanto para a produção quanto para o armazenamento e acesso aos OA.

Assim, os detalhes técnicos dos repositórios, somados com as diversas especificações dos OA dificultam o acesso aos objetos. Dessa forma o utilizador deve usar o sistema de recuperação do próprio repositório. Surge então a necessidade de desenvolver um sistema que fomente a busca e reutilização dos Objetos de Aprendizagem. Os autores propõe então, um modelo de sistema multiagente projetado para buscar e recuperar OA em diferentes repositórios, com heterogeneidade de acesso e de metadados.

O modelo é composto por uma interface web com o usuário, dois tipos de agentes, uma base de conhecimento e uma base de mapeamento

de metadados. Existe um papel de agente buscador, responsável por receber solicitações dos usuários e repassa-las aos agentes indexadores. Cada agente indexador é capaz de indexar um repositório, conhecendo seu mecanismo de disseminação de objetos e para lidar com a heterogeneidade de metadados, os agentes são capazes de realizar consultas a uma base de mapeamento de metadados, transformando os padrões LOM e Dublin Core em LOM. Os agentes indexadores são ainda capazes de consultar sinônimos dos termos submetidos a eles em uma base de conhecimento.

Para dar validade ao sistema, os autores realizaram testes em dois repositórios CESTA (CESTA, 2013) e LUME (LUME, 2013). Ao sistema foram submetidos 13 termos e coletados os resultados considerados relevantes pelos autores, esses resultados foram comparados os resultados produzidos com a ferramenta de busca de cada repositório indexado pelo sistema.

Os autores concluem que o modelo baseado em sistemas multiagentes permite, de forma simples, buscar objetos em repositórios federados e heterogêneos. A unificação da busca de diversos repositórios estimula o reuso dos objetos e o uso de bases de conhecimento para expandir os termos da busca provê uma melhor cobertura dos resultados.

D'Agostine (2009) questiona o uso de uma única representação de conhecimento, expresso em ontologias, pois as pessoas, em geral, tem diferentes visões sobre um determinado conceito, impedindo assim uma visão unificada de um conhecimento que satisfaça a todos. O autor destaca que uma solução para isso é permitir que pontos de vista diferentes possam existir e serem mapeados em ontologias subjacentes, gerando um conhecimento mais personalizado. Entretanto é um papel difícil para o usuário tentar se expressar de acordo com o formalismo exigido pelas ontologias.

Os autores se questionam então, em como capturar informações do usuário e emprega-las, juntamente com uma determinada ontologia, para melhorar a precisão, a cobertura e a classificação dos resultados de busca, de forma personalizada?

A resposta vem em forma de um sistema, Praestro, o qual traça um perfil de consultas do usuário, sendo possível definir, qualificar e desambiguar termos para pesquisa e filtrar os resultados, garantindo um grau semântico à resposta dada pelo sistema. O sistema inicia-se utilizando uma base de conhecimentos, e a partir da primeira interação do usuário, modifica essas bases, atribuindo pesos a determinadas relações, priorizando-as. Como exemplo, uma pesquisa ao termo "São Paulo" irá resultar uma série de resultados, como: à cidade de São Paulo,

time de futebol São Paulo, etc. Ao selecionar uma dessas opções, o sistema assume essa relação de São Paulo ao significado selecionado: cidade. Dessa forma é atribuído um peso a essa relação São Paulo - Cidade, esse peso irá direcionar os resultados da busca, priorizando as relações com maior peso.

Em se tratando de um sistema iterativo e personalizado, os autores relataram uma dificuldade de viabilizar experimentos, o que impossibilitou a validação do mesmo. Porém foi evidenciada uma crescente melhoria na relevância dos resultados, atribuída pelo sistema.

A análise destes três trabalhos evidencia a necessidade de um modelo de sistema de busca que seja capaz de realizar buscas em repositórios heterogêneos, de forma personalizada e utilizando aspectos semânticos. Por estudar empiricamente os sistemas, foram elencadas as seguintes características como fundamentais:

Realizar buscas em repositórios heterogêneos, de forma transparente ao usuário, como sugere Gil, De La Prieta e Rodríguez (2011), porém utilizando um modelo que trate a heterogeneidade de acesso e de metadados com robustez, como em Vian *et al* (2011) dessa forma o sistema tem o controle sobre os algoritmos que indexam e recuperam os objetos.

O sistema deve enriquecer a busca, através de uma base de conhecimento, expandindo e desambiguando os termos, como em Vian *et al* (2011), mas levando em conta os aspectos pessoais de cada utilizador, por meio de um perfil e esse perfil deve estar em constante aprendizagem, como sugere D'Agostine (2009).

3. METODOLOGIA

A pesquisa desenvolvida teve como objetivo geral propor um modelo de sistema inteligente de busca, recuperação e recomendação para Objetos de Aprendizagem armazenados em diferentes repositórios com heterogeneidade de metadados. Desta forma caracteriza-se como uma pesquisa de natureza aplicada, pois busca originar um conhecimento para auxiliar na resolução de um problema específico. Este tipo de pesquisa destina-se "a aplicar leis, teorias e modelos na solução de problemas que exigem ações e/ou diagnósticos de uma realidade" (SALOMON, 2008, p. 158).

Em relação aos seus objetivos caracteriza-se como exploratória e descritiva, pois busca "definir melhor o problema, proporcionar as chamadas instituições de solução, descrever comportamentos de fenômenos, definir e classificar fatos e variáveis" (Salomon, 2008, p.158). Quanto aos seus procedimentos esta é uma pesquisa bibliográfica e documental. Esta Pesquisa é bibliográfica, pois para atender o seu primeiro objetivo específico, que é definir uma arquitetura que possibilite a utilização de parâmetros de perfil do usuário como critério de recomendação, foi utilizando como base materiais já elaborados, constituídos por livros e trabalhos científicos (artigos, teses e dissertações). É documental, pois para atender ao segundo objetivo, que é elaborar um mapeamento de correspondência dos campos dos principais padrões de metadados para objetos de aprendizagem, foram analisados e comparados os principais padrões de metadados relacionados aos objetos de aprendizagem.

Para a execução da pesquisa, bem como a compilação dos dados que gerarão os subprodutos e resultados finais foram estabelecidas as etapas:

1. Identificar e analisar as principais técnicas de recuperação da informação, com a possibilidade de utilizar parâmetros do perfil do usuário;
2. Identificar os principais protocolos de comunicação de repositórios de objetos de aprendizagem.
3. Identificar e analisar semanticamente os principais padrões de metadados utilizados para descrever os objetos de aprendizagem.
4. Definir as características do sistema.
5. Desenvolver um modelo conceitual, focado nas ideias de adaptabilidade e distributividade.
6. Elaborar os planos dos agentes e suas competências, segundo a metodologia de descrição adotada.

7. Construir um protótipo para dar validade ao modelo.
8. Definir métricas de análise e realizar comparações.
9. Analisar os resultados.

A pesquisa iniciou-se com o levantamento da literatura nas principais bases relacionadas ao tema, o que justifica-se por "permitir uma compreensão mais ampla da produção nacional e internacional, mediante a identificação de obra efetivamente existentes em determinados acertos, bases ou bancos de dados" (PEREIRA; BUFREM, 2002, p.198). Para tanto foram levantados e abordados os temas que fundamentam o desenvolvimento do modelo proposto: sistemas de busca e recomendação existentes, passíveis de utilização em repositórios. Desta forma, foram identificados os protocolos de comunicação e padrões de metadados mais utilizados e ontologias existentes.

Foram identificadas as estruturas dos sistemas de recomendação e dentre elas foi escolhido a taxonomia definida por Schafer, Konstan e Riedl (2001). Composta por: identificação do usuário, coleta de informações sobre o usuário de forma implícita ou explícita (do perfil, sites, comunidades, por indicação de palavras chaves), estratégias de recomendação e visualização das recomendações. Esta taxonomia garante a flexibilidade necessária para a aplicação no modelo aqui proposto, desta forma definidos os quatro pontos da estrutura de recomendação.

Tais pontos são implementados em uma base de perfis, acessível ao Agente de Perfil. O qual faz uso dessas informações, construídas através da interação do usuário com o sistema. No protótipo a base foi construída sobre um modelo relacional.

O modelo conceitual foi idealizado para indexar qualquer repositório, utilizando um agente indexador especializado. Foi implementado no protótipo, um agente especializado no protocolo de comunicação OAI-PMH. Desta forma é prevista a heterogeneidade de metadados, interpretando qualquer padrão. Isso ocorre, pois o agente indexador é especializado em um determinado padrão, prevendo o uso dos mapeamentos expostos por Li *et al* (2008), tanto mapeamentos um-para-um, quanto n-para-um. Para isso foi realizado um mapeamento entre os principais padrões de metadados, apresentados no quadro 1.

Quadro 1: Mapeamento semântico de metadados

Nome/Padrão	Dublin Core (Qualificado)	LOM	OBAA
Título	dc.title	lom.general.Title	obaa.general.title
Autor	dc.creator dspace:dc.contributor.author	lom.lifeCycle.contribute.entity e lom.lifeCycle.contribute.role='author'	obaa.lifeCycle.contribute.entity e obaa.lifeCycle.contribute.role = 'author'
Descrição ou Resumo	dc.description ou dc.description.abstract	lom.general.description	obaa.general.description
Palavras-chave	dc.subject	lom.general.keyword	obaa.general.keyword
Status	-	lom.lifecycle.status	obaa.lifecycle.status
Data	dc.date dspace: dc.date.issued	lom.lifecycle.date	obaa.lifecycle.date
Versão	dc.relation.hasVersion dspace: dc.description.version	lom.lifecycle.version	obaa.lifecycle.version
Link ou Arquivo (URI, DOI, Handle)	dc.identifier ou dc.identifier.uri ou dc.source	lom.general.identifier.catalog='URI' (ISBN, ARIADNE) e lom.general.identifier.entry	obaa.general.identifier.catalog='...' e obaa.general.identifier.entry
Contexto Educatonal	-	lom.educational.description	obaa.educational.description
Área de conhecimento	-	lom.classification.keyword	obaa.classification.keyword
Tipo do objeto	dc.type	lom.educational.learningResourceType	obaa.educational.learningResourceType
Público alvo	-	lom.educational.intendedEndUserRole	obaa.educational.intendedEndUserRole
Interação	-	-	obaa.educational.interaction
Estratégia didática	-	-	obaa.educational.didacticstrategy
Relações	dc.relation.*	lom.relation.*	obaa.relation.*

Fonte: adaptado pelo autor de LOM, Dublin Core e OBAA

O mapeamento entre metadados dos padrões Dublin Core, LOM e OBAA, foi realizado sobre os aspectos semânticos dos padrões, informados por suas descrições formais. Levaram-se em conta os principais campos de metadados que identificam o objeto e os principais campos que descrevem a interação educacional. Nesse quadro foram analisados também os aspectos práticos de implementação no repositório Dspace (DSPACE, 2013).

As características do sistema foram definidas com base em uma revisão bibliográfica, resultando em uma análise sobre trabalhos correlatos. Foram analisadas tantas as características conceituais quanto técnicas que tornam o sistema capaz de lidar com a heterogeneidade de acesso e de descrição dos objetos, foram evidenciadas também os aspectos positivos em se incorporar características do perfil do usuário para ordenar e recomendar os resultados da busca.

O modelo conceitual foi elaborado contemplando as características definidas, dessa forma, tornando-o hábil para lidar com a heterogeneidade de metadados e de acesso aos objetos de aprendizagem, incorporando características do perfil do usuário e utilizando bases de conhecimento para enriquecer os termos para busca.

O modelo é então definido conceitualmente por sua descrição geral, evidenciando suas características, um fluxograma de informação, o qual define o comportamento desejado e uma sequência de diagramas do modelo de agentes, os quais definem seus objetivos, serviços, capacidades e comunicação.

A materialização do modelo, por meio da implementação, se deu com o framework JADE, no protótipo. A comunicação entre os agentes é estabelecida por troca de mensagens usando o padrão FIPA *Agent Communication Language* (FIPA-ACL). A comunicação entre o agente perfil e as bases de perfis e de conhecimento é realizada por meio de consultas (**query**).

A base de conhecimento é representada por ontologias. Essas ontologias são descritas conforme o modelo proposto e possuem dois propósitos: a expansão dos termos da busca, para garantir uma boa cobertura, e a filtragem dos resultados, evitando objetos fora de contexto. A ontologia de expansão de termos foi construída de forma que, sobre os termos, seja possível obter sinônimos e sua definição para desambiguação.

Para avaliar o modelo proposto foi implementado um protótipo com uma interface desenvolvida na linguagem PHP, com a possibilidade de identificação do usuário. O sistema é requisitado via protocolo, através da chamada ao método *Search* e com um vetor de texto contendo

informações de como identificar o usuário, no caso do protótipo o número de identificação dele no sistema, e os termos a serem pesquisados. Como resposta a chamada, o sistema web recebe também um vetor de texto, com informações dos objetos retornados. Foi utilizada a ontologia que represente o conjunto de termos do DECS (Descritores em Ciências da Saúde). Essa ontologia é utilizada pelos agentes para expandir as consultas solicitadas pelo usuário da aplicação. Para que isso seja possível foi utilizada uma base de conhecimento. Para tratar os diferentes mecanismos de acesso aos objetos e seus descritores, os agentes indexadores foram adaptados e estendidos para trabalhar com a biblioteca JDOM, além de implementar os protocolos OAI-PMH, OAI-ORE e de forma específica, o protocolo *ad-hoc* do Dspace.

Para a coleta de informações sobre o sistema multiagente, a fim de analisa-lo e valida-lo, foram definidas métricas tanto para a coleta quanto para análise dos dados. Nesta análise é verificada a pertinência de um dado objeto com os termos definidos para a consulta, bem como a relevância dos objetos retornados. As considerações sobre a relevância dos objetos ficou a cargo do julgamento do especialista. Foram elaborados tabelas e quadros comparativos dos dados obtidos pela pesquisa no SMA e em cada repositório por ele indexado.

Para utilização do sistema foi definido uma amostra não probabilística, com uma amostragem por julgamento, feita com especialistas da área. A amostragem não probabilística é utilizada em casos onde a seleção dos elementos da população para compor a amostra depende, ao menos em parte, do julgamento do pesquisador ou do entrevistador no campo. E mostra-se adequada ao problema em questão, devido às limitações em relação quantidade de especialistas com as competências necessárias para a validação do modelo (BARBETTA, 2007).

Foram elaborados três tipos de testes com o intuito de avaliar a busca de forma geral, a expansão dos termos e seus resultados e a busca específica utilizando o perfil. Ao avaliar a busca geral é analisada a indexação dos diversos repositórios. São avaliados os ganhos em se expandir os termos e a ordem em que os resultados são expostos, relevância. Já a avaliação da busca específica almeja validar as especificidades do perfil do utilizador do sistema, os filtros de metadados para o cálculo da relevância e os aspectos semânticos da busca.

Para analisar e validar a busca geral foi realizado teste com especialistas de domínio, definindo-se as seguintes métricas:

- Repositórios indexados: ARES UnA-SUS por sua

especificidade na área de saúde pública, podendo responder ao busca específica e LUME pois trata-se de um repositório multidisciplinar e com uma vasta gama de objetos.

- Termos para pesquisa: foi utilizada uma lista de dez termos, todos na área médica por aderirem à área de pesquisa do especialista.

Os testes foram realizados submetendo-se os termos aos repositórios, e coletando-se os dez primeiros objetos de cada. Em seguida os mesmos termos foram submetidos ao sistema multiagente em estudo, coletando-se seus resultados. Aos especialistas foram mostradas as listagens dos objetos, os dez primeiros, na ordem em que foram apresentados pelos repositórios juntamente com a listagem gerada pelo SMA para análise. Os especialistas também respondem, para cada objeto, se são pertinentes aos termos de busca ou se são irrelevantes à pesquisa. Para auxiliá-los foi desenvolvido um quadro de comparação da posição dos objetos nos repositórios, no SMA ordenado para cada repositório, para comparação direta, e no SMA de forma geral, com os resultados dos dois repositórios.

Para validar os recursos do perfil do usuário, foi utilizado um perfil com especialidade na área médica, contendo um repositório de busca preferido e um conjunto de valores de metadados pré-selecionado, simulando um perfil de busca.

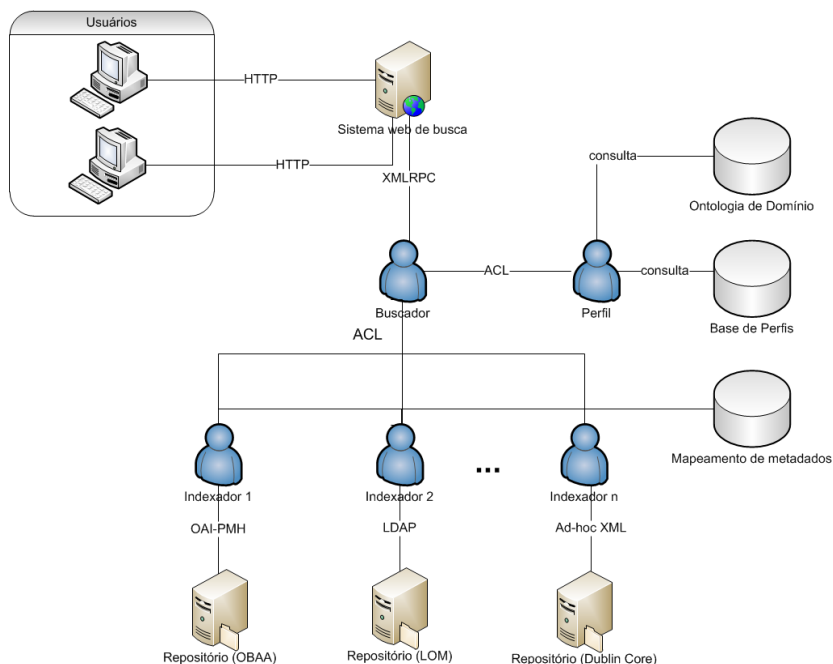
Os testes foram realizados pela interação de um especialista, ciente das informações contidas no perfil, com o sistema multiagente. Onde o mesmo elaborou uma lista de termos e as submeteu ao SMA, primeiramente com o perfil desativado, e depois com o perfil ativado. Ambas as pesquisas foram coletadas e em posse dos resultados, o especialista, com base em seu julgamento, avaliou os aspectos do perfil.

Ao final do experimento, tem-se uma lista de objetos recuperados para cada repositório e para o SMA, com a posição de cada objeto em relação ao SMA, destacando os objetos fora de contexto ou irrelevantes para pesquisa, bem como os objetos que foram recuperados pelo uso de algum sinônimo, neste caso que estavam presente no SMA e não constavam nos resultados dos repositórios. Tem-se também uma listagem com o total de objetos recuperado por cada base indexada, comprovando uma melhor cobertura de busca para os termos expandidos. E um comparativo de busca sem os aspectos do perfil do usuário e com o perfil.

4. O MODELO PROPOSTO

O objetivo deste modelo é desenvolver um sistema inteligente de busca, recuperação e recomendação para Objetos de Aprendizagem armazenados em diferentes repositórios com heterogeneidade de acesso e de metadados, por meio do apoio de agentes de software. O modelo de sistema multiagente proposto foi desenvolvido para ser capaz de indexar, classificar e recuperar objetos de aprendizagem em diferentes repositórios garantindo além de uma boa cobertura e recuperação por área de conhecimento, priorizar os resultados relevantes.

Figura 1: Arquitetura do sistema multiagente

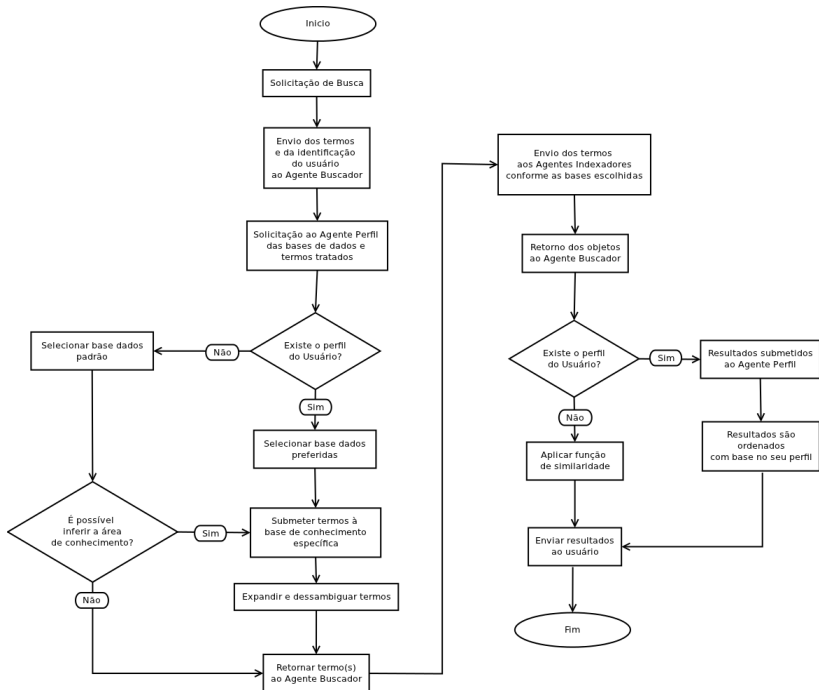


De acordo com a figura 1, os principais componentes são: o sistema multiagente, o sistema *web* de busca, o conjunto de repositórios de objetos de aprendizagem, e as bases de perfis de usuários e de conhecimento (expresso por ontologias). O sistema *web* é responsável pela interface do sistema multiagente com o usuário, recebendo informações de seu perfil, e termos para busca. Os repositórios podem ser totalmente diferentes, tanto em seu conteúdo quanto na tecnologia

em que foi desenvolvido, e para cada tipo de repositório, tem-se um Agente Indexador adaptado para lidar com um ambiente heterogêneo. As bases de conhecimento são acessadas pelo Agente de Perfil, e são responsáveis por expandir os termos para busca e filtrar os objetos retornados, já a base de perfil é responsável por armazenar as informações do usuário, sendo consultada pelo Agente Perfil.

Ao seguir o fluxo do utilizador do sistema, ou seja, o fluxo da informação tem-se: o Agente Buscador recebe um pedido de busca do sistema *web*. Esse pedido contém os termos da busca e, se disponível no sistema, informações de identificação do usuário; em posse dessas informações o Agente Buscador solicita informações ao Agente Perfil, que responde com as bases mais utilizadas, e os termos tratados (expandidos e detalhados); o Agente Buscador seleciona as bases (baseado no perfil) e solicita aos Agentes Indexadores dessas bases os objetos que atendam aos termos; em posse das respostas dos Indexadores, o Agente Buscador as submete ao Agente Perfil que ordena os objetos por similaridade textual entre os termos da pesquisa e os metadados do objeto, considerando também aspectos do perfil do usuário. Na resposta o sistema pode retroalimentar o banco de perfis, ajudando na construção do mesmo. Em detalhes na figura 2.

Figura 2: Fluxo da informação



De acordo com a figura 2, o fluxo de informação ocorre da seguinte maneira: O usuário acessa o sistema, podendo se identificar. Ao acessar o sistema, ele pode enviar uma requisição de busca, por meio de termos, palavras-chave. Dessa forma ele faz uma solicitação de busca. O sistema de busca web recebe essa solicitação e a repassa ao sistema multiagente via chamada remota de processo, contendo um vetor com os termos e a identificação do usuário, se realizada. O Agente Buscador recebe essa solicitação e envia uma mensagem contendo os termos e a identificação do usuário ao Agente Perfil, solicitando as bases de dados para pesquisa e os termos para busca tratados semanticamente.

Em posse dessas informações, o Agente Perfil envia mensagem ao banco de perfis solicitando se o perfil existe e suas informações. Existindo, ele pergunta sobre as bases de dados preferidas e solicita a base de conhecimento específica, submetendo a ela os termos da busca para serem expandidos e detalhados. Caso contrário, um conjunto de base de dados padrão é selecionado e ele tenta inferir a área de conhecimento dos termos, caso seja possível, os termos são submetidos a uma base de conhecimento para expansão e detalhamento e seguem o

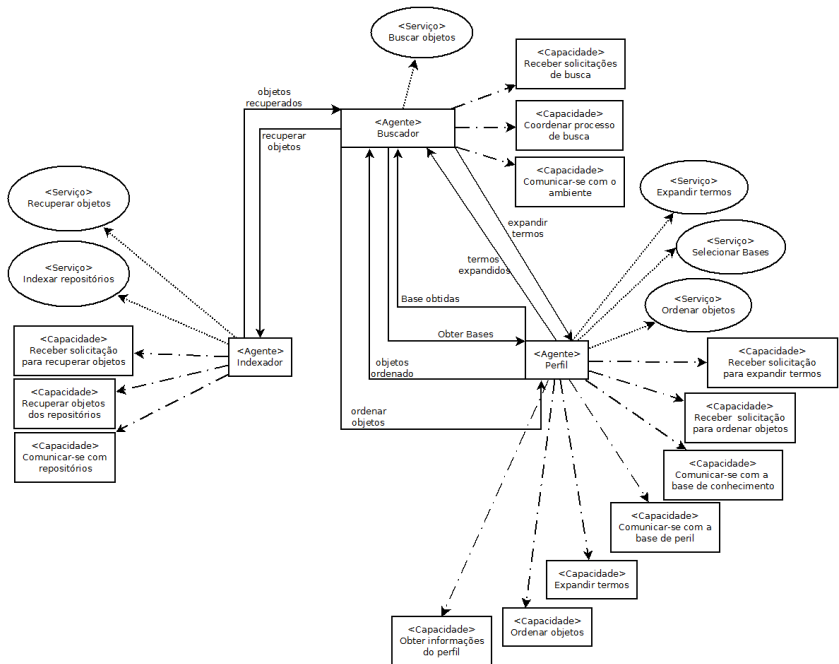
fluxo padrão, caso não seja, os termos não sofrem modificações.

Definidas as bases de dados e os termos para busca, o Agente Perfil os informa ao Agente Buscador que, conforme as bases selecionadas, aciona os Agentes Indexadores, enviando a eles mensagem com os termos de busca. Tais agentes vão realizar a busca nos repositórios por eles indexados, com base nos campos de resumo, título e palavras-chave, conforme o mapeamento do quadro 1. Realizada a busca, os agentes enviam mensagem ao Agente Buscador, contendo os objetos recuperados por eles.

O Agente Buscador recebe a lista de objetos e a envia ao Agente Perfil, verifica se há informações sobre o usuário para ordenar os objetos. Existindo, os objetos são ordenados com base no perfil do usuário e são enviados para visualização. Caso o perfil do usuário não seja localizado, os objetos são submetidos a uma função de similaridade, para que sejam evitados objetos fora de contexto e duplicados.

Conforme as descrições da arquitetura de agentes, do fluxo de informações e dos agentes, são definidos os serviços oferecidos pelos agentes, especificados em páginas amarelas, e as capacidades à eles atribuídas, com base no fluxo da informação e a interação entre os agentes, por troca de mensagens. Os serviços, capacidades e trocas de mensagens entre os agentes estão em detalhes na figura 3.

Figura 3: Diagrama de agentes.



Estão representados, na figura 3, os agentes identificados no modelo por retângulos marcados com a palavra "Agente", a interação (troca de mensagem) entre os agentes está representada por uma seta contínua e marcada por um texto que informa o tipo de mensagem. O Agente Buscador interage com os agentes perfil e indexadores e eles interagem somente com o Agente Buscador, o que caracteriza a coordenação das etapas de busca. A primeira etapa de troca de mensagens entre o Buscador e o Perfil visa obter as bases de pesquisa e os termos expandidos com a ontologia, a segunda interação é realizada para ordenar os objetos resultantes da recuperação dos indexadores. A interação entre o Agente Buscador e os agentes Indexadores, representada na figura por um Agente Indexador, é realizada por uma troca de mensagens.

Os serviços são representados por elipses marcadas com o texto "Serviço". Os serviços são competências que definem os objetivos gerais dos agentes. Sobre esses serviços, são realizadas as solicitações. Os serviços são registrados, no sistema multiagente, em uma lista de serviços para consulta (páginas amarelas), assim um agente que precisa de um determinado serviço pode consultá-las e obter uma listagem dos

agentes que oferecem tal tarefa. No diagrama, os serviços estão relacionados com os agentes por meio de uma seta pontilhada.

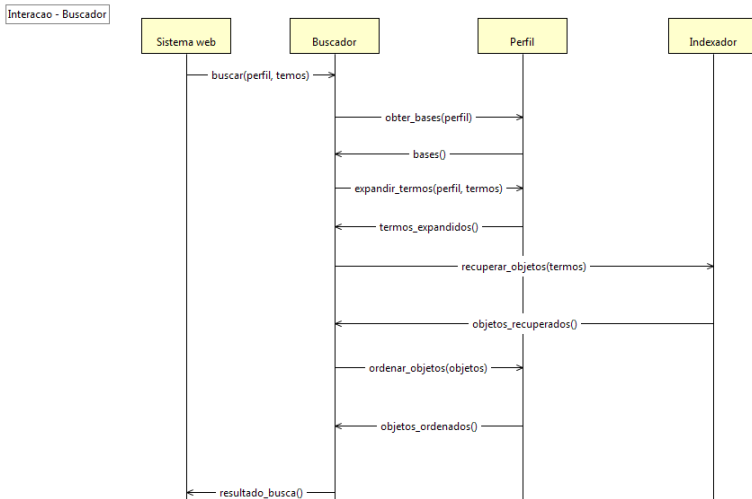
As capacidades do agente estão expressas por retângulos, marcados com "Capacidade" e os agentes que se relacionam com elas estão marcados por setas com tracejado e ponto. As capacidades representam as etapas para atingir os objetivos definidos na descrição da arquitetura geral.

4.1 AGENTE BUSCADOR

O Agente Buscador age como um coordenador de todo o processo, comunicando-se com o meio externo, o sistema web de busca, com os Agentes Indexadores e com o Agente Perfil. Todo o fluxo de informação passa por ele, a comunicação, então, torna-se importante para o agente, que é capaz de comunicar-se via *Agent Communication Language* (ACL) com os agentes e por chamada remota de processo com o sistema web de busca. Dessa forma o Agente Buscador é capaz de comunicar-se com o sistema *web* de busca, recebendo solicitações de busca e enviando os resultados; coordenar as etapas da busca, envio de mensagens aos agentes, manter o fluxo de informação e definir quais agentes indexadores acionar.

A representação da interação entre o agente buscador, os demais agentes e o sistema *web* de busca pode ser vista no diagrama de interação do Agente Buscador, na figura 4.

Figura 4: Diagrama de interação do Agente Buscador.



O diagrama de interação, representado na figura 4, contém o sistema *web* de busca e os três agentes: buscador, perfil e indexador representados por retângulos com suas linhas de vida (tempo). As setas ligando uma linha de vida a outra representa um envio de mensagem, onde o texto intitula o tipo da mensagem e se há a necessidade de informar algum parâmetro, ele é representado entre parênteses.

A busca se inicia por meio da interação do usuário com o sistema *web* e, a partir da solicitação de busca, com o sistema multiagente. Essa interação é estabelecida pela comunicação entre o sistema *web* e o Agente Buscador por meio de uma solicitação de processo de busca, onde o sistema *web* informa ao agente os termos para a busca e as informações sobre a identificação do usuário no banco de perfis. Tratando-se uma chamada remota de processo, o sistema *web* envia como parâmetros, juntamente com a requisição, os termos e a identificação do usuário, esses parâmetros são expressos por um vetor de dados.

Em posse dos dados, os quais vieram por parâmetros do sistema *web*, o agente precisa selecionar em quais bases pesquisar e se é possível expandir e detalhar os termos. Para isso, ele envia mensagem ao Agente Perfil, contendo os termos e a identificação do usuário, solicitando que o agente trate os termos, e diga em quais bases de dados ele deve pesquisar. Como resposta o agente recebe, com base no perfil do usuário identificado, os termos e as bases para pesquisa. Os termos

podem estar tratados ou não, se o Agente Perfil conseguiu identificar o usuário na base de perfis ou inferiu a área de conhecimento da busca, os termos estarão expandidos e detalhados, caso contrário estarão como foram enviados pelo usuário, sem detalhamento. De forma semelhante é feita a seleção das bases de dados para consulta, se estiverem presentes como informação no perfil, são indicadas ao Agente Buscador, caso contrário é informado ao agente que se use as bases previamente estipuladas, bases padrão.

A seleção das bases de dados para consulta é feita pelo Agente Buscador, que em posse de quais são as bases recomendadas para aquele usuário ou das bases padrão, envia mensagem aos Agentes Indexadores solicitando quais agentes indexam aquelas bases. E para esses agentes, são submetidos os termos, através de mensagem, e o agente aguarda como resposta uma lista de objetos que atendam aos termos submetidos.

Recebendo as listas de objetos dos Agentes Indexadores, o Agente Buscador as organiza em uma única listagem. Nesse ponto, o Agente Buscador tem em sua posse os objetos retornados da consulta às bases, os termos submetidos e se existe o perfil de usuário, já informado pelo Agente Perfil. Com essas informações, se existe o perfil do usuário, o agente toma uma decisão, se o usuário tem um perfil traçado ele envia a lista de objetos para o Agente Perfil e solicita que o agente a ordene por relevância, caso contrário o próprio Agente Buscador, com base em uma função de similaridade dos termos com os objetos, irá ordená-los.

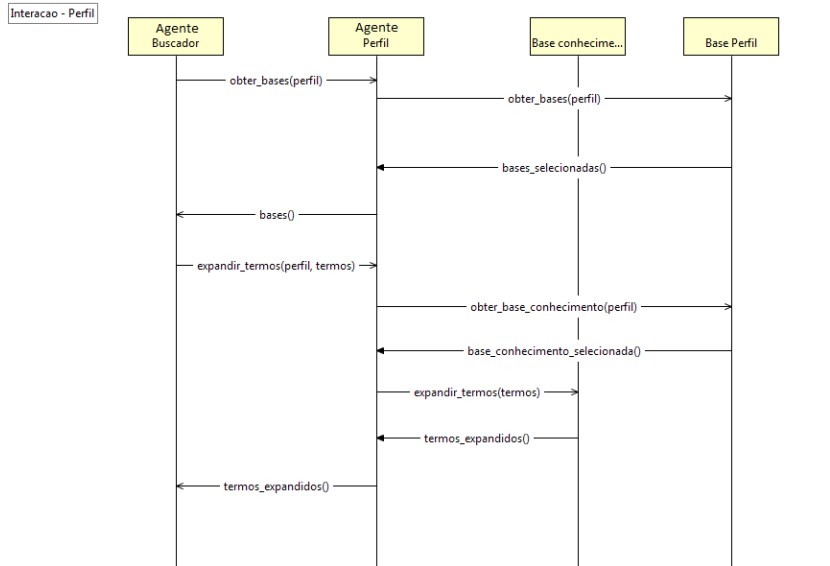
Na etapa final, o Agente Buscador em posse dos objetos ordenados, responderá a solicitação do sistema web, passando um vetor de dados contendo a informação de cada objeto, metadados, informações para atualizar o perfil e índice de relevância, para que o sistema *web* construa uma visualização para o usuário.

4.2 AGENTE PERFIL

No sistema multiagentes, o Agente Perfil é responsável por enriquecer a busca, produzindo resultados especializados e personalizados para o usuário. Para desempenhar tal tarefa, o agente interage com duas bases, uma de perfis e outra de dicionário de termos e conhecimento. O Agente Perfil é capaz, então, de obter o perfil do usuário; selecionar as bases de dados adequadas ao usuário; expandir e detalhar os termos para busca; ordenar e filtrar os objetos retornados,

por relevância; e auxiliar na construção e refinamento do perfil do usuário. O Agente Perfil possui duas etapas de interação. A primeira é definida por um diagrama de interação e é apresentada na figura 5.

Figura 5: Diagrama de interação do Agente Perfil - primeira etapa.



Estão representados na figura 5, os agentes: Buscador e Perfil, a base de conhecimento e a base de perfis. As trocas de mensagem entre agentes são representadas por uma seta, indicando o tipo da mensagem e seus parâmetros, já as chamadas internas ou protocolos que não seguem o padrão FIPA-ACL são expressos por uma seta preenchida. A comunicação entre o Agente Perfil e a base de conhecimento e a base de perfis é realizada por consulta, ou por uma linguagem *ad-hoc* de comunicação.

Inicialmente, o Agente Perfil entra em operação quando recebe uma solicitação do Agente Buscador, contendo informações sobre a identificação do usuário, e os termos submetidos à busca. Com a identificação do usuário, o Agente Perfil solicita para a base de perfis se o perfil do usuário existe. Para isso, a técnica que define a forma de identificação do usuário deve ser conhecida tanto pelo agente quanto pela base, de mesma forma ambos devem ser capazes de comunicar-se de uma forma padrão, por consulta ou usando uma linguagem *ad-hoc*.

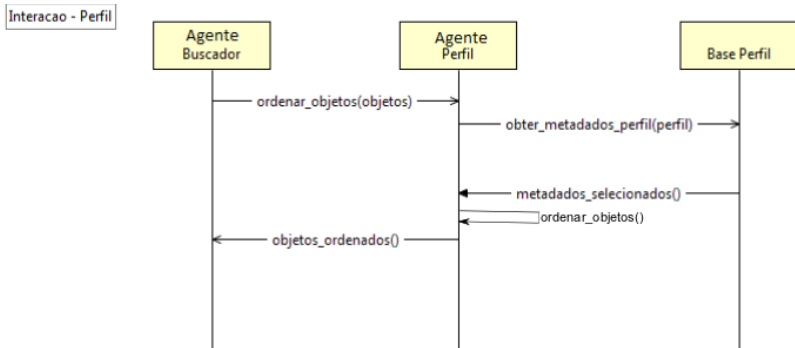
Sendo capaz de identificar e obter as informações do perfil do

usuário, o agente, quando recebe uma solicitação de termos e base de dados para pesquisa, envia mensagem à base de perfis solicitando, quais bases de dados o usuário costuma utilizar em sua busca e qual a base de conhecimento que ele deve utilizar para tratar os termos para busca.

Para a seleção da base de conhecimento, o Agente Perfil primeiro checa se existe o perfil do usuário, caso exista e em posse das informações sobre o perfil, o agente verifica se o utilizador tem estabelecido o uso de alguma base de conhecimento, e então seleciona tal base, caso não tenha sido estabelecida nenhuma base ou o perfil não foi encontrado, o agente então tenta inferir a base de conhecimento. Essa inferência é realizada submetendo-se os termos a uma função de similaridade textual com os termos contidos nas bases, dessa forma se alguma das bases possui um alto grau de similaridade dos seus termos com os termos da consulta, ela é uma provável candidata a base de conhecimento a ser utilizada. Em posse de qual base de conhecimento utilizar, o agente comunica-se com ela realizando uma consulta, a essa base o agente irá submeter os termos e receberá os termos expandidos para a busca. Como exemplo, o agente consulta a base especializada em termos médicos, submetendo os termos AVC e Circulação Sanguínea, a base então recebe os termos, expande os sinônimos e os comunica ao agente, nesse caso tem-se que o termo AVC possuiu como sinônimos: Derrame Cerebral, Ictus Cerebral, Acidente Vascular Cerebral, etc e que Circulação Sanguínea está relacionada com Sistema Cardiovascular.

A segunda interação do Agente Perfil, também definida por um diagrama de interação, é apresentada na figura 6.

Figura 6: Diagrama de interação do Agente Perfil - segunda etapa.



A figura 6 representa a interação do Agente Buscador e Perfil. A comunicação entre os agentes é expressa por uma seta com o tipo da mensagem e seus parâmetros, e a interação do Perfil com a base de perfis é feita por consulta ou por uma comunicação *ad-hoc* e é representada no diagrama por uma seta preenchida.

O Agente Perfil é incumbido de ordenar os objetos, resultantes da busca nos repositórios, por relevância. Não havendo um perfil de usuário selecionado, o agente irá utilizar uma função de similaridade dos termos com os objetos e, dessa forma, ordená-los. Existindo o perfil, o agente, também irá aplicar uma função de similaridade, mas utilizando os valores de metadados mais utilizados, informação contida no perfil do usuário como filtro, dando aos objetos que se enquadrem nesses valores maior relevância. Como exemplo, o perfil do usuário contém o campo Data definido como datas maiores que 2005, ao ordenar os objetos, havendo objetos cujo metadado Data se enquadre nessa definição, eles recebem um peso maior e conseqüentemente são ordenados como tendo uma maior relevância.

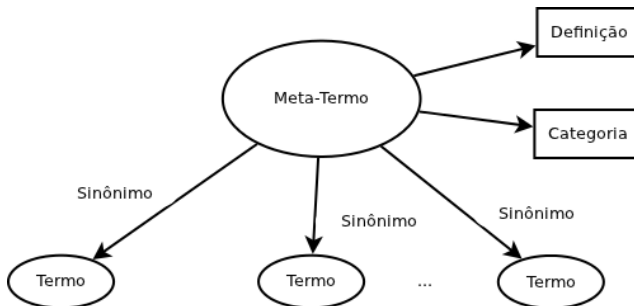
Ordenados os resultados, o agente prepara os objetos para envio montando uma lista com as informações sobre o objeto, campos de metadados e localização. A lista é submetida ao Agente Buscador, que a repassa ao sistema web de busca respondendo a sua solicitação. Com essas informações o sistema web monta a visualização dos objetos para o usuário, destacando os objetos com maior relevância e para cada objeto mostrado constrói uma função de atualização de perfil. Tal função atualiza a base de perfil do usuário conforme ele acessa os itens mostrados na interface web, analisando os campos do perfil que se referem aos metadados, estabelecendo padrões pelo cálculo da frequência dos valores dos campos e, assim, atualizando o perfil.

4.2.1 Ontologia

A expansão dos termos de busca, responsável por garantir uma maior cobertura na busca, é realizada pelo Agente de Perfil com uso de bases de conhecimentos específicas expressas por uma ontologia de domínio. Essa ontologia representa um dicionário de termos que além de expandir os termos, através de seus sinônimos, os define e categoriza afim de qualificá-los, evitando ambiguidades e termos fora de sua área de conhecimento. Como exemplo, AVC na área médica pode significar Acidente cerebral Vascular (termo mais comum) ou é referência para um protocolo de quimioterapia. Vírus pode ser vírus de computador ou vírus da biologia ou da área médica, ou até mesmo da área social.

Essa ontologia segue um modelo padrão, para ser consultada pelo Agente de Perfil, conforme a figura 7.

Figura 7: Modelo de ontologia de domínio



Segundo a figura 7 os termos estão relacionados a um meta-termo por relação de sinônimo. Esse meta-termo define um conceito ou ideia em alto nível, contendo sua definição, categoria para organização e termos que representam esse conceito. Como a relação do meta-termo com os termos é determinada por sinônimos, ao se localizar um termo, é possível obter seu meta-termo e por seguinte todos os seus sinônimos. Dessa maneira é realizada a expansão dos termos. Fazendo uso da mesma base de conhecimento, o Agente Perfil utiliza a definição e os sinônimos dos termos para o cálculo de similaridade dos objetos, dados os termos inseridos no sistema, sua definição e seus sinônimos.

4.2.2 Base de Perfis

A base de perfil é utilizada pelo agente perfil para obter informações do usuário. Pode ser implementada sobre diferentes técnicas, desde que a técnica aceite as mensagens do Agente de Perfil. Cada perfil contém uma série de informações específicas que definem a característica dos objetos a serem recomendados, são elas:

- Base de dados mais utilizadas (preferidas): Um dos pontos marcantes dos perfis é a área de conhecimento e pesquisa do utilizador, ao manter uma listagem das bases de dados mais utilizadas pelo usuário, a busca é direcionada a essas bases produzindo resultados mais relevantes. Dessa forma também se formam relações entre perfis e padrões de bases por áreas. A informação das bases mais utilizadas é obtida pelo agente de perfil, quando o utilizador, de forma explícita, seleciona as bases para consulta, ou encontrando perfis semelhantes e obtendo deles as bases de dados preferidas.
- Valores padrões para campos de metadados: Os sistemas de busca, em sua maioria, possuem sistemas de busca avançada. São compostos por filtros de campos de dados (ou metadados) ou busca truncada. O perfil do utilizador armazena esses filtros de campos avançados, conforme a utilização do sistema, para sugerir, durante a busca, esses filtros. Assim o esforço para a descrição da busca é significativamente reduzido. Da mesma forma os valores padrões são obtidos pela interação do usuário com o sistema. O agente de perfil em posse das informações interativas do utilizador constrói os padrões de metadados. Por exemplo, o utilizador realiza algumas consultas em que o metadado de data é superior a 2005, logo o agente de perfil infere que ele busca documentos recentes e já inclui tal filtro na próxima pesquisa.
- Relações com outros perfis: Quando não ha informações suficientes, ou para enriquecer o perfil do utilizador, é traçado um relação com outros perfis semelhantes. Dessa forma é traçado um relacionamento por conteúdo e a função social é evidenciada. Perfis com informações semelhantes são agrupados em um meta-perfil, isto é, um perfil genérico organizado por área de conhecimento. Esse perfil é construído com base em uma função social, e serve para ajudar a traçar o perfil inicial do usuário.

- Bases de conhecimento: Identificam ontologias de domínio, e dicionário de termos para que sejam, na etapa da busca, realizadas a expansão e desambiguação dos termos da pesquisa. O modelo, para auxiliar na busca e filtragem de resultados, foi concebido para utilizar bases de conhecimento organizadas por ontologias de domínio. Essa ontologia deve estar expressa em OWL e organizada em bases de conhecimento.

Inicialmente o sistema deve conter um conjunto de perfis especializados, construídos com foco nas especificidades de cada área de conhecimento abordadas pelo sistema. Esse perfil especializado contém uma lista de bases de dados para pesquisa, voltadas para a área em questão, e informações sobre qual base de conhecimento utilizar para tratar os termos. Exemplificando, um perfil especializado em medicina irá conter uma lista de base de dados para pesquisa que sejam voltadas à área médica, e que os termos submetidos a esse perfil deverão ser tratados por uma base de conhecimento específica em medicina. Dessa forma, todo usuário que se inscreve no sistema indica a qual área pertence, e com isso herda as informações específicas desse perfil, atendendo dessa forma a característica de relações com outros perfis.

Como o sistema indexa diferentes bases de dados, o utilizador pode optar por realizar a pesquisa em um sub-conjunto delas, retirando bases que julga ser de menor relevância, ou adicionando bases que possam lhe trazer mais resultados, dessa forma ele constrói uma lista das bases de dados mais utilizadas. Essa interação é realizada diretamente pelo sistema web de busca, através de campos de busca avançada. O sistema web é, então, responsável por atualizar a base de perfis.

Dentre as informações contidas no perfil do usuário, está a identificação da base de conhecimento. Essa base foi previamente construída através do modelo especificado na pesquisa. Essa base é referenciada pelo perfil especializado e é herdada pelo usuário quando ele cria o seu perfil. Tecnicamente ela é uma referencia, podendo ser implementada de várias formas, desde que seja acessível ao Agente de Perfil via troca de mensagens.

Os valores padrões são importantes, é principalmente nessa etapa que o sistema, através das interações do usuário, define um conjunto de metadados relevantes para a busca e seus valores padrões, agindo como filtro de objetos, reduzindo o número de resultados fora de conformidade com a busca. Como exemplo, um usuário que costuma recuperar objetos recentes, através das interações com o sistema ele aprende esse padrão e estipula um filtro de resultados baseado no campo de metadado Data e utiliza a Data maior que 2005.

No protótipo os seguintes campos são armazenados e refinados:

- Data (intervalo)
- Contexto Educacional
- Área de conhecimento
- Tipo
- Público alvo
- Interação
- Estratégia didática

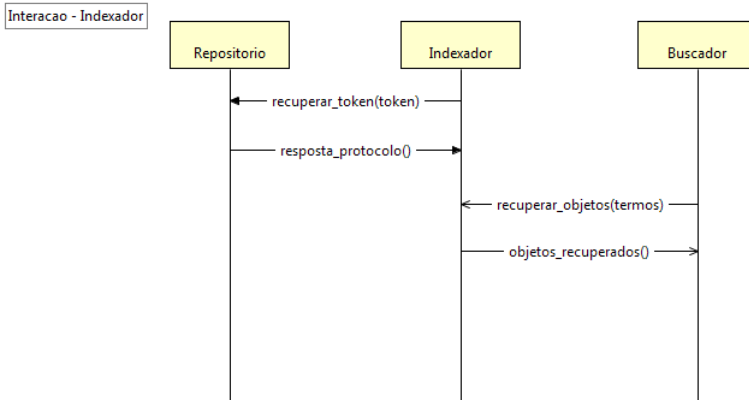
Esses campos foram definidos com base em entrevista com dois especialistas, que os consideraram os mais relevantes para identificação e filtragem dos objetos de aprendizagem.

4.3 AGENTE INDEXADOR

O Agente Indexador foi modelado para indexar diferentes repositórios de objetos, com heterogeneidade de acesso e de descrição de objetos. Dessa forma, para cada repositório indexado no sistema, existe um Agente Indexador específico. Para desempenhar tal tarefa o agente deve ser capaz de acessar os repositórios de objetos; obter os objetos nele contidos; obter a descrição desses objetos, ou seja, os metadados; obter, se disponíveis, informações adicionais colaborativas, como número de acessos e qualificações; e construir um índice com as informações coletadas.

A interação do Agente Indexador com o repositório e com o Agente de Buscador é definida na figura 8, por um diagrama de interação.

Figura 8: Diagrama de interação do Agente indexador.



Segundo a figura 8, os elementos do diagrama são: os agentes Buscador e Indexador e o Repositório. A troca de mensagem entre os agentes é representada por setas com o tipo da mensagem e seus parâmetros. Já a comunicação entre o Agente Indexador e o Repositório é realizada por um mecanismo interno do agente, o qual é capaz de reconhecer o repositório no ambiente, acessá-lo por meio de um protocolo e recuperar seus objetos e dados. Esse mecanismo é representado por uma seta preenchida, onde do agente para o repositório é realizada uma chamada ao protocolo de acesso implementado, de forma iterativa, obtendo os dados atualizados dele. O repositório responde, via protocolo, as informações necessárias para o agente atualizar seu índice.

O acesso ao repositório, aos objetos e sua descrição, é realizado por um protocolo de disseminação. A tendência dos repositórios é utilizar um dos padrões de protocolos de disseminação e coleta, como por exemplo, o OAI-PMH. Porém existem casos, como os repositórios da BVS (Biblioteca Virtual em Saúde), em que o mecanismo é desenvolvido de forma específica (*ad-hoc*). Existem repositórios em que o acesso se dá através de uma estrutura de dados em LDAP, ou por protocolos de busca como o SPARQL. Cada Agente Indexador deve, então ser especializado em um repositório e em seu protocolo ou mecanismo de disseminação, garantindo o acesso aos objetos nele contidos.

Relacionado ao protocolo de coleta estão as capacidades de acessar o repositório e obter os objetos e suas descrições, pois para cada tipo de protocolo tem-se um tipo diferente de técnica de descrição de metadados, por exemplo os metadados coletados pelo protocolo OAI-

PMH são representados em uma estrutura em XML. Ao definir o protocolo utilizado pelo agente, também devem ser definidas as estruturas de descrição dos objetos e seus metadados, tornando o agente capaz de entendê-los e recuperá-los.

Em posse dos objetos e sua estrutura conhecida, o agente constrói um índice com as informações coletadas. É através desse índice que o Agente Indexador retorna os objetos solicitados pelo Agente Buscador. Esse índice precisa estar padronizado para que os diversos objetos, oriundos dos diversos repositórios, estejam descritos e acessíveis de uma forma padrão. Para tal, os metadados devem ser mapeados para um conjunto padrão, dessa forma, pode-se:

- usar uma base de mapeamento, selecionando um conjunto de metadados e mapeando os objetos para esse conjunto;
- usar um padrão, como o LOM, e mapear os objetos para esse padrão;
- fazer a recuperação por campos específicos, como título e palavras-chaves, e mandar os metadados no formato original.

Ao utilizar uma base de mapeamento, elabora-se uma lista dos metadados relevantes para busca e mapeiam-se os padrões conhecidos para essa lista, formando um conjunto padrão de metadados. Assim, nem todos os metadados contidos nos objetos serão mapeados, somente os definidos como relevantes, tornando o índice mais simples e robusto. Como exemplo pode-se checar o quadro 1.

Ao se adotar um padrão, como o LOM, e mapear os objetos a esse padrão, opta-se por se trabalhar com todo o sistema em um único padrão, simplificando a etapa de implementação do modelo e o tornando passível de disseminação. Porém, objetos que estejam pouco descritos ou que estejam representados em um padrão mais simples, como o Dublin Core, não estarão com a maioria dos campos descritos. Também objetos que contenham campos que não existam no padrão a ser mapeados, ficam com a sua descrição incompleta, impactando nos resultados da busca.

Utilizar o formato original de descrição do objeto, ao invés do mapeamento, e manter sua descrição nesse formato, simplifica o agente, e garante a total representatividade do objeto. Porém traz uma complexidade a mais aos outros agentes do sistema, pois eles devem ser capazes de entender todos os padrões de metadados conhecidos pelo sistema.

4.3.1 Desempenho

Os repositórios podem conter um grande número de objetos. Analisar todos eles para selecionar os mais relevantes para busca pode ser uma tarefa muito custosa. Em sistemas em que o tempo de resposta não é um fator limitante o problema pode se tornar nulo, porém em sistemas com compromisso em desempenho, grandes índices podem dificultar o tempo de resposta dos agentes. Para amenizar tal questão, têm-se duas alternativas, usar algoritmos de pesquisa em índices grandes ou utilizar um sistema de leilão para os agentes indexadores.

Para diminuir o tempo de pesquisa do agente no índice de objetos pode-se adotar algum algoritmo específico de busca em índices grandes, como por exemplo, um grafo de índices invertidos. Dividir o índice em subíndices ordenados por áreas ou subáreas de conhecimento, ou armazenar os resultados das buscas mais realizadas, para uma consulta mais rápida.

Ao utilizar um sistema de leilão garante-se o tempo de resposta, pois o Agente Buscador abre um leilão de uma quantidade de objetos com os termos selecionados e com termino em um tempo determinado. Assim, cada agente indexador consulta a tabela de leilão e conforme for consultando seu índice, responde com os objetos com maior grau de relevância. Ao final do leilão o Agente Buscador coleta os resultados..

Seguem um exemplo, onde o Agente Buscador abre um leilão para um tempo determinado de resposta, com quatro propostas para o termo vírus.

Quadro 2: Exemplo de Leilão

	Vírus	Vírus	Vírus	Vírus
Agente 1	9,7	8,35	8,2	7,69
Agente 2	2,1	3,7	4	5,62
Agente 3	7,8	9,35	8,9	8,2

Esquema de leilão com "n" compradores e "m" vendedores. Matriz onde as linhas são os agentes e as colunas os termos requisitados repetidos 4 vezes. O conteúdo das células é o coeficiente de similaridade do objeto com o termo. Com base na tabela os agentes escalonam seus objetos a fim de tentar vencer o leilão.

Com a finalidade de avaliar e validar o modelo proposto, foi

desenvolvido um protótipo do sistema multiagente, fazendo uso de uma ontologia que representa termos na área médica, para enriquecer os termos da busca, e de uma base de perfis com informações do usuário, para classificar e qualificar os resultados. Tal protótipo seguiu as etapas do modelo proposto, tanto para a construção do SMA quanto para a elaboração da base de conhecimento e perfil.

4.4 IMPLEMENTAÇÃO E VALIDAÇÃO DO MODELO

Com a finalidade de avaliar e validar o modelo proposto, foi desenvolvido um protótipo do sistema multiagente, fazendo uso de uma ontologia que representa termos na área médica, para enriquecer os termos da busca, e de uma base de perfis com informações do usuário, para classificar e qualificar os resultados. Tal protótipo seguiu as etapas do modelo proposto, tanto para a construção do SMA quanto para a elaboração da base de conhecimento e perfil.

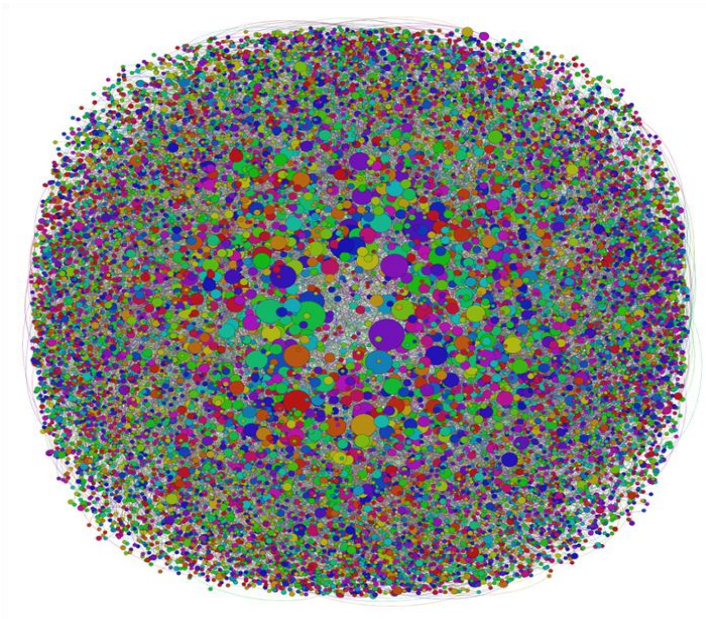
A interface homem-máquina, o sistema web de busca, foi implementado na linguagem PHP, fazendo uso da biblioteca XMLRPC (XMLRPC, 2013) para comunicação com o sistema multiagente. O sistema web é composto por um campo de texto, para receber os termos para busca, e um seletor de perfis especializados. Quando acionado, ele verifica se o perfil foi identificado e coleta os termos de busca, submetendo-os ao sistema multiagente via chamada remota de processo. As informações, termos e perfil, são codificadas em um conjunto de palavras (*String*) em formato JSON (JSON, 2013) e enviadas como parâmetro para a chamada *busca*, implementada pelo Agente Buscador, que também responde, ao final do processo, um conjunto de palavras em JSON contendo os resultados.

O sistema multiagente foi implementado utilizando o JADE (JADE, 2013) um framework de desenvolvimento de sistemas multiagentes que, segundo sua descrição, simplifica o processo de implementação e está de acordo com o padrão FIPA (FIPA, 2013). Com isso é possível traçar um paralelo com as características do modelo referentes à comunicação e à integração dos agentes com os sistemas periféricos que dão suporte ao modelo.

A base de conhecimento, representada pela ontologia descrita na sessão 4.2.1, traz um vocabulário de termos médicos, o DECS, contendo

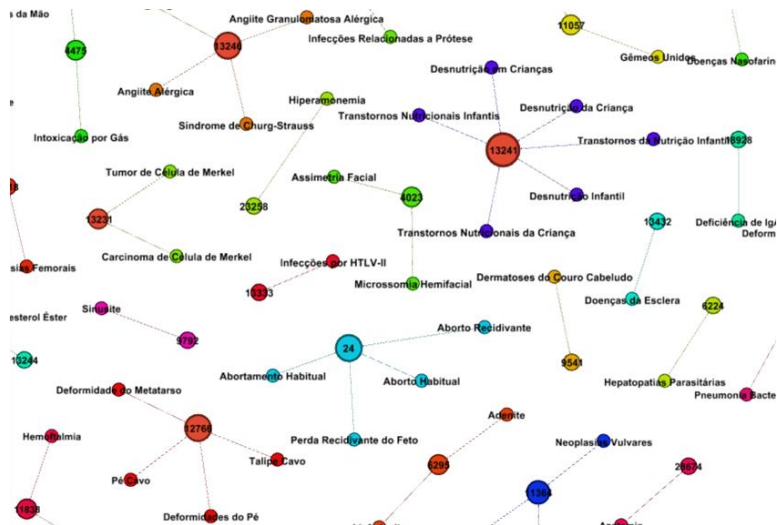
30369 índices (meta-termos) e mais de cem mil termos através de relações de sinônimos. A figura 9 trás uma representação gráfica desse vocabulário.

Figura 9: Representação gráfica do vocabulário DECS



A figura 9 representa o vocabulário DECS onde cada círculo corresponde a um termo, e a relação entre os termos é representada por uma aresta. A figura 10 representa o mesmo vocabulário, em um nível de detalhamento maior.

Figura 10: Representação gráfica do vocabulário DECS - Visão detalhada



Para validação apenas um perfil especialista foi criado, contendo o perfil de um especialista na área médica. Dessa forma ele faz uso da base de dados ARES UnA-SUS como base preferida e como metadados padrão apresentado no quadro 3, representando um profissional na área de saúde cujo interesse é maior em objetos que sejam recentes à 2005 e que atendam tanto ao nível técnico de conhecimento quanto ao superior.

Quadro 3: metadados padrão para perfil especializado.

Metadado	Valor Padrão
Data	Maior que 2005
Público Alvo	Mestrado
Área de Conhecimento	Ciências da Saúde
Tipo de objeto	Material Multimídia

Os repositórios indexados foram LUME (LUME, 2013), tendo seu padrão de metadados o Dublin Core e como protocolo de disseminação o OAI-PMH e o ARES UnA-SUS que utiliza como descritor de objetos um perfil do OBAA e como disseminador o protocolo OAI-PMH. Ambos foram indexados no dia 23 de janeiro de 2013, sendo coletados um sub-conjunto de metadados que atendem ao mapeamento do quadro 1 e que estão presentes em ambos os repositórios, resultando

nos seguintes itens: Identificador (URI), Título, Autor, Resumo, Palavras-chave, Data, Tipo do Objeto, Público Alvo e Relações. Foram coletados 375 objetos no ARES e 60.663 objetos no LUME.

Tanto para obter, quanto para ordenar os objetos, foi utilizada uma função de classificação por cobertura de densidade (*cover density ranking*) (CLARKE; CORMACK; TUDHOPE, 1999) indexada sobre os metadados:

- Título, com peso A (1.0), ou seja, cada ocorrência é multiplicada por 1
- Palavras-chaves, com peso B (0.4), cada ocorrência é multiplicada por 0.4
- Resumo, com peso C (0.2), cada ocorrência é multiplicada por 0.2

Tal função trabalha separando o texto em palavras, armazenando a sua posição dentro do texto. Quando uma busca é realizada, o algoritmo encontra as ocorrências dos termos da busca, criando um conjunto

$$\mathcal{C} = \{(p_1, q_1), (p_2, q_2), \dots, (p_n, q_n)\}$$

Onde p indica a posição do primeiro termo e q é a ocorrência do último termo da busca. Sobre esse conjunto é aplicada a fórmula:

$$S(\mathcal{C}) = \sum_{j=1}^n I(p_j, q_j),$$

$$I(p, q) = \begin{cases} \frac{\mathcal{K}}{q - p + 1} & \text{Se } q - p + 1 > \mathcal{K}, \\ 1 & \text{caso contrário} \end{cases}$$

Onde k é uma constante de cobertura de tamanho. Sobre essa pontuação S é aplicada então a multiplicação dos pesos A, B e C e esse valor é utilizado como resultado da função de similaridade.

Dessa forma o índice de classificação de um objeto, pela aderência dos termos a ele aplicados por essa função, é normalizado em valores entre 0 (não aderentes) e próximos de 1 (fortemente aderentes ou similares).

Os agentes indexadores utilizam essa função para obter objetos

com índice maior que 0. Já o Agente Perfil utiliza tal função para obter os índices dos objetos e sobre eles realizar uma pontuação, caso julgue necessário, ao observar o perfil do utilizador, adicionando, para cada metadado do objeto que tenha correspondência com o padrão contido no perfil, uma ocorrência de um termo com peso B, ou seja é acrescido ao índice uma pontuação de 0.4.

4.4.1 Testes

Para avaliar as vantagens trazidas pela expansão dos termos, foi definida uma lista de dez termos, onde cada termo foi pesquisado nos repositórios LUME e ARES ao sistema multiagente. Cada termo foi aplicado ao:

- Sistema de busca simples do LUME
- Sistema de busca avançada do LUME, buscando no Título ou no Assunto
- Sistema de busca do ARES, que utiliza um motor próprio do DSPACE chamado *Discover*
- Sistema de busca do SMA

Foi coletado o total de objetos encontrados para cada termo em cada repositório, bem como o total de objetos encontrado no SMA. Os resultados da submissão de cada termo aos repositórios LUME, ARES e ao SMA, destacando o número total de objetos recuperados pelo sistema de busca de cada ferramenta, estão representado na tabela 1.

Tabela 1: Total de objetos recuperados por repositório.

Termo	LUME	ARES	SMA/LUME	SMA/ARES
Acidente Cerebral	34	00	1110	00
Gripe	02	00	15	00
Atenção Primária	144	311	307	121
Odontopediatria	23	03	25	03
Dor de Cabeça	01	00	48	00

Doença de Chagas	04	00	07	00
Doenças Respiratórias	12	02	146	04
Sistema Único de Saúde	83	112	2927	133
Dor de Garganta	00	00	02	00
Dengue	08	06	09	06

A primeira coluna da tabela representam os termos submetidos aos repositórios, a segunda coluna traz o número de objetos encontrado pelo repositório LUME relativo ao termo que se encontra na mesma linha, a terceira coluna indica o número total de objetos encontrado pelo ARES. As duas ultimas colunas representam, respectivamente, o número total de objetos encontrados pelo SMA no repositório LUME, e no repositório ARES. Dessa forma, para o termo "Acidente Cerebral", o repositório LUME encontro 34 objetos, o ARES não teve resultados, o SMA indexado com os objetos do LUME encontrou 1110 objetos e nenhum objeto no índice do ARES. Essa diferença é bem expressiva, e está relacionada a expansão dos termos da busca, realizada pelo Agente Perfil, dessa forma o termo "Acidente Cerebral" foi expandido para 14 termos, tal expansão pode ser observada no quadro 4, e esses termos foram então submetidos a uma busca no índice do LUME. Observando a tabela 1, pode-se perceber que a expansão dos termos, para o repositório LUME, trouxe, em todos os casos, mais objetos, proporcionando assim uma boa cobertura dos objetos contidos no repositório. Já para o repositório ARES, apenas em dois casos a expansão trouxe mais resultados que o mecanismo do próprio repositório. Nesse caso, fazendo uma análise sobre o mecanismo de busca e os resultados obtidos pelo ARES, observou-se que seu mecanismo é capaz de indexar o texto contido nos objetos em formatos de texto digital (Microsoft Word e Adobe PDF).

A expansão de cada termo, utilizada pelos agentes do SMA para realizar a busca pode ser observada no quadro 4.

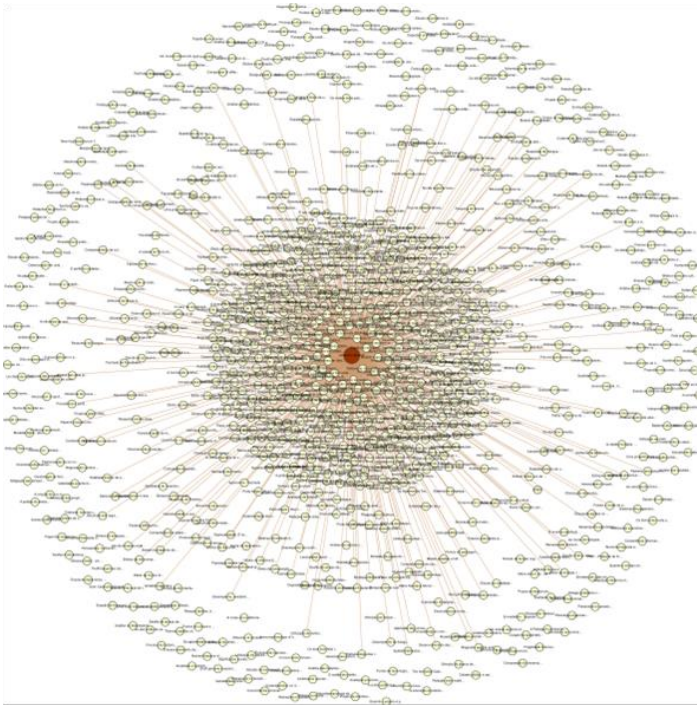
Quadro 4: expansão dos termos.

Termo	Termo Expandido
Acidente Cerebral	Acidente Cerebral; Acidente Cerebral Vascular; Derrame Cerebral; Ictus Cerebral; AVC; Apoplexia; Acidente Cerebrovascular; Apoplexia Cerebral; Apoplexia Cerebrovascular; Icto Cerebral; Acidente Vascular Encefálico; AVE; Acidente

	Vascular do Cérebro; Acidente Vascular Cerebral
Gripe	Gripe; Influenza Humana; Gripe Humana; Influenza em Humanos
Atenção Primária	Atenção Primária; Atenção Primária à Saúde; Atenção Básica; Atenção Básica à Saúde; Atenção Primária de Saúde; Atendimento Primário; Cuidados Primários; Cuidados Primários de Saúde
Odontopediatria	Odontopediatria; Pedodontia; Odontologia Pediátrica; Pedodontia
Dor de Cabeça	Dor de Cabeça; Cefaléia; Cefalgia; Hemicrania
Doença de Chagas	Doença de Chagas; Tripanossomose Sul-Americana
Doenças Respiratórias	Doenças Respiratórias; Doenças do Aparelho Respiratório; Doenças do Sistema Respiratório; Doenças do Trato Respiratório
Sistema Único de Saúde	Sistema Único de Saúde; SUS
Dor de Garganta	Dor de Garganta; Faringite
Dengue	Dengue; Febre da Dengue; Febre Quebra-Ossos

A expansão dos termos, exemplificada no quadro 4, é realizada pelo Agente de Perfil ao consultar uma base de conhecimento na área médica. Essa base foi construída utilizando o modelo abordado no capítulo de modelagem e está de acordo com o vocabulário contido nos Descritores em Ciências da Saúde (DECS) em versão datada de 2011. Observa-se que em sete casos os termos foram expandidos para três ou mais termos, e em três casos para dois termos sendo que em um deles, Sistema Único de Saúde, foi acrescentado uma abreviatura (SUS) comumente utilizada. Os resultados da consulta ao sistema multiagentes, utilizando o termo Acidente Cerebral, é representado na figura 11.

Figura 11: Representação gráfica dos resultados obtidos para Acidente Cerebral



Para avaliar a ordem e relevância dos resultados obtidos pela consulta ao SMA com termos expandidos, foram selecionados três termos, os quais foram submetidos aos repositórios e ao SMA. Para análise, foram coletados os dez primeiros resultados de cada repositório e do SMA. O primeiro termo analisado foi "Doenças Respiratórias", os nomes dos objetos retornados foram codificados, para simplificação e melhor entendimento do quadro, como Objeto 1, Objeto 2, etc. O Quadro 5 representa a busca do termo "Doenças Respiratórias" ao SMA indexado com o repositório ARES, e do próprio ARES.

Quadro 5: comparação entre os resultados do SMA e ARES para o termo Doenças Respiratórias.

Objeto	SMA/ARES	ARES
Objeto 1	01	01
Objeto 2	02	-
Objeto 3	03	-
Objeto 4	04	02

O quadro 5 apresenta os objetos recuperados com sua posição no SMA ao indexar o repositório ARES e a terceira coluna representa a posição do objeto no sistema de busca do próprio ARES. Pelo quadro 5, percebe-se que o SMA trouxe dois objetos a mais (Objeto 2 e Objeto 3) e os classificou com uma relevância maior do que o Objeto 4, o qual foi trazido pelo ARES na segunda posição. Fazendo uma análise mais profunda sobre os resultados, tem-se que os objetos 2 e 3 trazem em sua descrição termos que são sinônimos de "Doenças Respiratórias", e por isso foram encontrados e recuperados com uma relevância maior. Já o Objeto 4 trata de um assunto diferente e não relacionado diretamente com o termo pesquisado, mas que contém em seu resumo uma referência ao termo.

O mesmo termo, Doenças Respiratórias, também foi submetido ao sistema de busca do LUME e ao SMA indexado com os objetos do LUME. A busca no repositório LUME foi realizada de duas formas, utilizando o mecanismo de busca simples, e a busca avançada para o termo no título ou no assunto. Os resultados estão expressos no quadro 6.

Quadro 6: comparação entre os resultados do SMA e LUME para o termo Doenças Respiratórias.

Objeto	SMA/LUME	LUME - Simples	LUME - Avançado
Objeto 1	01	04	06
Objeto 2	02	14	05
Objeto 3	03	79	-
Objeto 4	04	48	-
Objeto 5	05	62	07
Objeto 6	06	48	-
Objeto 7	07	-	-
Objeto 8	08	05	09
Objeto 9	09	77	-
Objeto 10	10	81	11

O quadro 6 apresenta a codificação do objeto encontrado seguido da posição do objeto nos resultados do SMA e pela posição em que o objeto foi encontrado na busca simples e avançada do LUME, respectivamente. Quanto a busca simples, em apenas dois casos objetos entre os 10 primeiros do SMA também foram encontrados entre os 10 primeiros do LUME, um deles não foi retornado pelo LUME e os demais estão em posições muito longe, o que mostra que a busca

simples traz diversos objetos fora de contexto.

No quadro 7, em uma análise mais acurada dos objetos, percebe-se que a busca simples retorna objetos que atendam ao termo "Doenças" ou "Respiratórias", por isso o grande número de objetos fora de contexto. No que tange a busca avançada, o SMA recuperou 5 objetos que não foram listados pelo LUME, isso deve-se ao fato de que a busca no SMA expandiu os termos e realizou uma cobertura maior de objetos. Nos outros cinco resultados recuperados, apenas um deles ficou em uma posição além dos dez primeiros, em décimo primeiro, o que não afeta muito o cálculo de relevância, pois nesses casos os objetos estão listados com uma boa proximidade.

Quadro 7: comparação entre os resultados do LUME e SMA para o termo Doenças Respiratórias na busca simples do LUME.

Objeto	LUME	SMA/LUME
Objeto 1	01	28
Objeto 2	02	14
Objeto 3	03	Fora de contexto
Objeto 4	04	01
Objeto 5	05	09
Objeto 6	06	06
Objeto 7	07	Fora de contexto
Objeto 8	08	Fora de contexto
Objeto 9	09	13
Objeto 10	10	Fora de contexto

O quadro 7 mostra os dez primeiros resultados para busca simples do LUME. A primeira coluna representa a codificação do objeto, em seguida a ordem dos objetos retornados pela busca do LUME e a posição em que o mesmo objeto foi encontrado nos resultados do SMA. Em análise, três objetos estão entre os dez primeiros do SMA, três estão em posições medianas e quatro resultados apresentados pelo LUME estão fora de contexto e não foram classificados pelo SMA.

Para evidenciar o cálculo da similaridade entre termos e objetos, foi utilizada a busca de um termo que produziu o mesmo número de resultado tanto no SMA quanto no repositório; nesse caso foi utilizado o termo "Dengue", e foram escolhidos para análise os resultados da busca no repositório ARES. Consultando a tabela 1, pode-se ver que ambos, SMA e ARES, apresentam um total de seis objetos retornados. A ordem comparativa dos resultados pode ser vista no quadro 8.

Quadro 8: Comparação entre os resultados do SMA e ARES para o termo Dengue.

Objeto	SMA/ARES	ARES
Objeto 1	01	02
Objeto 2	02	03
Objeto 3	03	05
Objeto 4	04	06
Objeto 5	05	04
Objeto 6	06	01

O que se pode notar no quadro 8, onde a segunda coluna representa a posição do objeto nos resultados do SMA e a terceira coluna representa a posição do mesmo objeto na busca do ARES, é que os objetos estão presentes em diferente ordem. O último objeto encontrado no SMA, ou seja, com menor relevância, aparece como objeto mais relevante na busca no ARES e ao analisar o objeto, após uma análise do objeto percebe-se que ele possui uma baixa relevância ao termo.

Para analisar a principal característica do Agente Perfil, que é enriquecer e produzir resultados especializados e personalizados para o usuário, foi elaborado um teste com um perfil na área médica, contendo como refinamento os metadados, gerados pela interação do usuário com o sistema. Ou seja, o perfil para testes do SMA é um médico que busca por objetos de aprendizagem com data posterior a 2005, o publico alvo mestrado, a Área do Conhecimento Ciências da Saúde e Tipo de Objeto Material Multimídia, como visto anteriormente no quadro 3. Para o teste, além dos metadados, foi selecionado pelo agente o repositório ARES como base de dados preferida para busca, dessa forma tem-se:

- Um conjunto de base de dados preferidas, nesse caso somente o repositório ARES.
- Uma base de conhecimento especializada, para expandir os termos, o DECS.
- Um conjunto de metadados mais utilizados pelo usuário, contidos no quadro 3.

Inicialmente o teste foi realizado sem o perfil, submetendo ao SMA o termo "Enfermagem na atenção primária". Analisando as mensagens trocadas pelos agentes, foi constatado que o Agente Perfil inferiu o termo como pertencente ao vocabulário DECS, a busca foi realizada nos dois repositórios, uma vez que ambos armazenam objetos da área médica. Os resultados foram coletados e armazenados na tabela 2.

Tabela 2: Comparativo entre a busca sem perfil no SMA.

SMA sem perfil	Objeto	Função de Similaridade
01	Objeto 1	0,99966
02	Objeto 2	0,999507
03	Objeto 3	0,998867
04	Objeto 4	0,998373
05	Objeto 5	0,998373

Para simplificar a tabela 2, foram armazenados somente os resultados obtidos pelo repositório ARES. Em sua primeira coluna a tabela traz a ordem que foram retornados os objetos, em seguida a codificação do objeto e o resultado da função de similaridade atribuída pelo Agente Perfil ao objeto.

Em seguida o perfil foi ativado, com as informações já relatadas, e o mesmo termo, Enfermagem na atenção primária, foi submetido a busca. Já em posse do perfil do usuário, o Agente Perfil selecionou a base de conhecimento do DECS, o repositório ARES e os metadados contidos no perfil para ordenar os objetos. Os resultados podem ser vistos na tabela 3.

Tabela 3: Comparativo entre a busca com perfil no SMA.

SMA com perfil	Objeto	Pontuação
01	Objeto 3	1,078867
02	Objeto 1	1,03966
03	Objeto 2	1,039507
04	Objeto 4	1,038373
05	Objeto 5	1,038373

Comparando as tabelas 2 e 3, sem e com perfil, respectivamente, de forma inicial, nota-se a diferença no resultado da função de similaridade, resultante da interação do Agente Perfil com os objetos. Os resultados são os mesmos, porém em diferente ordem. O Objeto 3, na busca com perfil ativado, saiu da terceira posição e passou a ocupar a primeira, deslocando os objetos 1 e 2, para as posições 2 e 3, respectivamente. Com isso, pode-se concluir que o Agente Perfil, ao ordenar os objetos, prioriza objetos que atendam ao padrão pesquisado pelo usuário, dando a eles um peso maior e consequentemente uma posição mais alta na ordem dos objetos. Nesse exemplo, o Objeto 3 passou para a primeira posição, pois analisando seus metadados, pode-se ver que ele é um objeto com data maior que 2005, do tipo "Multimídia" e tem como público alvo alunos de mestrado e ainda tem uma alta

similaridade em relação ao termo buscado.

5. CONSIDERAÇÕES FINAIS

Em um ambiente de repositórios digitais com heterogeneidade de metadados e de acesso ao recurso, o uso de um sistema multiagente se justifica, pois como demonstrado, as características de distributividade e adaptabilidade vão ao encontro da resolução desta questão. Por serem facilmente extensíveis e versáteis, os modelos baseados em sistemas multiagentes são adequados para tratar os aspectos heterogêneos dos repositórios.

A possibilidade do sistema fazer busca em diversas bases de uma forma homogênea e centralizada é atrativa ao usuário, e ao definir informações sobre os termos da busca, minimizando a interação com o sistema usuário, identificando de forma inteligente tais informações e padrões de pesquisa do usuário, define-se um dos objetivos pretendidos por essa pesquisa.

O uso de informações colaborativas, como descritas por Gil, De La Prieta e Rodríguez (2011), agrega qualidade e relevância aos resultados da busca e recuperação. Essas informações foram definidas através de parâmetros de perfil do usuário, como critérios de recomendação. E conforme apontam os teste, esses parâmetros produzem bons resultados, minimizando também a descrição do usuário fornecida ao sistema. Porém Gil cobre de forma muito simples a heterogeneidade de acesso aos objetos, utilizando o próprio sistema de busca dos repositórios indexados. Dessa forma ficam-se sujeito as limitações desses sistemas de busca, não sendo possível inferir a qualidade dos resultados retornados por esses sistemas, nem a cobertura dos objetos.

Para sanar tal limitação, o modelo aqui apresentado tem como característica, realizar busca em repositórios com heterogeneidade de acesso. Para isso existe o agente especializado em comunicar-se com um repositório e recuperar seus objetos. Assim, o agente constrói um índice de todos os objetos e suas descrições, e sobre esse índice ele realiza a busca. Dessa forma é possível garantir uma boa cobertura e qualidade dos resultados. Também é possível adaptar o agente para incluir novos algoritmos de busca e cobertura de objetos, tornando o sistema mais robusto.

Vian *et al* (2011) destaca os ganhos trazidos pelo uso de uma base de conhecimento para expandir os termos da busca. Essa base armazena conhecimento expresso por ontologias, trazendo relações entre termos e conceitos. Essa base é acessada pelo agente *Searcher* no momento em

que ele recebe uma solicitação de busca, e a ela é feita uma consulta sobre os sinônimos e suas possíveis relações com os termos submetidos. Porém, devido ao fato do agente *Searcher* concentrar grande parte da inteligência do sistema a seleção de quando e que base de conhecimento usar ficou a cargo do usuário.

Para minimizar a descrição fornecida pelo usuário, conforme citado acima, e fazer uso de bases de conhecimentos para expandir termos, o modelo, aqui proposto, possui um agente Perfil que é capaz de tomar decisões com base no perfil do utilizador do sistema, podendo receber indicações de quais bases utilizar ou inferi-las.

Baseado nas características de construção de um perfil por iterações com o usuário como sugere D'Agostine (2009) o agente Perfil é capaz de construir um perfil para cada utilizador, partindo de um perfil especializado em uma determinada área de conhecimento, e através das iterações com o usuário ir adaptando tal perfil para que este atenda as necessidades do usuário. A opção em se utilizar um protocolo aberto baseado em XML para realizar a comunicação entre o sistema multiagente e a aplicação web de busca incorporou características distribuídas ao sistema, uma vez que qualquer tipo de aplicação que implemente o protocolo é capaz de se comunicar com o sistema, tornando-o mais acessível. Essa característica permite o uso da ferramenta de forma transparente por parte de outros sistemas, tais como AVA (Ambiente Virtuais de Aprendizagem), possibilitando a busca e incorporação de conteúdos por meio do próprio sistema.

Por realizar buscas em ambientes heterogêneos utilizando agentes especializados em coletar objetos, em repositórios com padrões distintos de metadados e de acesso aos objetos e seus descritores, a inclusão de novos repositórios ou bases para pesquisa é simplificada, resultando apenas na elaboração e inclusão de um agente especializado no sistema, o que facilita a expansão do mesmo.

Como o modelo prevê o uso de uma base de conhecimento para realizar inferências sobre os termos de busca e sobre os resultados da pesquisa, tecnologias de *web* semântica, como por exemplo, redes RDF e OWL, podem ser facilmente incorporadas e utilizadas pelos agentes.

A qualidade dos resultados atribuída as capacidades do Agente Perfil pode também ser facilmente estendida, uma vez que o agente é capaz de selecionar as bases de conhecimento e obter o perfil do utilizador, somente é necessário que mais bases sejam incluídas e que o perfil seja refinado pelas interações do usuário com o sistema.

Durante a pesquisa e elaboração do modelo, algumas dificuldades surgiram. A primeira delas diz respeito a interpretação dos campos e

valores dos metadados, as normas definem os campos e são acompanhadas de uma breve explicação, valores padrões e exemplos de uso, mas não são tão claras e há casos em que a explicação é simplificada, dando margem a diversas interpretações. Como exemplo o metadado Tipo, que no LUME é utilizado para definir o tipo de arquivo digital (formato), como um PDF, por exemplo, já o repositório ARES utiliza o metadado Tipo para representar o tipo de recurso digital, como por exemplo, o valor "Material Multimídia".

A respeito dos valores dos campos, nem todos possuem um conjunto padrão de valores, são referenciados na norma como campo de texto, e isso dificulta a interação com as ferramentas de coleta de metadados, pois valores diferentes podem significar a mesma coisa, como no metadado público alvo pode-se ter, com o mesmo significado, valores "Ensino Superior", "Terceiro Grau" ou ainda "Nível Superior". Em segundo, a definição das ontologias e bases de dados, pois requerem um conhecimento especializado e um formalismo descritivo, gerando um grande esforço para sua elaboração. Quanto a coleta dos objetos, nem sempre é possível garantir que o mesmo atinja um padrão de qualidade, corroborando com Gil, De La Prieta e Rodríguez (2011) que estabelece uma relação entre qualidade do objeto e completude e confiabilidade dos metadados preenchidos.

Foi identificado também a dificuldade em se acertar os parâmetros das funções de similaridade, por serem muito sensíveis a variações, a escolha de quais metadados recebem um peso maior, juntamente com a normalização da pontuação dos índices de similaridade e a implementação dos algoritmos, devem ser realizadas por um processo iterativo, por meio de testes, a fim de obter resultados mais precisos.

Conforme apresentado das 20 consultas submetidas aos repositórios e comparadas com as consulta ao SMA, em 60% casos o SMA produziu mais resultados, em 30% os resultados foram os mesmos, e somente em 10% casos foi menor. Constatando-se que o uso de uma base de conhecimento para tratar e expandir os termos da busca a enriquece, pois acaba cobrindo sinônimos e termos relacionados, trazendo mais resultados a consulta e encontrando objetos que dificilmente seriam recuperados pelos meios convencionais de busca.

Desta forma, os resultados apontam para uma melhora, tanto na cobertura quanto na relevância dos objetos recuperados. Ao fazer uso do mecanismo de expansão, mais objetos foram recuperados e para reduzir o problema da sobrecarga de informações foram incorporados aspectos de recomendação, trazidos pelo uso de algoritmos que, além de tratar a

densidade de cobertura dos objetos, lidam com aspectos semânticos presentes nos metadados, garantindo um grau de qualidade aos resultados.

5.1 SUGESTÃO PARA TRABALHOS FUTUROS

Pelo constante aperfeiçoamento das técnicas de busca e recuperação da informação, há a necessidade de uma continuidade de pesquisas relacionadas ao tema abordado.

Desta forma, como continuidade de pesquisa sugere-se testar a distributividade dos agentes, em especial os indexadores, rodando-os em diferentes plataformas de acordo com as especificações FIPA.

O uso de informações colaborativas, expresso pelo perfil do usuário, podem ser aperfeiçoados e expandidos, incluindo sugestões de objetos e assuntos entre perfis e compartilhamento de resultados. Para isso, sugere-se um estudo empírico das técnicas, heurísticas e algoritmos relacionados a recomendação por construção de um perfil.

Como o sistema utiliza um protocolo aberto para comunicação externa, é possível utiliza-lo em outros sistemas, além do sistema de busca *web*, de diversas formas. Dessa forma sugere-se o desenvolvimento de um canal de comunicação com um AVA, para que seja possível buscar e recuperar objetos através do próprio AVA, de forma transparente ao utilizador, tornando o SMA mais acessível e ampliando seu espectro de possibilidades.

Os algoritmos de similaridade textual, utilizados nas buscas, estão em constante evolução. O uso de técnicas que utilizam aspectos semânticos vem sendo explorado pelos sistemas de busca e, aliados a algoritmos estatísticos, tem produzido resultados satisfatórios. Por fim, sugere-se então o constante estudo e teste dos novos algoritmos, para complementar e qualificar ainda mais os resultados.

6. REFERÊNCIAS

- ADL. Advanced Distributed Learning [online]. Disponível em <http://www.adlnet.gov/Pages/Default.aspx> Acessado: 2 jan 2013.
- AICC. Aviation Industry CBT Committee [online]. Disponível em <http://www.aicc.org/dev/> Acessado: 2 jan 2013.
- ARIADNE. Ariadne Foundation, 2013. Disponível em <http://www.ariadne-eu.org/> Acessado: 2 Jan 2013.
- BARBETTA, Pedro Alberto. Estatística Aplicada às Ciências Sociais. 7. ed. Florianópolis: Editora da UFSC, 2007.
- BARGMEYER B.E., GILLMAN D.W. Metadata Standards and Metadata Registries: An Overview. Disponível em <http://www.bls.gov/ore/pdf/st000010.pdf> Acessado: 10 Jul 2011
- BIOE. Bank for International Educational Objects, 2013 Disponível em <http://objetoseducacionais2.mec.gov.br/> [Accessed: 10 May 2013].
- CESTA. Collection of Entities Support the use of Technology in Learning. Center for Interdisciplinary Studies in New Technologies in Education (CINTED), 2013. Disponível em <http://www.cinted.ufrgs.br/CESTA/>
- CLARKE, C.L.A.; CORMACK, Gordon V.; TUDHOPE, Elizabeth A. (2000). Relevance ranking for one to three term queries, Information Processing & Management, Volume 36, Issue 2, 1 March 2000, Pages 291-311, ISSN 0306-4573, 10.1016/S0306-4573(99)00017-5.
- CORDRA. Content Object Repository Discovery and Registration. Disponível em <http://cordra.net/> Acessado: 2 Jan. 2013
- COSTA S.M.; LEITE F.C.L. Repositórios institucionais: potencial para maximizar o acesso e o impacto da pesquisa em universidades. In: *Conferência Iberoamericana de Publicações Eletrônicas no Contexto da Comunicação Científica*, 1. Brasília, Universidade de Brasília, 25 a 28 de abril de 2006.
- D'Agostine , Caio S. Captura e Gerência de Informações de Contexto Semântico para Recuperação de Informação. 2009. 114 f. Dissertação (Mestrado em Ciências da Computação) - Programa de Pós-graduação em Ciências da Computação, Universidade Federal de Santa Catarina, Florianópolis, 2009.
- DCMI. The Dublin Core Metadata Initiative, 2013. Available in: <http://dublincore.org/> . Acessado: 4 Jan. 2013.
- DOWNES, S. Learning Objects: Resources for distance education worldwide. [online] The International Review of Research in Open and Distance Learning, v. 2, n. 1, 2001.
- DSpace. DSpace [online] Disponível em <http://www.dspace.org/> Acessado: 2 Jan. 2013.
- Dublin Core (2013) The Dublin Core Metadata Initiative. [online] Disponível em <http://dublincore.org/> Acessado: 2 Jan. 2013.
- FABRE, M. C. J. M.; TAROUCO, L. M. R. and TAMUSIUNAS, F. R. Reusabilidade de objetos educacionais. RENOTE - Revista Novas Tecnologias na Educação. Porto Alegre: Centro Interdisciplinar de Novas Tecnologias na Educação (UFRGS), v. 1, n. 1, 2003.
- FACHIN, G.R.B.. Ontologia de referência para periódico científico digital. 2011. Tese (Doutorado em Engenharia e Gestão do Conhecimento) - Programa de Pós-graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2011.

- FEB. Federation of Educa Brazil. [online] Disponível em <http://feb.ufrgs.br/> Acessado: 2 Jan 2013.
- FIPA. Foundation for Intelligent Physical Agents. [online]. Disponível em: <http://www.fipa.org/> Acessado: 10 Jan 2013.
- GIL A.B., DE LA PRIETA F. and RODRÍGUEZ S. Automatic Learning Object Extraction and Classification in Heterogeneous Environments. In: Highlights in Practical Applications of Agents and Multiagent Systems 9th International Conference on Practical Applications of Agents and Multiagent Systems, 2011, Volume 89/2011, 109-116.
- GRUBER, T. What is na ontology? 1996. Disponível em: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html> Acessado em 05 jun 2012.
- HEERY R.; ANDERSON S. Digital repositories review. [online] In: UKOLN. Ahd: arts and humanities data service. Fev. 2005. Disponível em http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf. Acessado: 2 May 2011.
- HOWES T.A.; SMITH M.C.; GOOD, G.S. Understanding And Deploying Ldap Directory Services. 2ª ed. Addison-Wesley Professional, 2003.
- IEEE-LTSC. Draft Standard for Learning Object Metadata [Online]. IEEE Learning Technology Standards Committee, 2002. Disponível em http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf Acessado: 10 mai 2011
- IEEE-LTSC. WG12: Learning Object Metadata [online] IEEE Learning Technology Standards Committee. Disponível em <http://ltsc.ieee.org> Acessado: 2 Jan 2013.
- IMS. IMS Global Learning Consortium, 2013. [online] Disponível em <http://www.imsglobal.org/> Acessado: 2 jan 2013.
- ISO. International Organization for Standardization [online] Disponível em <http://www.iso.org/> Acessado em: 13 jan 2013.
- JADE. Java Agent Development Framework), 2013. [online] Disponível em <http://jade.tilab.com/> Acessado: 3 Jan 2013.
- JOHNSON R.K. Partnering with faculty to enhance scholarly communication. D-Lib Magazine, v. 8, n. 11, nov. 2002. Disponível em <http://www.dlib.org/dlib/november02/johnson/11johnson.html>>. Acessado em 3 Jan 2011.
- JSON. Introducing JSON. 2013. [online] Disponível em: <http://json.org/> Acessado em 05 Jan 2013.
- LEE, Ming-Che; TSAI, Kun Hua; WANG, Tzone I. A practical ontology query expansion algorithm for semantic-aware learning objects retrieval. In: Computers & Education. Volume 50, N. 4, 2008, p. 1240-1257. Disponível em: sciencedirect Acessado em: 23 nov 2011.
- LI S.; YANG Z.; LIU Q.; HUANG T. Research of web information retrieval based on metadata and OAI, *Granular Computing, 2008. GrC 2008. IEEE International Conference on*, vol., no., pp.383-386, 26-28 Aug. 2008 doi: 10.1109/GRC.2008.4664693
- LUME. Lume - Repository of the Federal University of Rio Grande do Sul.[online] Disponível em <http://www.lume.ufrgs.br/> Acessado em: 10 jan 2013.
- LYNCH, C.A. Institutional Repositories: essential infrastructure for scholarship in the digital age. ARL [Association Research Libraries], n. 226, feb. 2003, p 1-7. Disponível em <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>. Acessado: 3 jan. 2011

- MCGREAL, R. *Online Education Using Learning Objects*. London: Routledge, 2004.
- NASH, SS. *Learning Objects, Learning Object Repositories, and Learning Theory: Preliminary Best Practices for Online Courses*. *Interdisciplinary Journal of Knowledge and Learning Objects*, New York, 2005.
- NISO (2007) ANSI/NISO Z39.85 - The Dublin Core Metadata Element Set. National Information Standards Organization. Disponível em http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=52142 Acessado em: 10 Jun 2011.
- NISO. *Understanding Metadata*. [online] National Information Standards Organization. Press. ISBN: 1-880124-62-9. 2004 Disponível em <http://www.niso.org/publications/press/UnderstandingMetadata.pdf> Acessado em: 10 Jun 2011
- OAI-ORE. *Open Archives Initiative*. 2013 [online] Disponível em <http://www.openarchives.org/> Acessado em 05 jan 2013.
- OBAA. *Projeto Objetos de Aprendizagem Baseados em Agentes (Obaa)*. [online] Disponível em <http://www.portalobaa.org/obaac/padrao-obao/relatorios-tecnicos/RT-OBAA-01.pdf> Acessado: 10 jan 2013.
- PEREIRA, E.C.; BUFREM L.S. Fontes de informação especializada: uma prática de ensino-aprendizagem com pesquisa na Universidade Federal do Paraná. *Perspect. cienc. inf.*, Belo Horizonte, v. 7, n. 2, p. 197-206, jul./dez. 2002
- PRIMO, T.; LOH, S. Técnicas de Recomendação para usuários de Bibliotecas Digitais. in *Simpósio Brasileiro de Sistemas de Informação*. Curitiba, PR, novembro de 2006.
- PubMed. US National Library of Medicine: PubMed. Disponível em <http://www.ncbi.nlm.nih.gov/pubmed> Acessado em 05 jan 2013.
- RESNICK, P.; VARIAN, H. R. Recommender systems. *Commun. ACM*, New York, NY, USA, v.40, n.3, p.56-58, 1997.
- ROUSSET, M.C., REYNAUD, C. Knowledge representation for information integration, *Information Systems*, Volume 29, Issue 1, March 2004, Pages 3-22, ISSN 0306-4379
- SCHAFER, J.B., Konstan, J.A., and Riedl, J. "E-Commerce Recommender Applications," *Data Mining and Knowledge Discovery*, vol 5 nos.1/2, 2001, pp 115-152. Disponível em: <http://www.grouplens.org/papers/pdf/ECRA.pdf> Acessado em 15 jun 2012.
- SCORM. SCORM 2004 Overview for Developers. [online] Disponível em <http://scorm.com/scorm-explained/technical-scorm/scorm-2004-overview-for-developers/> Acessado em: 3 Jan 2013.
- SOLOMON, Délcio V. *Como fazer uma monografia*. 11. Ed. São Paulo: Martins Fontes, 2008.
- TARRANT D., O'STEEN B., BRODY T., HITCHCOCK S., JEFFERIES N. CARR L. (2009). Using OAI-ORE to Transform Digital Repositories into Interoperable Storage and Services Applications. [online] In: *Journal*. Issue 6. March 2009. Disponível em <http://journal.code4lib.org/articles/1062> Acessado: 14 Jul 2011.
- UnA-SUS. Repositório da UnA-SUS. Disponível em <http://repositorio.unasus.ufsc.br/> Acessado em 05 jan 2013.
- VIAN, J. ; CAMPOS, Ronaldo ; PALOMINO, C. E. G. ; SILVEIRA, R. A. . A Multiagent Model for Searching Learning Objects in Heterogeneous Set of Repositories. In: 11th IEEE International Conference on Advanced Learning Technologies, 2011, Athens. proceedings. New York: IEEE, 2011.
- VIAN, J. ; SILVEIRA, R. A.. Sistema Multiagente para Indexação e Recuperação de Objetos de Aprendizagem. In: XXI Simpósio Brasileiro de Informática na Educação,

- 2010, João Pessoa. Anais do XXI Simpósio Brasileiro de Informática na Educação. Porto Alegre: SBC, 2010.
- VICARI, Rosa Maria; RIBEIRO, Alexandre ; SILVA, Júlia Marques Carvalho da; SANTOS, Elder Rizzon; PRIMO, Tiago; BEZ, Marta. Brazilian Proposal for Agent-Based Learning Objects Metadata Standard – OBAA. IN: Metadata and Semantic Research. Communications in Computer and Information Science Volume 108, 2010, pp 300-311
- WILEY, D. Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley (Ed.), The instructional use of learning objects [Online]. Disponível em: <http://reusability.org/read/chapters/wiley.doc> [Accessed: 10 jul 2011].
- WOOLDRIDGE, M. An Introduction to Multiagent Systems. England: John Wiley, 2002.
- XMLRPC. XML-RPC Home Page. [online] Disponível em <http://www.xmlrpc.com/> Acessado em 10 Jan 2013.