



UNIFACS
UNIVERSIDADE SALVADOR
LAUREATE INTERNATIONAL UNIVERSITIES®

UNIFACS UNIVERSIDADE SALVADOR
MESTRADO PROFISSIONAL EM SISTEMAS E COMPUTAÇÃO

MÁRCIO MELO DE OLIVEIRA

**PROCESSO DE CONSTRUÇÃO DA ONTOLOGIA DE REGISTROS ACADÊMICOS
DO MODELO EDUCACIONAL BRASILEIRO – ONTO-RA**

Salvador
2014

MÁRCIO MELO DE OLIVEIRA

**PROCESSO DE CONSTRUÇÃO DA ONTOLOGIA DE REGISTROS ACADÊMICOS
DO MODELO EDUCACIONAL BRASILEIRO – ONTO-RA**

Dissertação apresentada ao Programa de Pós-Graduação em Sistemas e Computação de UNIFACS Universidade Salvador, Laureate International Universities como requisito parcial à obtenção do título de Mestre.

Orientador: Prof. Dr. Paulo Caetano da Silva.

Salvador
2014

FICHA CATALOGRÁFICA

Elaborada pelo Sistema de Bibliotecas da UNIFACS Universidade Salvador, Laureate International Universities)

Oliveira, Márcio Melo de

Processo de construção da ontologia de registros acadêmicos do modelo educacional brasileiro – ONTO-RA./ Márcio Melo de Oliveira.- Salvador: UNIFACS, 2014.

123 f. : il.

Dissertação Programa de Pós-Graduação em Sistemas e Computação de UNIFACS Universidade Salvador, Laureate International Universities como requisito parcial à obtenção do título de Mestre.

Orientador: Prof. Dr. Paulo Caetano da Silva.

1. Ontologias. 2. Engenharia de Ontologias. 3. Interoperabilidade. I. Silva, Paulo Caetano da, orient. II. Título.

CDD: 004.22

MÁRCIO MELO DE OLIVEIRA

PROCESSO DE CONSTRUÇÃO DA ONTOLOGIA DE REGISTROS ACADÊMICOS
DO MODELO EDUCACIONAL BRASILEIRO – ONTO-RA

Dissertação de Mestrado aprovada como requisito parcial para obtenção do grau de mestre em Sistemas e Computação, UNIFACS Universidade Salvador, Laureate International Universities pela seguinte banca examinadora:

Paulo Caetano da Silva – Orientador _____
Ph.D. in Computer Science, Universidade Federal de Pernambuco - UFPE
UNIFACS Universidade Salvador, Laureate International Universities

Glauco de Figueiredo Carneiro _____
Doutor em Ciência da Computação pela Universidade Federal da Bahia - Ufba
UNIFACS Universidade Salvador, Laureate International Universities

Vaninha Vieira dos Santos _____
Doutora em Ciências da Computação pela Universidade Federal de Pernambuco,
UFPE
Universidade Federal da Bahia - UFBA

Salvador, 14 de novembro de 2014.

Dedico este trabalho à memória da minha saudosa, querida e inesquecível mãe que através do seu exemplo me ensinou que o triunfo só vem após muita luta.

AGRADECIMENTOS

Tarefa árdua a de lembrar de tantos parentes, amigos, colegas de trabalho, professores, etc que de maneira direta ou indireta contribuíram para a concretização desta etapa tão importante da minha vida sem correr o risco de deixar de citar alguém. Muitos foram os que contribuíram e todos eles tiveram o seu grau de importância, entretanto, seria impossível alcançar este patamar sem a ajuda, companheirismo, incentivo, compreensão das pessoas que citarei a seguir.

Meus queridos amigos da “velha guarda” da DGTI, Allan, Edna, Thiago, Lázaro, Marcelo e Rogério que muitas vezes acreditaram em mim muito mais do que eu próprio e compreenderam cada momento de ausência e aflição e mais do que isso me incentivaram em diversos momentos em que o horizonte parecia não querer se mostrar para mim. Vocês foram fundamentais.

Ao grande Professor Paulo Caetano que com toda a simplicidade consegue conduzir seus orientandos a um Porto seguro mesmo que por vezes precisemos navegar por águas turbulentas.

Meus pais pelo exemplo de luta e de valorização da educação de qualidade como o maior legado que poderiam deixar para mim. Sem o esforço que vocês fizeram no passado eu jamais teria chegado aonde cheguei. Vocês são os grandes artífices desta conquista.

À minha amada esposa Tati que certamente foi a pessoa que mais se sacrificou por mim durante todo este período e que além de tudo foi cúmplice das minhas angústias, incertezas e aflições sempre me transmitindo uma palavra de conforto e incentivo nos momentos mais importantes. Prometo recompensar em dobro todas as faltas em eventos sociais e todos os momentos que deixei de compartilhar da sua presença. Te amo !

Aos meus pequeninos amores Tainá e Isabel. Fontes inesgotáveis de alegria e inspiração. Agora papai é todo de vocês!

RESUMO

A autonomia das instituições federais de ensino no Brasil dificultou a criação de um sistema centralizado que fosse responsável pela gestão acadêmica das instituições da alçada administrativa do Ministério da Educação - MEC. Este cenário propiciou que cada instituição buscasse desenvolver sua própria solução de sistema de gestão acadêmica e que o MEC utilizasse diversos sistemas de coleta de dados para obter informações sobre as ações das instituições. Tudo isso ocorreu sem que houvesse alguma iniciativa ou recomendação para a definição de diretrizes para a construção destes sistemas ou algo que gerasse uniformidade semântica no domínio dos registros acadêmicos brasileiros. Esta situação gerou sistemas heterogêneos e dados estatísticos pouco confiáveis. Com base nos leiautes de migração para os sistemas de coleta de dados, na Lei de Diretrizes e Bases da Educação – LDB e em outros artefatos como o Termo de Acordo de Metas e Compromissos – TAMC, é proposta nessa dissertação a ONTOLOGIA DE REGISTROS ACADÊMICOS DO MODELO BRASILEIRO DE EDCUAÇÃO – ONTOR-RA. Esta ontologia utilizou estes artefatos, classificados na Engenharia de Ontologias como Recurso não Ontológico, visando gerar a uniformização semântica necessária para que os sistemas de coleta de dados e de gestão acadêmica possam ser ajustados e novos sistemas a serem construídos se baseiem neste modelo. A construção da ONTO-RA, contou com a definição de um método de avaliação das metodologias de construção de ontologias disponíveis, que resultou na escolha da Metodologia NEON. Essa metodologia foi aplicada com pequenas adaptações para a construção da ontologia proposta. Este trabalho ainda apresentou uma aplicação da ontologia que visa evidenciar a sua eficácia como base para o ajuste dos sistemas a partir da uniformidade semântica obtida. Nesta aplicação, para cada requisito da ONTO-RA foi analisado o atendimento do mesmo em relação aos sistemas de coleta de dados SISTEC, EDUCANCENSO, e CENSO DO ENSINO SUPERIOR – CENSUP. Foram analisados também os sistemas SICAD e SIGA, que são sistemas de gestão acadêmica em funcionamento atualmente no Instituto Federal de Educação da Bahia IFBA –. Este estudo revelou que diversos indicadores acadêmicos importantes devem estar sendo gerados de forma incorreta em função das diversas inadequações identificadas, visto que o percentual de requisitos não atendidos variou entre 2% e 37% nos sistemas analisados.

Palavras-Chave: Ontologias. Registros Acadêmicos. Engenharia de Ontologias. Processo de Construção de Ontologias. Interoperabilidade

ABSTRACT

The autonomy of federal educational institutions in Brazil hindered the creation of a centralized system that was responsible for the academic management of the administrative purview of the Ministry of Education institutions - MEC. This scenario provided that each institution should seek to develop their own solution of academic management system and that the MEC would use different data collection systems to obtain information about the actions of the institutions. All this occurred without there being any motion or recommendation for the definition of guidelines for the construction of these systems or something that generates semantic uniformity in the field of Brazilian academic records. This has created heterogeneous systems and unreliable statistics. Based on the layouts for systems migration data collection, the Law of Guidelines and Bases of Education - LDB and other artifacts such as the Statement of Goals and Commitments Agreement - TAMC is proposed in this dissertation ONTOLOGY OF ACADEMIC RECORDS BRAZILIAN MODEL EDCUAÇÃO - ONTOR-RA. This ontology used these artifacts, ranked in Engineering Ontology as a Resource Ontological not, in order to generate semantic uniformity required for systems of data collection and academic management can be adjusted and new systems to be built is based on this model. The construction of the ONTO-RA, had the definition of a method of evaluation of methodologies for building ontologies available, which resulted in the choice of methodology NEON. This methodology was applied with minor adjustments to the construction of the proposed ontology. This paper also presented an application of the ontology which aims to highlight its effectiveness as a basis for setting the systems from the semantic uniformity obtained. In this application, for each requirement of ONTO-RA was analyzed care of the same for the collection SISTEC, EDUCANCENSO data systems, and CENSUS OF HIGHER EDUCATION - CENSUP. The SICAD and SIGA systems, which are academic management systems currently in operation at the Federal Institute of Education Bahia - IFBA were also analyzed. This study revealed that several important academic indicators are being generated incorrectly due to the various inadequacies identified, since the percentage of requirements not met ranged between 2% and 37% in the analyzed systems.

Key-words: Ontologies. Academic Records. Ontology Engineering. Building Ontology Process. Interoperability.

LISTA DE QUADROS

Quadro 1 - Lista de Sistemas de Coletas de Dados do MEC.....	15
Quadro 2 - Metodologias de Construção de Ontologias em Ordem Cronológica.....	34
Quadro 3 - Valores desejáveis para os critérios de avaliação das metodologias.....	37
Quadro 4 - Avaliação das metodologias de construção de ontologias conforme critérios escolhidos.....	38
Quadro 5 - OSRD da ONTO-RA	63
Quadro 6 - Lista de questões de competência por grupo.....	67
Quadro 7 - Termos extraídos das perguntas das questões de competência	72
Quadro 8 - Termos extraídos das respostas das questões de competência.....	73
Quadro 9 - Tabela de avaliação dos candidatos a NORs	80
Quadro 10 - Análise da QC-01	92
Quadro 11 - Análise da QC-02	93
Quadro 12 - Análise da QC-03	94
Quadro 13 - Análise da QC 04	94
Quadro 14 - Análise da QC-05	95
Quadro 15 - Análise da QC-06	96
Quadro 16 - Análise da QC-07	97
Quadro 17 - Análise da QC-08	98
Quadro 18 - Análise da QC-09	99
Quadro 19 - Análise da QC - 10.....	100
Quadro 20 - Análise da QC-11	101
Quadro 21 - Análise da QC-12	102
Quadro 22 - Análise da QC-13, QC-14 e QC-15.....	103
Quadro 23 - Análise da QC-16, QC-17 e QC-18.....	104
Quadro 24 - Análise das QCs do grupo pessoa.....	105
Quadro 25 - Análise da QC-26	106
Quadro 26 - Análise das QCs do grupo situação acadêmica	106
Quadro 27 – Análise das QCs do grupo vinculo institucional	108
Quadro 28 – Análise das QCs do grupo vinculo institucional para docentes	109
Quadro 29 - Percentual das QCs não atendidas por sistema analisado	110

LISTA DE FIGURAS

Figura 1 - Processo de Escolha da Metodologia para Construção da ONTO-RA	32
Figura 2 - Atividades por Fase do Ciclo de Vida da Metodologia NEON.....	41
Figura 3 - Cenários da Metodologia NEON.....	43
Figura 4 - Estrutura dos processos, atividades e tarefas da metodologia NEON.....	45
Figura 5 - Processos e atividades executados	46
Figura 6 - Tarefas da geração da OSRD.....	49
Figura 7 - Fluxo de atividades e tarefas de reuso de NORs.....	52
Figura 8 - Fluxo de atividades e tarefas do processo de reengenharia de NORs	54
Figura 9 - Categorização dos modelos de dados para esquemas de classificação ..	58
Figura 10 - Fluxo geral de atividades para geração da ONTO-RA.....	61
Figura 11 - Fluxo de Atividades Executadas no Levantamento de Requisitos da ONTO-RA.....	65
Figura 12 - Classe Curso e seus Relacionamentos ou Propriedades	84
Figura 13 - Estrutura Hierárquica dos Cursos em ONTO-RA.....	85
Figura 14 - Classes da ONTO-RA que se relacionam com a classe curso	85
Figura 15 - Classes da ONTO-RA que se relacionam com a classe aluno	86

LISTA DE ABREVIATURAS E SIGLAS

CEFET	Centro Federal de Educação Ciência e Tecnologia
IF	Instituto Federal de Educação, Ciência e Tecnologia
IFBA	Instituto Federal de Educação, Ciência e Tecnologia da Bahia
MEC	Ministério da Educação
NORs	Recursos Não Ontológicos
OWL	WEB Ontology Language
PR-NOR	Padrão de Reengenharia de Recursos Não Ontológicos
RDF	Resource Description Framework
SETEC	Secretaria de Educação Profissional e Tecnológica
TAMC	Termo de Acordo de Metas e Compromissos
TCU	Tribunal de Contas da União
W3C	World Wide Web Consortium
ONTO-RA	Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro
CENSUP	Censo do Ensino Superior
SISTEC	Sistema Nacional de Informações da Educação Profissional e Tecnológica
XML	Extensible Markup Language
MCC	Microeletronic and Computer Technology
AIAI	Instituto de Aplicações em Inteligência Artificial
TOVE	Toronto Virtual Enterprise
KIF	Knowledge Interchange Format
SABIO	Systematic Approach for Building Ontologies
UML	Unified Modeling Language
ODM	Ontology Definition Metamodel
OMG	Object Managemnt Group
MDA	Model Driven Architeture
OSRD	Documento de Especificação de Requisitos da Ontologia
IDEB	Índice de Desenvolvimento da Educação Básica
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
FUNDEB	Fundo de Manutenção e Desenvolvimento da Educação Básica

PDE	Programa de Desenvolvimento Educacional
LDB	Lei de Diretrizes e Bases da Educação
IES	Instituições de Ensino Superior
CPC	Cálculo Preliminar de Curso
IGC	Índice Geral de Cursos
SICAD	Sistema de Controle Acadêmico
SIGA-EPCT	Sistema Integrado de Gestão Acadêmica da Rede de Educação Profissional, Científica e Tecnológica

SUMÁRIO

1 INTRODUÇÃO	14
1.1 CONTEXTUALIZAÇÃO	14
1.2 JUSTIFICATIVA	16
1.3 MOTIVAÇÃO	17
1.4 OBJETIVOS	17
1.5 METODOLOGIA	18
1.6 ESTRUTURA DA DISSERTAÇÃO	19
2 FUNDAMENTAÇÃO TEÓRICA	21
2.1 ONTOLOGIAS: DEFINIÇÕES	21
2.1.1 O que é ontologia	21
2.1.2 Tipos de Ontologia	22
2.1.3 Componentes de uma Ontologia	23
2.2 PRINCIPAIS APLICAÇÕES E USOS DE ONTOLOGIAS	23
2.3 RECURSO NÃO ONTOLÓGICO – NOR	25
2.4 PROCESSO DE TRANSFORMAÇÃO DE RECURSOS NÃO ONTOLÓGICOS	25
2.5 CONSIDERAÇÕES FINAIS	26
3 REVISÃO BIBLIOGRÁFICA	27
3.1 PROCESSO DE PESQUISA	27
3.2 ONTOLOGIAS COMO SOLUÇÃO PARA INTEROPERABILIDADE	28
3.3 ONTOLOGIAS PARA O DOMÍNIO EDUCACIONAL	29
3.4 CONSIDERAÇÕES FINAIS	29
4 PROCESSO DE ESCOLHA DA METODOLOGIA PARA CONSTRUÇÃO DA ONTO-RA	31
4.1 VISÃO GERAL DO PROCESSO	31
4.2 METODOLOGIAS DE DESENVOLVIMENTO DE ONTOLOGIAS	33
4.3 CRITÉRIOS DE AVALIAÇÃO	35
4.4 MONTAGEM DA TABELA DE AVALIAÇÃO E ESCOLHA DA METODOLOGIA	37
4.5 METODOLOGIA NEON	39
4.6 CENÁRIO DOIS DA METODOLOGIA NEON: REUTILIZAÇÃO E REENGENHARIA DE RECURSOS NÃO ONTOLÓGICOS	44
4.6.1 Especificação dos Requisitos	47
4.6.2 Reuso de Recursos não Ontológicos	50
4.6.3 Reengenharia de Recursos não Ontológicos	53
4.7 CONSIDERAÇÕES FINAIS	59
5 CONSTRUÇÃO DA ONTOLOGIA DE REGISTROS ACADÊMICOS DO MODELO EDUCACIONAL BRASILEIRO – ONTO-RA	60
5.1 VISÃO GERAL DO PROCESSO DE CONSTRUÇÃO	60
5.2 ESTUDO DE VIABILIDADE E AQUISIÇÃO DO CONHECIMENTO	61
5.3 ESPECIFICAÇÃO DA ONTOLOGIA	62
5.3.1 Levantamento dos Requisitos em ONTO-RA	64
5.3.2 Extração da Terminologia e Frequência	71
5.4 BUSCA POR RECURSOS NÃO ONTOLÓGICOS PARA A ONTO-RA	78
5.5 AVALIAÇÃO DO CONJUNTO DE CANDIDATOS A RECURSOS NÃO ONTOLÓGICOS	79

5.6 SELEÇÃO DO RECURSO NÃO ONTOLÓGICO MAIS APROPRIADO.....	80
5.7 TRANSFORMAÇÃO DE RECURSO NÃO ONTOLÓGICO EM ONTO-RA.....	82
5.8 ENGENHARIA <i>FORWARD</i> DE ONTOLOGIA EM ONTO-RA.....	83
5.9 CONSIDERAÇÕES FINAIS	86
6 AVALIAÇÃO DA ADEÇÃO DOS SISTEMAS EM RELAÇÃO À ONTO-RA	88
6.1 SISTEMAS DE COLETA DO MEC	88
6.2 SISTEMAS DE GESTÃO ACADÊMICA DO IFBA.....	90
6.3 ANÁLISE POR QUESTÃO DE COMPETÊNCIA.....	91
6.3.1 Requisitos do Grupo Vínculo Institucional para docentes.....	108
6.4 CONSIDERAÇÕES FINAIS	109
7 CONCLUSÃO	111
7.1 CONSIDERAÇÕES FINAIS	111
7.2 PRINCIPAIS CONTRIBUIÇÕES	112
7.3 TRABALHOS FUTUROS	113
7.4 ARTIGOS PUBLICADOS	114
REFERÊNCIAS.....	115
ANEXO A - Tabelas do esquema de classificação resultante da fusão dos candidatos a recursos não ontológicos.....	119
ANEXO B - Modelo Conceitual	122

1 INTRODUÇÃO

Este Capítulo tem como objetivo apresentar o contexto no qual foi idealizada a construção da Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro – ONTO-RA, assim como as justificativas, motivação e objetivos deste trabalho. Também descreve a sua estrutura e a metodologia utilizada na realização deste trabalho.

1.1 CONTEXTUALIZAÇÃO

A educação profissional e tecnológica faz parte do cotidiano dos brasileiros há mais de cem anos e durante este período passou por momentos de maior e menor evidência, oscilando em importância e volume de recursos aplicados ao seu desenvolvimento de acordo com a visão estratégica dos governantes. Desde o ano de 2002 os investimentos neste ramo da educação têm crescido e a partir de 2008 se intensificaram com a transformação da maioria dos antigos Centros Federais de Educação Tecnológica (CEFETs) em Institutos Federais de Educação, Ciência e Tecnologia (IFs).

As políticas públicas implantadas fizeram com que o número de alunos e instituições crescesse. Até o ano de 2002, existiam 140 unidades de ensino federais de educação profissional e tecnológica. Segundo dados do Ministério da Educação - MEC (<http://redefederal.mec.gov.br/expansao-da-rede-federal>). Atualmente, existem 354 unidades que ofertam aproximadamente 400 mil vagas, com previsão de crescimento para 562 unidades com capacidade de oferta de 600 mil vagas até o final do ano de 2014. Com este crescimento, aumentou também a necessidade por do MEC gerenciar, medir e controlar os resultados gerados. Os IFs, assim como as Universidades Federais são autarquias com autonomia administrativa, portanto, livres para estabelecer as regras que norteiam as suas relações acadêmicas. Este cenário dificultou que no âmbito do MEC existisse um sistema de informação centralizado capaz de realizar a gestão acadêmica (e.g. Matrícula, Movimentação, Emissão de Documentos, Registro de Desempenho de Avaliações, etc.). Ao contrário, a realidade é que cada Instituição Federal de Ensino, inclusive as universidades, tenha soluções próprias, a partir de desenvolvimento próprio ou buscando por meio do mercado de softwares de gestão acadêmica. Isto criou um

ambiente heterogêneo entre as instituições na alçada administrativa do MEC. A falta de informações centralizadas e o ambiente heterogêneo dos registros acadêmicos entre as instituições fez com que o MEC recorresse à utilização de sistemas de coleta de dados, em que as instituições informam dados dos alunos, cursos e outros. Com este intuito diversos sistemas de informação foram desenvolvidos e os dados são periodicamente fornecidos pelas instituições federais de ensino, seja através de informações de dados estatísticos gerenciais, seja através de dados primários, i.e. dados pessoais de docentes e discentes, além de informações sobre cursos, turmas e da própria instituição. A carga de dados nestes sistemas é feita de forma manual ou automática a partir de modelo de dados disponibilizado pela secretaria do MEC responsável pela coleta. O Quadro 1 lista os sistemas de coleta utilizados pelo MEC no âmbito acadêmico.

Quadro 1 - Lista de Sistemas de Coletas de Dados do MEC

Sistema de Coleta	Nível de Ensino
Educacenso	Educação Básica
Censo do Ensino Superior	Educação Superior
SISTEC	Educação Básica, Educação Superior e Formação Inicial e Continuada
PINGIFES	Educação Superior

Os sistemas de coleta de dados citados no Quadro 1 são utilizados para a geração de indicadores oficiais da educação brasileira e como base para o acompanhamento do Termo de Acordo de Metas e Compromissos - TAMC, pactuado entre o MEC e todos os IFs, no qual estão estabelecidas algumas metas a serem atingidas por parte das instituições e algumas contrapartidas que deverão ser entregues pelo MEC. As metas constantes do TAMC referem-se, principalmente, a indicadores de desempenho e produtividade acadêmicos. Outra finalidade a que se destinam os sistemas de coleta é a definição da matriz orçamentária dos IFs. Tal matriz também é definida a partir de indicadores de produtividade acadêmica.

O cenário heterogêneo das Instituições de Ensino, aliado à falta de uniformidade semântica para registros acadêmicos, têm gerado distorções importantes nos levantamentos de informações realizados pelo MEC e até mesmo por órgãos de controle como o Tribunal de Contas da União - TCU. Nas coletas

realizadas nos sistemas responsáveis pelos dados da educação básica, SISTEC E Educacenso, no ano de 2013, o número de alunos matriculados identificados em cada um dos sistemas, apresentou uma diferença da ordem de 26%, conforme relatório da Secretaria de Educação Profissional e Tecnológica – SETEC do MEC. Tal diferença pode ser explicada por alguns fatores que vão desde a falta de uniformidade entre os calendários acadêmicos até a indefinição de termos importantes no contexto dos sistemas de coleta, como por exemplo, aluno efetivo, turma, turma de ingresso, evasão, etc.

1.2 JUSTIFICATIVA

Apesar da existência de legislação específica que estabelece limites para todos os níveis de ensino no Brasil, como citado anteriormente, as Instituições Federais de Ensino possuem autonomia administrativa, o que implica liberdade, inclusive para gestão acadêmica dentro dos limites legais. Entretanto, embora o MEC possua um *Thesaurus*, não há por parte do órgão qualquer iniciativa que objetive padronizar os conceitos utilizados na educação brasileira, a fim de que os sistemas acadêmicos utilizados pelas instituições, assim como os sistemas de coleta projetados pelo próprio MEC pudessem manter uma unidade semântica capaz de impedir algumas distorções geradas por este cenário.

Diante deste quadro, existe a necessidade de um ambiente de unidade semântica para o Domínio dos Registros Acadêmicos do Modelo Educacional Brasileiro a partir dos dados como Thesaurus Brasileiro da Educação, modelo de dados dos sistemas de coleta do MEC, premissas estabelecidas no TAMC e outros artefatos que possam ser identificados. Ademais, estabelecer um mapeamento semântico desse domínio servirá de base para o desenvolvimento de novos sistemas acadêmicos, permitirá que ajustes sejam realizados nos sistemas acadêmicos já em execução, servirá como base para a manutenção dos sistemas de coleta do MEC e até mesmo possibilitará que no futuro possa ser desenvolvida e proposta uma linguagem de integração de dados e informações acadêmicas.

1.3 MOTIVAÇÃO

A principal motivação para a realização deste trabalho é promover um ambiente de unidade semântica no âmbito do domínio dos registros acadêmicos do modelo brasileiro de educação. Visto que, no levantamento bibliográfico realizado não foram encontradas iniciativas que visassem trazer uniformidade semântica para o ambiente da educação brasileira.

Para alcançar essa uniformidade semântica, é proposta a construção de uma ontologia do domínio dos registros acadêmicos para o modelo brasileiro de educação que poderá possibilitar que os sistemas de coletas de informações do MEC sejam aperfeiçoados, assim como os sistemas acadêmicos utilizados pelas instituições federais de ensino. Tais ajustes trarão um ganho de qualidade às informações coletadas pelo MEC, possibilitando a geração de indicadores mais confiáveis, o que permitiria a elaboração de matrizes orçamentárias mais justas, controle mais efetivo da produtividade acadêmica, além da possibilidade de realização de um melhor planejamento das políticas públicas na área da educação.

1.4 OBJETIVOS

Esta dissertação tem como principal objetivo a criação de um ambiente de uniformidade semântica para o domínio dos registros acadêmicos do modelo educacional brasileiro, a partir da utilização das definições do *Thesaurus* brasileiro da educação e do modelo de dados dos sistemas de coleta do MEC. Para isto é proposta uma ontologia que possibilite que sejam feitos ajustes nos sistemas de coleta de dados e sistemas de gestão acadêmica existentes nas instituições federais de ensino, assim como, servir como base para o desenvolvimento de novos sistemas de gestão acadêmica.

Para atingir este objetivo, algumas metas foram estabelecidas. São elas:

- Pesquisar por trabalhos correlatos a fim de encontrar outras ontologias desenvolvidas que pudessem ser estendidas ou reutilizadas na construção da proposta de ontologia para o domínio dos registros acadêmicos do modelo educacional brasileiro;
- Caso sejam encontradas ontologias que possam ser reutilizadas ou estendidas, pesquisar na área de engenharia de ontologias métodos

ou recomendações de melhores práticas para a reutilização de recursos ontológicos para construção de uma ontologia;

- Caso não sejam encontradas ontologias para reutilização ou para se estendida, então, deverá ser realizada pesquisa na área de engenharia de ontologias as principais metodologias de construção de ontologias e realizar um estudo comparativo entre elas para identificar a metodologia mais adequada para a construção da proposta, levando em conta as especificidades do cenário descrito na Seção 1.1;
- Desenvolver a Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro – ONTO-RA utilizando a metodologia de construção identificada como mais adequada;
- Validar a proposta através da aplicação da ontologia nos sistemas SICAD e SIGA, que são sistemas de gestão acadêmica em funcionamento atualmente no IFBA – Instituto Federal de Educação da Bahia e no EDUCACENSO, CENSO DA EDUCAÇÃO SUPERIOR e SISTEC, que são sistemas de coleta de dados atualmente utilizados pelo MEC, para verificar o atendimento destes sistemas em relação aos requisitos implementados na ONTO-RA. A validação consistirá na avaliação dos sistemas citados em relação ao atendimento dos mesmos em relação às questões de competência elaboradas no processo de desenvolvimento da ONTO-RA. Como a ontologia implementada deverá ser capaz de responder as questões de competência elaboradas, desta forma, a avaliação proposta deverá evidenciar o quanto cada um dos sistemas avaliados está de acordo com o modelo ontológico especificado.

1.5 METODOLOGIA

A metodologia utilizada para o desenvolvimento deste trabalho inclui uma revisão bibliográfica na área de ontologias e suas aplicações e engenharia de ontologias, com ênfase para as metodologias de desenvolvimento de ontologias. Em seguida, foi realizada uma pesquisa em trabalhos correlatos com o intuito de identificar ontologias desenvolvidas no domínio dos registros acadêmicos, que

pudessem ser estendidas ou reaproveitadas na construção da proposta de ontologia para o domínio dos registros acadêmicos do modelo educacional brasileiro.

A seguir, com objetivo de identificar a metodologia de construção de ontologias mais indicada para a construção da proposta, foi necessário realizar um estudo comparativo entre as principais metodologias encontradas na literatura. Para tanto, os seguintes passos foram executados: Identificação dos trabalhos que abordassem o tema metodologia para a construção de ontologias, Análise e seleção de documentos, identificando as metodologias abordadas nos mesmos, Análise comparativa e crítica das metodologias pré-selecionadas a partir de critérios adequados ao cenário descrito e definição da metodologia a ser utilizada.

A partir da metodologia de construção de ontologias selecionada foi definida a proposta de ontologia com base nas atividades: especificação dos requisitos da ontologia; identificação, avaliação e seleção de recursos não ontológicos; engenharia reversa de recursos não ontológicos, transformação de recursos não ontológicos e geração da ontologia.

Por fim, a ontologia construída foi aplicada no Instituto Federal da Bahia – IFBA, através da análise de dois sistemas de gestão acadêmica utilizados nesta instituição federal de ensino, além da análise dos sistemas de coleta de dados EDUCACENSO, CENSUP e SISTEC. Nesta estudo, foram identificados os requisitos implementados em ONTO-RA que os sistemas analisados divergiam ou não implementavam, evidenciando assim, a necessidade de ajuste.

1.6 ESTRUTURA DA DISSERTAÇÃO

O restante deste trabalho está estruturado em mais cinco capítulos. No Capítulo 2 há a fundamentação teórica necessária para a compreensão do problema, assim como para a proposta de solução. Desta forma foram abordados os temas Ontologias e suas aplicações e Engenharia de Ontologias. No Capítulo 3 é feita uma revisão bibliográfica na qual são apresentados os principais trabalhos correlatos em que foram desenvolvidas ontologias para o domínio da educação e em que ontologias foram utilizadas para promover a interoperabilidade entre sistemas.

O Capítulo 4 apresenta o processo de escolha da metodologia de desenvolvimento de ontologia que será utilizada na construção da ONTO-RA. O capítulo 5 apresenta e discute o desenvolvimento da proposta de ontologia para o

domínio dos registros acadêmicos do modelo educacional brasileiro, desde a especificação até a conclusão da construção da ontologia. O Capítulo 6 relata a avaliação realizada para validar a efetividade da ontologia construída. Por fim, o Capítulo 6 apresenta as conclusões e indica as possibilidades de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados os principais fundamentos teóricos que deram sustentação ao processo de construção da Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro. Serão discutidos trabalhos que se dedicaram a definir ontologias e serão apresentadas as principais aplicações e usos das mesmas.

2.1 ONTOLOGIAS: DEFINIÇÕES

Nesta seção serão apresentadas as principais definições e conceitos sobre Ontologias, bem como suas aplicações e usos no universo da ciência da computação.

2.1.1 O que é ontologia

Ontologia é uma palavra de origem grega e é formada pela junção do prefixo ontos, cujo significado é ser, com o sufixo logos, cuja definição é palavra. Ontologias são utilizadas na filosofia desde o século XIX, de modo a fazer uma distinção entre o estudo do ser do estudo dos vários tipos de seres vivos existentes no mundo natural. Outra definição do termo, ainda no campo da filosofia, é focada no fornecimento de sistemas de categorização para a organização da realidade (GUARINO, 1998).

O conceito de ontologia em Ciência da Computação possui outro significado, adaptado da Filosofia, entretanto, mantendo relação com o significado original. Em Ciência da Computação, uma ontologia é um tipo de objeto de informação ou artefato computacional que incorpora algum tipo de visão do mundo em relação a um dado domínio (GUARINO et al., 2009). Ainda segundo (GUARINO et al., 2009), ontologias são meios de modelar formalmente a estrutura de um sistema, ou seja, as entidades e relações que são extraídas a partir de sua observação. (STUDER et al, 1998) afirma que uma ontologia é uma especificação formal e explícita de uma conceituação compartilhada. Uma conceituação é uma visão abstrata e simplificada do mundo que se deseja representar para algum propósito (GUARINO et al., 2009).

Assim, uma ontologia, além de representar os conceitos e relações extraídas do domínio específico de estudo, deverá especificar de maneira formal, os

conceitos. Portanto, uma ontologia é formalizada através de linguagens específicas, com sintaxe e semântica definidas, provendo assim, a restrição da interpretação dos seus conceitos através de axiomas. Desta maneira, uma ontologia pode ser considerada como um conjunto de axiomas que definem e restringem as possibilidades de interpretações dos conceitos representados (SCATALON et. al, 2010).

Segundo (MAEDCHE e STAAB, 2001) uma ontologia, pode ser descrita através de uma 5-tupla composta dos elementos primitivos de uma ontologia, i.e., conceitos ou classes, relacionamentos, hierarquia de conceitos ou classes, função que relaciona conceitos de forma não taxonômica (relacionamentos cujos tipos são diferentes de subclasse e, portanto, não especificados na hierarquia dos conceitos ou classes) e um conjunto de axiomas expressos em linguagem lógica. Este conjunto é responsável pelas restrições estabelecidas para cada conceito ou classe.

2.1.2 Tipos de Ontologia

Existem várias propostas de classificação de ontologias na literatura, entretanto, um sistema de classificação que utiliza a generalidade da ontologia como o critério principal para a classificação foi proposto por (GUARINO, 1998) além de ser adotado com mais frequência na literatura, possui uma maior adequação para o contexto deste trabalho. Este sistema de classificação organiza as ontologias em quatro tipos, são eles:

- Ontologias de nível superior, que descrevem conceitos muito genéricos, tais como espaço, tempo, e eventos. Estes conceitos são independentes de domínio e podem ser reutilizados na confecção de novas ontologias;
- Ontologias de domínio, que descrevem o vocabulário relativo a um domínio específico através da especialização de conceitos presentes na ontologia de nível superior;
- Ontologias de tarefas, que descrevem o vocabulário relativo a uma tarefa genérica ou atividade através da especialização de conceitos presentes na ontologia de nível superior;

- Ontologias de aplicação, que são as ontologias mais específicas. Conceitos em ontologias de aplicação correspondem, de maneira geral, a papéis desempenhados por entidades do domínio no desenvolvimento de alguma tarefa.

Outra forma de classificação de ontologias se refere à forma como as ontologias modelam o domínio. O objetivo desta forma de classificação é distinguir as ontologias que são basicamente uma taxonomia das ontologias que fazem uma modelagem do domínio de forma mais aprofundada, através de mais restrições sobre a semântica do domínio. Assim, as ontologias podem ser classificadas em:

- Ontologias Leves ou *LightWeight Ontologies*, que incluem conceitos, hierarquia dos conceitos, relacionamentos entre os conceitos e propriedades que descrevem os conceitos;
- Ontologias Completas ou *Heavyweight Ontologies* ou Ontologias, que adicionam axiomas e restrições para as ontologias leves.

2.1.3 Componentes de uma Ontologia

Uma ontologia pode ter diferentes formas, mas, necessariamente, inclui um vocabulário de termos e alguma especificação de seus significados. Os componentes básicos de uma ontologia são classes, propriedades e instâncias (USCHOLD e GRUNINGER, 1996). Segundo (HORRIDGE, 2009) instâncias são os objetos do domínio que está sendo representado pela ontologia, propriedades são relações binárias entre instâncias e classes são conjuntos que contêm instâncias e devem estar organizadas em uma hierarquia.

Uma classe é definida por meio de descrições matemáticas formais que declaram os requisitos de pertinência da classe. Tais requisitos delimitam a interpretação do conceito através de restrições de propriedades.

2.2 PRINCIPAIS APLICAÇÕES E USOS DE ONTOLOGIAS

Através das principais definições conceituais sobre ontologias, discutidas na Seção 2.1, pode-se perceber que o compartilhamento de informações e conhecimento acerca de um determinado domínio é o principal motivador para a

construção de ontologias. As ontologias possuem um arcabouço explícito para a conceitualização de alguma realidade (Noy, 2005).

Assim, atualmente uma das principais aplicações das ontologias, ocorre na área de Engenharia de Software, mais especificamente no processo de levantamento de requisitos. Neste processo, as ontologias podem ser utilizadas como base para definição de requisitos de sistemas de informação. Outro uso comum das ontologias, devido ao seu poder de representação do conhecimento, é o de servir como base para estabelecimento de interoperabilidade entre sistemas de informação.

Ontologias também estão sendo utilizadas como base para a organização dos conteúdos da WEB, possibilitando que os mesmos sejam processados por máquinas, facilitando assim a busca por tais conteúdos. Tal iniciativa é denominada Web Semântica, na qual cada recurso de informação deve ser representado de maneira a incorporar o seu significado, apresentando metadados que permitam que agentes de software processem os dados. Ontologias são consideradas a infraestrutura da Web Semântica. OWL, a linguagem padrão da W3C para desenvolvimento de ontologias, estende padrões já existentes da Web, como *Extensible Markup Language - XML* (BRAY, 1998). Há uma grande variedade de ferramentas disponíveis para a OWL, esta é uma das possíveis razões para a sua ampla adoção em meios acadêmicos e no mercado (Horridge e Bechhofer, 2009). Essas ferramentas têm suportado a criação e edição de ontologias OWL (Protégé³, Swoop⁴), a computação de suas vinculações (Hermit⁶, Pellet⁷) e o uso de ontologias em aplicações (Jena Ontology API⁸ e OWLAPI⁹).

2.3 RECURSO NÃO ONTOLÓGICO – NOR

Segundo (SUÁREZ-FIGUEROA, 2008), recursos não ontológicos – NORS são conhecimentos existentes em recursos cuja semântica ainda não foi formalizada por meio de uma ontologia. Existem diversos recursos não ontológicos que armazenam conhecimento sobre um domínio em particular, estes recursos podem ser apresentados em forma de textos livres, páginas web, padrões, catálogos, classificações, thesaurus, folksonomias, entre outros. (ANGELETOU, 2009). Existem diversas formas de classificar recursos não ontológicos. (Maedche e Staab, 2001) e (SABOU, 2007) classificam como estruturados, não estruturados (e.g. texto livre), semi-estruturados (e.g. folksonomias) e estruturados (e.g. base de dados), (HODGE, 2000) propõe características como estrutura, complexidade, relacionamentos entre os termos para classificar os recursos não ontológicos. Não há, portanto, um consenso em relação à forma de classificação dos recursos não ontológicos (SUÁREZ-FIGUEROA, 2008).

2.4 PROCESSO DE TRANSFORMAÇÃO DE RECURSOS NÃO ONTOLÓGICOS

(VILLAZÓN-TERRAZAS, 2009) faz uma revisão dos principais trabalhos que abordam métodos e ferramentas utilizadas para realizar a reengenharia de recursos não ontológicos. Neste levantamento, foram identificadas as principais características deste processo em diversas abordagens. Segundo (VILLAZÓN-TERRAZAS, 2009) o processo de transformação de recursos não ontológicos em ontologia apresenta características gerais e específicas. Uma das características gerais é a forma como ocorre a transformação, que pode ocorrer de maneira completa, conhecida como *one step transformation*, ou de maneira incremental, a *incremental transformation*, que converte componentes específicos dos recursos em ontologias. Outra característica é a abordagem de transformação, a qual pode (1) transformar o esquema do recurso em ontologia e o conteúdo do recurso em instâncias; (2) transformar o conteúdo do recurso em esquema de ontologia ou (3) transformar o conteúdo do recurso em instâncias de uma ontologia já existente. Além disso, o processo de transformação poderá ser automático ou manual. Por fim, ainda segundo (VILLAZÓN-TERRAZAS, 2009) o processo de transformação deverá

controlar a semântica dos relacionamentos e executar uma completa conversão dos recursos, o que significa dizer que qualquer consulta executada no recurso não ontológico também poderá ser executada na ontologia gerada e durante o processo de transformação poderão ser utilizados recursos adicionais.

Como características gerais, (VILLAZÓN-TERRAZAS, 2009) define que os aspectos de transformação devem ser contemplados nos níveis sintático e semântico. No nível sintático trata da habilidade de estruturar a representação em sentenças estruturadas, fórmulas ou afirmações. Inclui ainda, a transformação de definições de acordo com as gramáticas da fonte e formatos alvo. Já o nível semântico, objetiva construir o significado proposicional da representação.

A ontologia resultante do processo de transformação deverá ter como componentes classes, atributos, relacionamentos e instâncias, e ser construída utilizando a linguagens OWL.

2.5 CONSIDERAÇÕES FINAIS

Neste capítulo foram discutidos os conceitos básicos de ontologias, como definições, tipos e componentes. Foram apresentadas também as principais aplicações e usos das ontologias no ramo da ciência da computação. Evidenciando assim que o uso de ontologias para compartilhamento de informações, assim como interoperabilidade de sistemas a partir de uniformidade semântica, tem sido discutido e validado por meio de trabalhos acadêmicos. No próximo Capítulo será feita uma discussão, através de uma revisão bibliográfica a respeito da Engenharia de Ontologias, com foco específico nas metodologias de construção de ontologias.

3 REVISÃO BIBLIOGRÁFICA

Neste capítulo será apresentada a pesquisa em fontes bibliográficas que possibilitou a seleção dos trabalhos correlatos, que além de servir para validar o uso de ontologias para a resolução do problema descrito no Capítulo 1, identificar a possibilidade, ou não, de reutilização de outra ontologia para servir de base para a construção da ONTO-RA, irá posicionar este trabalho perante aos demais trabalhos correlatos produzidos. Em seguida, serão discutidos os principais trabalhos correlatos que abordem o uso de ontologias para resolução de problemas de interoperabilidade entre sistemas de informação a partir da uniformidade semântica e por fim serão discutidos os trabalhos que apresentem ontologias para o domínio da educação.

3.1 PROCESSO DE PESQUISA

O processo de pesquisa bibliográfica com o intuito de identificar trabalhos correlatos foi dividido em duas áreas de busca específicas. A primeira delas visou identificar trabalhos recentes que tenham como foco a utilização de ontologias para a resolução de problemas de interoperabilidade de sistemas de informação, baseado em uniformidade semântica. Assim foram estabelecidas as seguintes *strings* de busca :

1. ("ontologias") and ("interoperabilidade entre sistemas") and ("semantica");
2. ("ontologias") and ("interoperabilidade") and ("sistemas") and ("semantica");
3. ("ontology") and ("systems interoperability") and ("semantic")

A segunda área de busca tinha como foco a construção de ontologias que estivessem relacionadas com o domínio da educação. Para tanto, foram utilizadas as seguintes *strings* de busca :

1. ("Ontologias") and ("Educação")
2. ("ontologia") and ("registros acadêmicos")
3. ("Ontologias") and ("Domínio da Educação")
4. ("Ontology") and ("Education")

Estas *strings* foram aplicadas nas fontes de pesquisa *Google* e *Google Scholar*. Os trabalhos extraídos da aplicação das strings informadas acima passaram por um processo de avaliação e seleção visando identificar os trabalhos de maior representatividade e correlação em relação a este trabalho. Assim, os principais trabalhos correlatos identificados nas duas áreas pesquisadas foram discutidos nas Seções 3.2 e 3.3.

3.2 ONTOLOGIAS COMO SOLUÇÃO PARA INTEROPERABILIDADE

A seleção dos trabalhos que retornaram após a aplicação das strings de busca resultou na indicação de três principais trabalhos que tratam do tema da interoperabilidade entre sistemas a partir da uniformidade semântica. O trabalho realizado por (Panetto et. al, 2012) propõe uma abordagem para facilitar a interoperabilidade de sistemas em um ambiente de indústrias. O Trabalho toma como base a afirmação que um modelo ontológico de um produto pode ser considerado como um facilitador para promover a interoperabilidade entre o produto e o software aplicativo que compartilha informações durante o ciclo de vida do produto físico. Todos os dados e conceitos técnicos foram formalizados para a definição de uma Ontologia que deverá ser incorporada a cada produto produzido pela fábrica, tornando-o assim passível de interoperabilidade com aplicações, a partir da minimização das perdas semânticas.

Já Farinelli et. al. (2013) estabelece um paralelo entre ontologias estabelece um paralelo entre a interoperabilidade de sistemas e a utilização de ontologias com esta finalidade, com um foco nas iniciativas do governo brasileiro neste sentido, esclarecendo o papel as ontologias nesse contexto, uma vez que ontologias vem sendo recomendadas como alternativa para prover interoperabilidade. Desta forma, são apresentadas iniciativas do governo brasileiro. Por fim, conclui que ontologias podem desempenhar papel relevante na busca pela interoperabilidade entre sistemas.

O trabalho desenvolvido por (Plastiras, 2014) apresenta uma solução para o problema de interoperabilidade entre sistemas de prontuários a idéia é estabelecer um ambiente de uniformidade semântica que permita a integração de registros pessoais de saúde oriundos de quaisquer sistemas. Para tanto, foi desenvolvido um modelo de informação que utiliza uma ontologia para garantir a integridade

semântica entre os conceitos. A Ontologia proposta funciona como uma camada intermediária entre os sistemas de registro e, portanto, não está vinculado a qualquer sistema específico.

3.3 ONTOLOGIAS PARA O DOMÍNIO EDUCACIONAL

No processo de busca de trabalhos que abordassem a construção de ontologias para o domínio da educação, não foram encontrados trabalhos que se dedicassem a construir modelos que representassem o domínio dos registros acadêmicos, entretanto, foram selecionados alguns trabalhos que guardam alguma similaridade com o trabalho aqui apresentado, uma vez que, se dedicaram a construir ontologias do domínio educacional, porém com foco para o ensino a distância.

O trabalho desenvolvido por (Bucos, et. al, 2011) apresenta uma ontologia que modela sistemas de ensino à distância baseado na web para o ensino superior. O principal objetivo do trabalho foi criar uma ontologia de domínio para representar os conceitos de ensino superior e auxiliar sistemas de ensino à distância. Já (Malviya, 2011) constrói uma ontologia do domínio da educação, entretanto tal modelo ontológico construído é destinado a mapear especificamente a área de avaliações de aprendizado e conceitos relacionados ao ensino à distância.

Por fim, o trabalho de (Chung e Kim, 2012) tem como objetivo principal desenvolver um sistema de suporte ao ensino à distância baseado em ontologias que permita ao aluno a construir caminhos de aprendizagem adaptáveis de acordo com o currículo do curso, os programas de ensino e áreas específicas do curso.

3.4 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os principais trabalhos correlatos que abordam o tema da interoperabilidade entre sistemas e que proponham a construção de ontologias para o domínio da educação. No levantamento bibliográfico não foram encontrados trabalhos que discutissem a construção de ontologias para o domínio dos registros acadêmicos. Entretanto, foram encontrados trabalhos que geraram ontologias para o domínio da educação à distância, além de

trabalhos que buscam a integração de sistemas de informação a partir da uniformidade semântica estabelecida por uma ontologia.

O próximo capítulo apresenta e discute o processo de escolha da metodologia para a construção da ONTO-RA.

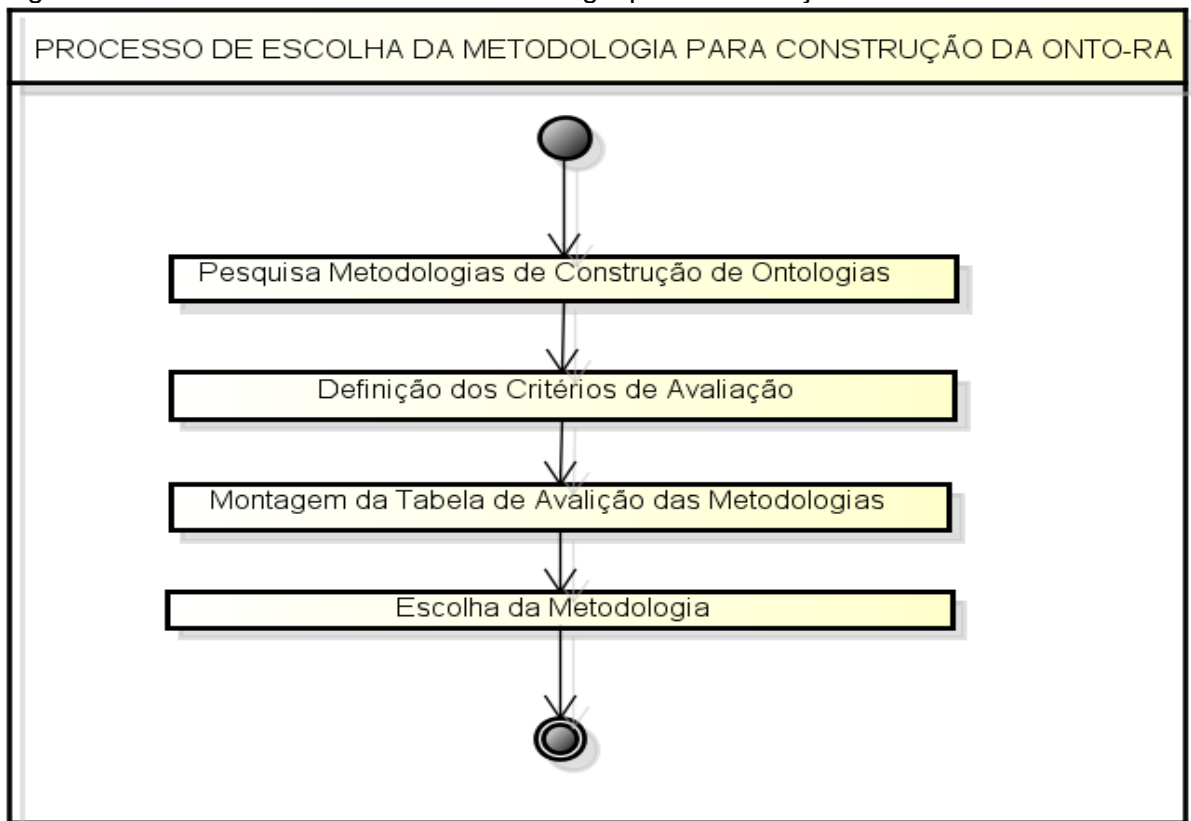
4 PROCESSO DE ESCOLHA DA METODOLOGIA PARA CONSTRUÇÃO DA ONTO-RA

Neste capítulo será apresentado o processo de escolha da metodologia para a construção da Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro – ONTO-RA, a partir dos resultados gerados no Capítulo 3 que indicaram a ausência de ontologias do domínio educacional que pudessem ser reutilizadas no processo de construção da ONTO-RA. Inicialmente serão discutidos trabalhos que abordam o tema de Engenharia de Ontologias, especificamente, no que se refere a metodologias de construção de ontologias, onde as principais metodologias serão apresentadas numa perspectiva cronológica. Em seguida serão discutidos todos os passos que foram executados para a definição pela metodologia NEON, avaliada como mais adequada perante aos critérios de avaliação estabelecidos. Por fim, a metodologia NEON é apresentada em detalhes.

4.1 VISÃO GERAL DO PROCESSO

Após a constatação de que não seria possível reutilizar ou estender outra ontologia já construída para a construção da ONTO-RA, seguindo o planejamento estabelecido Seção 1.4, foi necessário identificar a metodologia de construção de ontologia mais adequada para ser utilizada na construção da ONTO-RA. Assim, foi definido um processo a ser executado com o objetivo de realizar uma pesquisa sobre as metodologias de construção de ontologia disponíveis e realizar uma avaliação das mesmas a partir de critérios pré-estabelecidos, para determinar a metodologia mais adequada para a construção da ONTO-RA. A Figura 1 mostra todas as tarefas executadas durante este processo.

Figura 1 - Processo de Escolha da Metodologia para Construção da ONTO-RA



O processo inicia com a pesquisa que objetiva levantar as principais metodologias de construção de ontologias e classificá-las de acordo com as suas características. A partir destas características, ocorreu uma pré-seleção das metodologias e o consequente descarte de outras. Esta atividade será detalhada na Seção 4.2. Em seguida, são pesquisados trabalhos que se dedicam a realizar avaliação de metodologias de construção de ontologias. Destes trabalhos, são identificados os critérios utilizados. Tais critérios foram analisados e avaliados para identificar se serão ou não utilizados para avaliar as metodologias pré-selecionadas. O conjunto de atividades que definiu os critérios a serem utilizados para a avaliação das metodologias está descrito na Seção 4.3.

A partir dos critérios definidos, foi realizada uma avaliação de cada metodologia pré-selecionada em função dos critérios escolhidos e os resultados da avaliação de cada critério foi inserido na tabela de avaliação das metodologias e após a avaliação de todos os critérios a metodologia NEON foi escolhida como a metodologia mais adequada para a construção da ONTO-RA. A montagem da tabela de avaliação das metodologias e a definição pela mais adequada estão descritas em detalhes na Seção 4.4.

4.2 METODOLOGIAS DE DESENVOLVIMENTO DE ONTOLOGIAS

Até a primeira metade da década de 1990, no processo de desenvolvimento de ontologias cada desenvolvedor utilizava os seus próprios princípios e critérios, bem como seguia os passos que entendesse necessários para o desenvolvimento (GOMEZ-PEREZ; FERNÁNDEZ-LÓPEZ; CORCHO, 2004a). Como consequência, havia uma grande dificuldade em desenvolver ontologias, assim como reutilizar e realizar extensões em qualquer ontologia já construída.

Engenharia de Ontologias refere-se ao conjunto de atividades que compreende o processo de desenvolvimento de ontologias, o ciclo de vida das ontologias, bem como as metodologias, ferramentas e linguagens necessárias para a construção de ontologias (GOMEZ-PEREZ; FERNÁNDEZ-LÓPEZ; CORCHO, 2004a). As primeiras publicações que tratavam de tais assuntos começaram a surgir na primeira metade da década de 1990, inicialmente com abordagens que abrangiam o processo de construção de ontologias, sem insumos de entrada, ou seja, sem aproveitamento ou reuso de informações sejam ontológicas ou não. Os estudos na área se intensificaram e atualmente, além das abordagens tradicionais surgidas na década de 1990, existem estudos que abordam a criação de ontologias a partir do reuso de recursos ontológicos e não ontológicos, a utilização de trabalho colaborativo e distribuído, construção de redes de ontologia e processos de engenharia reversa para obtenção de ontologias.

Segundo (JIMÉNEZ-RUIZ e BERLANGA, 2006), as metodologias de construção de ontologias podem ser divididas em dois grupos. O primeiro representa um conjunto de metodologias clássicas que propõe uma abordagem centralizada de desenvolvimento de ontologia, sem levar em conta atividades de colaboração e a possibilidade de desenvolvimento distribuído. Os trabalhos do segundo grupo não propõem metodologias completas, mas soluções para a construção consensual da definição do conhecimento. O Quadro 2 apresenta as metodologias encontradas na pesquisa bibliográfica realizada para este trabalho, com suas respectivas classificações, em ordem cronológica, além de indicar o resultado da análise de pré-seleção, indicado se a metodologia foi pré-selecionada, garantindo assim que ela participará do processo de avaliação descrito na Seção 4.3, ou indicando que a mesma foi descartada e portanto não será analisada em detalhes.

Quadro 2 - Metodologias de Construção de Ontologias em Ordem Cronológica

Nº	Metodologia	Classificação	Pré-Seleção
1	CYC (LENAT e GUHA, 1989)	Clássica	Selecionada
2	Enterprise (USCHOLD e KING, 1995)	Clássica	Selecionada
3	TOVE (GRÜNINGER e FOX, 1995)	Clássica	Selecionada
4	Kactus (BERNARAS; LARESGOITI; CORERA, 1996)	Clássica	Selecionada
5	METHONTOLOGY (GÓMEZ-PÉREZ; FERNÁNDEZ; VICENTE, 1996)	Clássica	Selecionada
6	CO4 (EUZENAT, 1996)	Moderna	Descartada
7	SENSUS (SWARTOUT et al., 1996)	Clássica	Selecionada
8	SABiO (DE ALMEIDA FALBO; DE MENEZES; DA ROCHA, 1998)	Clássica	Selecionada
9	KA ² (RICHARD BENJAMINS et al., 1999)	Moderna	Descartada
10	On-To-Knowledge (OTK) (Staab SURE; STAAB; STUDER, 2002)	Clássica	Selecionada
11	Ontology Development 101 (NOY et al., 2001)	Clássica	Selecionada
12	ONTOCLEAN (GUARINO e WELTY, 2002)	Avaliação	Descartada
13	Metodologia de Sugumaran e Storey (2002)	Engenharia Reversa	Descartada
14	Metodologia de Dogan e Islamaj (2002)	Engenharia Reversa	Descartada
15	Metodologia de Saias e Quaresma (2003)	Engenharia Reversa	Descartada
16	DILIGENT (VRANDECIC, 2005)	Moderna	Descartada
17	Metodologia de Bueno (2005)	Engenharia Reversa	Descartada
18	Metodologia de Bortoleto (2010)	Engenharia Reversa	Descartada
19	NeOn (SUAREZ-FIGUEROA, 2010)	Mista	Selecionada

No grupo das metodologias Clássicas estariam as metodologias CYC (LENAT e GUHA, 1989), Enterprise (USCHOLD e KING, 1995), TOVE (GRÜNINGER e FOX, 1995), Kactus (BERNARAS; LARESGOITI; CORERA, 1996), METHONTOLOGY

(GÓMEZ-PÉREZ; FERNÁNDEZ; VICENTE, 1996), SENSUS (SWARTOUT et al., 1996), SABIO (DE ALMEIDA FALBO; DE MENEZES; DA ROCHA, 1998), OTK (Staab SURE; STAAB; STUDER, 2002) e Ontology Development 101 (NOY et al., 2001). No grupo das metodologias Modernas, estariam as metodologias KA² (RICHARD BENJAMINS et al., 1999), CO4 (EUZENAT, 1996) e DILIGENT (VRANDECIC, 2005), que preconizam a utilização de algum método clássico para a fase de construção da ontologia, ou deixa tal escolha a cargo do projetista. Essas metodologias centram as suas atenções para a fase do ciclo de vida da ontologia, responsável pela manutenção. Por este motivo, foi descartada a possibilidade de utilizar qualquer uma das metodologias modernas para o processo de construção da proposta de ontologia para o domínio dos registros acadêmicos do modelo brasileiro de educação, entretanto, há que se fazer uma nova avaliação para um trabalho futuro de manutenção da ontologia a ser criada. Todavia, a metodologia NeOn (SUAREZ-FIGUEROA, 2010) se enquadra em ambos os grupos, uma vez que aborda o processo de construção de ontologias, assim como o processo de manutenção colaborativa.

É conveniente ressaltar que durante a pesquisa outras metodologias foram encontradas, entretanto, foram descartadas em função da não adequação para o projeto de construção da ontologia de registros acadêmicos do modelo educacional brasileiro – ONTO-RA. Assim, não foram estudadas as metodologias de Sugumaran e Storey (2002), Bueno (2005), Saias e Quaresma (2003), Dogan e Islamaj (2002) e Bortoleto (2010), por tratar-se de metodologias de engenharia reversa, ou seja, inicia-se a criação a partir de artefatos gerados durante o ciclo de vida do desenvolvimento de sistemas e se gera a ontologia. Já a metodologia ONTOCLEAN (GUARINO e WELTY, 2002), caracterizou-se como um método para avaliação de ontologias. Portanto, sem apresentar um processo de construção, desta forma, tal metodologia também não teve aproveitamento neste trabalho.

4.3 CRITÉRIOS DE AVALIAÇÃO

A maioria dos trabalhos que se propõem a realizar estudos comparativos a respeito de metodologias para construção de ontologias, como (DA SILVA, SOUZA e ALMEIDA, 2008), (FERNÁNDEZ-LÓPEZ e GÓMEZ-PÉREZ, 2002) e (FERNÁNDEZ-LOPEZ, 1999), utilizam critérios que avaliam aspectos da

metodologia, como detalhamento das técnicas e processos propostos, recomendações para a formalização do conhecimento, dependência do método em relação ao desenvolvimento de aplicações, estratégias para identificar conceitos e comparação com o padrão IEEE 1074-1995, que descreve o processo de desenvolvimento de software. Outros critérios avaliam o nível de maturidade das propostas, como, recomendação de ciclo de vida, se técnicas foram desenvolvidas ou recomendadas nas diversas atividades da proposta e quais ontologias foram desenvolvidas utilizando a metodologia.

Entre os critérios identificados, considerando-se a necessidade de identificação da metodologia mais adequada para o desenvolvimento da Ontologia para o Domínio dos Registros Acadêmicos do Modelo Educacional Brasileiro, conforme cenário descrito na Seção 1.1, o único descartado foi a comparação com o padrão IEEE 1071-1995. Entre os critérios selecionados consideram-se os mais importantes os seguintes, em ordem crescente de relevância:

1) Ontologias desenvolvidas: é importante que já exista um volume significativo de ontologias construídas utilizando a metodologia, para que o processo tenha sido validado;

2) Detalhamento das Técnicas e Processos Propostos: um alto nível de detalhamento é importante para orientar o desenvolvimento da ontologia. Outros critérios, que serão incluídos na avaliação, porém considerados com menor grau de importância são:

3) Recomendações para a Formalização do Conhecimento;

4) Técnicas Recomendadas

5) Dependência de Aplicação: é importante que o método não dependa necessariamente da construção de uma aplicação, uma vez que no cenário descrito na Seção 1.1, a ontologia a ser criada deverá servir como base para aplicações diversas já existentes tanto no âmbito das instituições federais de ensino, quanto no MEC. Entretanto, tal critério não é determinante

6) Ciclo de Vida: é desejável que a metodologia aborde aspectos relacionados ao ciclo de vida, uma vez que o cenário descrito na Seção 1.1 é dinâmico e a ontologia precisa ser mantida e atualizada para que o seu propósito de utilização seja mantido;

7) Estratégias para Identificação dos Conceitos: Este critério, embora seja desejável, entra como último na ordem de relevância em função do cenário descrito

na Seção 1.1 identificar um ambiente em que os conceitos já estão identificados na sua maioria.

Para estabelecer uma comparação entre as diversas metodologias foram definidos os possíveis valores para cada critério a ser utilizado na avaliação. Para o critério 2) Detalhamento, foi estabelecido que os valores possíveis na avaliação fossem, em ordem crescente de detalhamento, Muito Baixo, Baixo, Mediano e Alto. Para os demais critérios os valores possíveis foram Sim e Não. O Quadro 4 apresenta os valores desejáveis para cada critério de avaliação.

Quadro 3 - Valores desejáveis para os critérios de avaliação das metodologias

CRITÉRIO DE AVALIAÇÃO	VALOR DESEJÁVEL
1) Ontologias Desenvolvidas	SIM
2) Detalhamento	ALTO
3) Recomendações para Formalização do Conhecimento	SIM
4) Dependência de Aplicação	NÃO
5) Estratégias para Identificação de Conceitos	SIM
6) Ciclo de Vida	SIM
7) Técnicas Recomendadas	SIM

4.4 MONTAGEM DA TABELA DE AVALIAÇÃO E ESCOLHA DA METODOLOGIA

Após a identificação e seleção dos critérios de avaliação a serem utilizados na Seção 4.3, cada uma das metodologias pré-selecionadas na Seção 4.2 foi avaliada a partir destes critérios. O Quadro 5 ilustra a avaliação das metodologias a partir dos critérios de avaliação selecionados. Dos critérios encontrados no levantamento bibliográfico, apenas o critério comparação com o padrão IEEE 1074-1995 foi descartado em função da pouca relevância em relação ao cenário descrito na Seção 1.1.

Quadro 4 - Avaliação das metodologias de construção de ontologias conforme critérios escolhidos

METODOLOGIAS	CRITÉRIOS DE AVALIAÇÃO EM ORDEM DE IMPORTÂNCIA							Número de Critérios Satisfeitos
	Ontologias Desenvolvidas	Detalhamento	Recomendações para a Formalização do Conhecimento	Técnicas Recomendadas	Dependência de Aplicação	Ciclo de Vida	Estratégia para Identificação de Conceitos	
Valores Desejados	Sim	Alto	Sim	Sim	Não	Sim	Sim	-
CYC	Sim	Baixo	Não	Não	Não	Não	Sim	2
Enterprise	Sim	Baixo	Não	Não	Não	Não	Sim	3
TOVE	Sim	Baixo	Sim	Não	Sim	Não	Sim	3
Kactus	Sim	Muito Baixo	Não	Não	Sim	Não	Sim	2
METHONTOLOGY	Sim	Alto	Sim	Sim	Não	Sim	Sim	7
SENSUS	Sim	Médiano	Sim	Não	Sim	Não	Não	2
SABiO	Sim	Alto	Sim	Sim	Não	Não	Sim	6
On-To-Knowledge	Sim	Alto	Sim	Sim	Sim	Sim	Sim	6
Ontology Development 101	Sim	Baixo	Não	Não	Não	Não	Não	1
NeOn	Sim	Alto	Sim	Sim	Não	Sim	Sim	7

Os Quadros 3 e 4 permitem estabelecer uma análise comparativa de abordagens distintas, sob a perspectiva da ordem de importância dos critérios estabelecidos para a escolha da metodologia a ser utilizada para a construção da ONTO-RA. A partir desta análise, pode-se perceber que, as metodologias mais indicadas são METHONTOLOGY e NeOn, que atendem a todos os critérios de forma satisfatória. SABiO e On-To-Knowledge deixaram de atender a apenas um critério. A SABiO não possui recomendações explícitas a cerca do ciclo de vida da ontologia criada e a On-To-Knowledge possui dependência em relação à aplicação desenvolvida em conjunto com a ontologia. Já as metodologias CYC, Enterprise, TOVE, Kactus, SENSUS e Ontology Development 101 não detalham os procedimentos a serem realizados em cada etapa da construção. Sendo este um critério importante, nenhuma destas metodologias poderá ter indicação para o desenvolvimento da ontologia proposta.

A literatura encontrada no âmbito das metodologias de construção de ontologias aponta para um cenário de muita diversidade entre as abordagens. Não havendo, portanto, um senso comum a respeito das melhores práticas e direções a serem seguidas para a construção de ontologias. Ao contrário disso, têm-se propostas que se adequam a propósitos e aplicações em contextos distintos, portanto, é preciso analisar a metodologia a ser utilizada na construção de ontologias, a partir do cenário em que esta construção se dará.

Os critérios discutidos e avaliados de forma comparativa neste trabalho indicam NeOn e METHONTOLOGY como as metodologias mais indicadas para a construção da ONTO-RA. Entretanto, a metodologia NeOn no seu primeiro cenário, faz uso de uma abordagem que utiliza diversas das propostas que constam na metodologia METHONTOLOGY adicionando, porém, possibilidade de utilização de recursos não ontológicos e tratando a possibilidade de utilização de trabalho colaborativo. Desta maneira, levando em conta que o cenário descrito na Seção 1.1 deste trabalho é aderente ao cenário dois da metodologia NeOn e que a mesma utiliza como base a metodologia METHONTOLOGY para o processo de construção da ontologia, é que se conclui que a metodologia NeOn é a mais indicada para o trabalho de construção da proposta de Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro – ONTO-RA.

4.5 METODOLOGIA NEON

O projeto NeOn (<http://www.neon-project.org/>) (SUAREZ-FIGUEROA, 2010) desenvolveu uma metodologia homônima para criação de ontologias que parte das premissas de que existe uma série de lacunas nas abordagens tradicionais das metodologias de desenvolvimento de ontologias e que existe um paradigma de construção de ontologias baseado no reuso e na reengenharia de recursos de conhecimento, desenvolvimento colaborativo e construção de redes de ontologias.

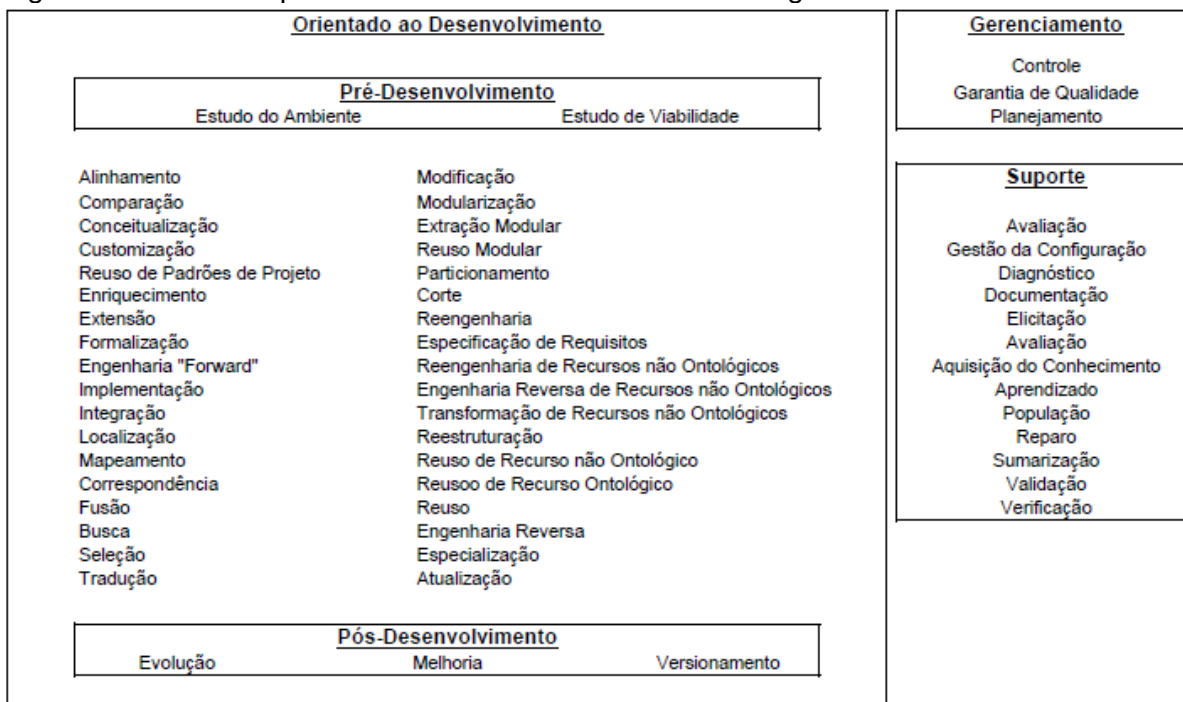
As principais lacunas identificadas pela metodologia são:

- Não há detalhamento nem diretrizes claras a respeito do reuso e reengenharia de recursos de conhecimento do domínio em particular que se deseja representar;
- Não são considerados os cenários em que a reutilização é o aspecto chave;
- Não são levados em consideração diferentes modelos de ciclo de vida relacionados com os possíveis cenários de reutilização;
- O processo de construção da ontologia não é exposto com o mesmo nível de granularidade, por exemplo, de algumas metodologias de desenvolvimento de software, o que poderia facilitar a adoção das metodologias.

Desta maneira, com o objetivo de cobrir as principais desvantagens das metodologias tradicionais, mas ao mesmo tempo se beneficiar das vantagens das mesmas, além de cobrir um novo paradigma de desenvolvimento de ontologia, baseado em reuso, desenvolvimento colaborativo e redes de ontologias, é que foi criada a metodologia NeOn, que fornece um conjunto de nove cenários que podem ser combinados entre eles. Cada cenário é decomposto em diferentes processos e atividades. Todos os processos e atividades previstos em cada um dos cenários constam do Glossário de Processos e Atividades da metodologia. Cada atividade é detalhada de acordo com a sua possível utilização a depender do cenário ou conjunto de cenários escolhidos para o desenvolvimento da ontologia.

A metodologia NEON identificou um conjunto de 59 processos e atividades que, a depender do cenário de desenvolvimento poderão ser obrigatórias ou opcionais. Tais atividades e processos foram classificados e distribuídos em cinco grupos principais, que são Gerenciamento de Processos e Atividades (Gerenciamento), Processos e Atividades Orientadas ao Desenvolvimento (Orientado ao Desenvolvimento) que é composto por atividades de Pré-desenvolvimento, Desenvolvimento e Pós-Desenvolvimento. Por fim, estão agrupados as atividades e processos de Suporte. As 59 atividades identificadas no glossário da metodologia NEON estão listadas e agrupadas na Figura 2.

Figura 2 - Atividades por Fase do Ciclo de Vida da Metodologia NEON



Fonte: Suárez-Figueroa adaptado (2010).

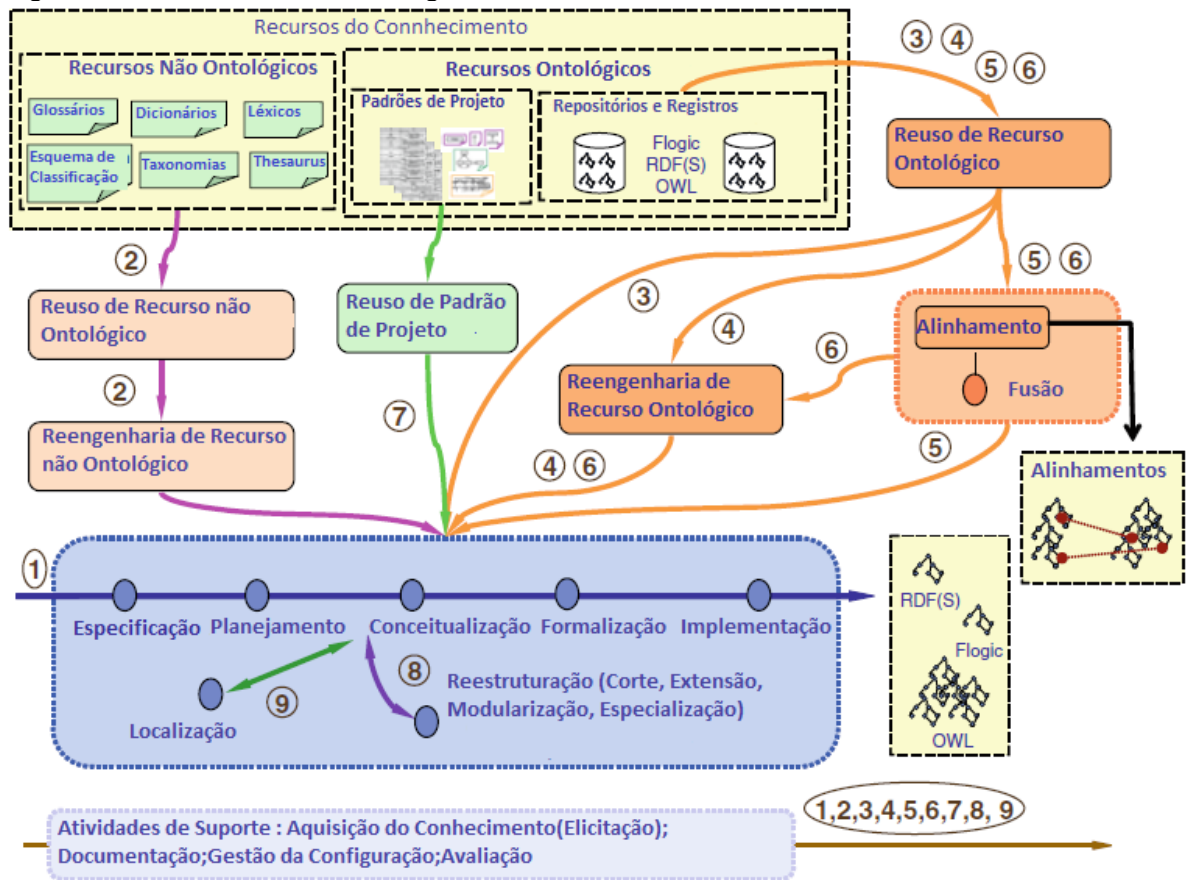
O agrupamento e a classificação destes processos e atividades basearam-se nas classificações propostas pelo IEEE (IEEE, 1997) e pela metodologia METHONTOLOGY. O conjunto completo de processos e atividades ilustrados na Figura 3 abrange todo o ciclo de vida de uma ontologia. Os grupos fornecem uma classificação administrativa que apresenta os processos e atividades de uma forma mais coerente.

Os nove cenários identificados e propostos na metodologia NEON, assim como as relações entre os mesmos são descritos a seguir e estão ilustrados na Figura 3:

- Cenário um: Da Especificação a Implementação. A ontologia ou rede de ontologias são desenvolvidas sem levar em conta recursos de conhecimento já existentes;
- Cenário dois: Reutilização e Reengenharia de Recursos não Ontológicos (NORs). Os desenvolvedores da ontologia deverão realizar um processo de reutilização de recursos não ontológicos (Esquemas de Classificação, Thesaurus, etc.), de acordo com o Documento de Especificação dos Requisitos da Ontologia - ORSD quais os NORs poderão ser reutilizados para a construção da ontologia;

- Cenário três: Reutilização de Recursos Ontológicos. Desenvolvedores utilizam recursos ontológicos (Ontologias completas, módulos de ontologias, etc.);
- Cenário quatro: Reutilização e Reengenharia de Recursos Ontológicos. Os desenvolvedores fazem o reuso e a reengenharia de recursos ontológicos;
- Cenário cinco: Reutilização e Fusão de Recursos Ontológicos. Este cenário ocorre somente nos casos em que diversos recursos ontológicos do mesmo domínio são selecionados para reuso e quando os desenvolvedores necessitam criar um novo recurso ontológico a partir de dois ou mais recursos.
- Cenário seis: Reutilização, Fusão e Reengenharia de Recursos Ontológicos. Este cenário é similar ao cenário cinco, entretanto, os desenvolvedores decidem não utilizar o conjunto de recursos originados da fusão, mas realizar um processo de reengenharia com eles;
- Cenário sete: Reutilização de Padrões de Projeto de Ontologias (ODPs). Os desenvolvedores buscam ODPs em repositórios para poderem reutilizá-los;
- Cenário oito: Reestruturação de Recursos Ontológicos. Os desenvolvedores reestruturam (modularizam, suprimem, entendem e/ou especializam) os recursos ontológicos para serem integrados à ontologia que está sendo construída;
- Cenário nove: Adaptação Local de Recursos Ontológicos: Desenvolvedores adaptam uma ontologia para outros idiomas, obtendo assim uma ontologia multilíngüe.

Figura 3 - Cenários da Metodologia NEON



Fonte: Suárez-Figueroa adaptado (2010).

As atividades são definidas de acordo com o cenário específico utilizado. O cenário um é utilizado por todos os outros cenários, uma vez que é o núcleo das atividades que precisam ser executadas por qualquer desenvolvimento de ontologia. Este cenário é baseado nas metodologias tradicionais da década de 90. Para cada cenário e tipo de ontologia desenvolvida (ontologia simples, interconectada ou rede de ontologias) existe um grupo de atividades e processos definidos no glossário da metodologia. Para a construção da ONTO-RA o cenário da metodologia NEON a ser utilizado será o cenário dois. Este é o cenário que prevê a reutilização e a reengenharia de recursos não ontológicos, estando, portanto, mais adequado ao cenário de construção da ONTO-RA, uma vez que para a sua construção, como estabelecido no Capítulo 1 objetiva-se utilizar os recursos não ontológicos disponíveis no domínio da educação brasileira.

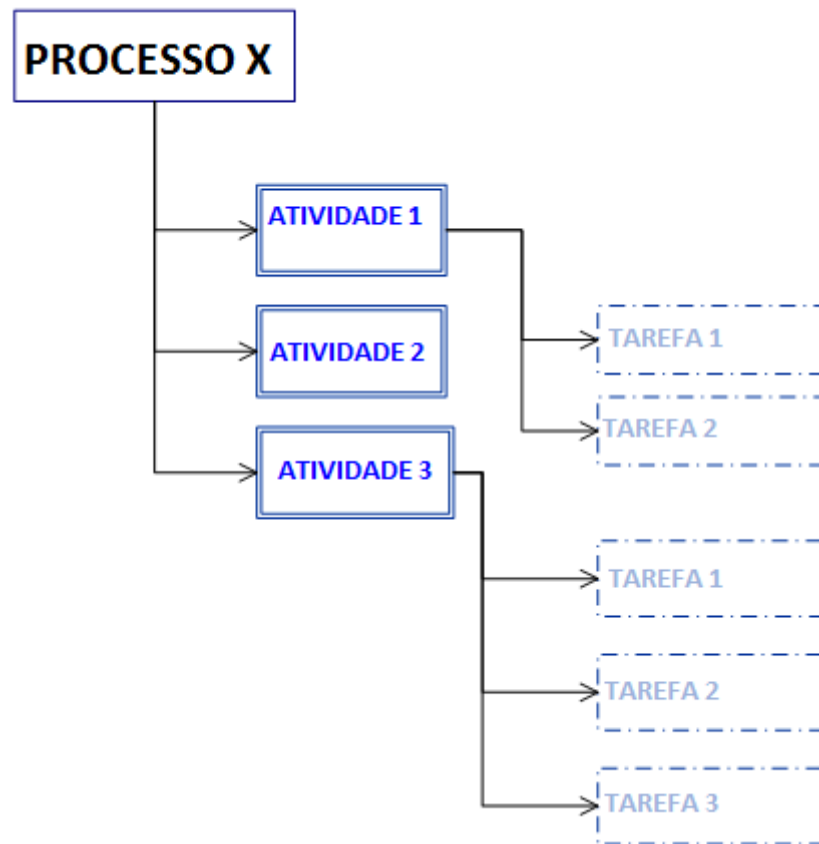
4.6 CENÁRIO DOIS DA METODOLOGIA NEON: REUTILIZAÇÃO E REENGENHARIA DE RECURSOS NÃO ONTOLÓGICOS

Reutilização de recursos tem sido uma prática bastante comum entre os desenvolvedores de ontologia atualmente (SUÁREZ-FIGUEROA, 2010). Recursos como Esquemas de Classificação, *Thesaurus*, Dicionários e Folksonomias já construídos têm levado agilidade ao processo de desenvolvimento de ontologias. O reuso destes recursos envolve necessariamente a reengenharia dos mesmos para que sejam agregados à ontologia a ser criada. Portanto, O cenário dois da metodologia NEON é aquele em que os desenvolvedores reutilizam recursos não ontológicos.

No caso específico da construção da ONTO-RA, serão aplicadas as atividades envolvidas nos cenários um (desde a Especificação até a Implementação), que é comum a todos os outros cenários, e dois (Reutilização e Reengenharia de Recursos não Ontológicos), uma vez que este cenário é o mais similar ao cenário descrito na Seção 1.1.

As atividades que compõem o Glossário de Processos e Atividades da metodologia NEON são chamadas de processos que podem ser compostos por uma ou mais atividades. As atividades podem ser compostas por zero ou mais tarefas (SUAREZ-FIGUEROA, 2008), conforme ilustrado na Figura 4.

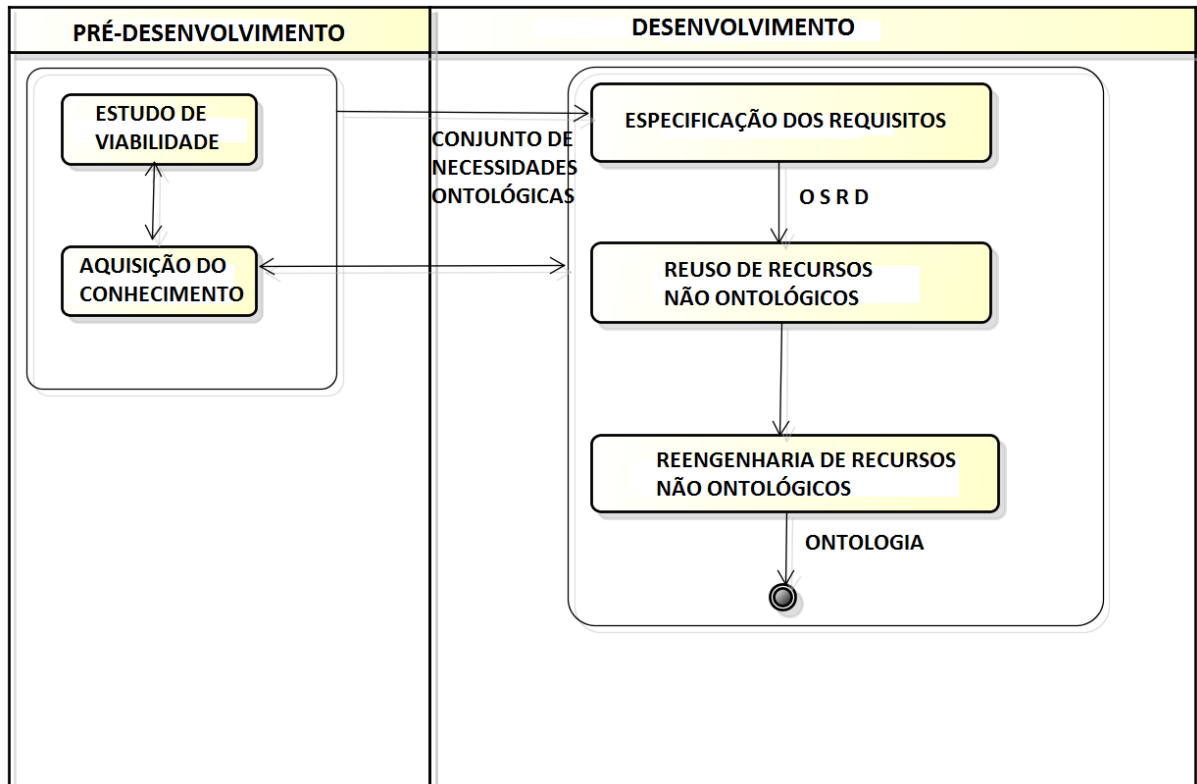
Figura 4 - Estrutura dos processos, atividades e tarefas da metodologia NEON



Fonte: Suárez-Figueroa adaptado (2008).

O fluxo de atividades e processos previstos pela metodologia NEON e executados durante a construção da ontologia, que é objeto deste trabalho, é ilustrado na Figura 5. Inicialmente foram executadas as atividades de pré-desenvolvimento (Estudo de Viabilidade e Aquisição do Conhecimento). Em seguida, a partir do conjunto de necessidades levantadas no pré-desenvolvimento, são executados nesta ordem, o processo de especificação dos requisitos, responsável por gerar o OSRD, o processo de reuso de recursos não ontológicos, que gera como saída um conjunto de recursos não ontológicos e por fim é executado o processo de reengenharia de recursos não ontológicos que resultará na ontologia.

Figura 5 - Processos e atividades executados



A primeira atividade do pré-desenvolvimento a ser realizada é o Estudo de Viabilidade e do Ambiente. Este estudo objetiva identificar se o desenvolvimento da ontologia irá conseguir resolver os problemas identificados como causas para o desenvolvimento da ontologia. Caso se decida por construir a ontologia, deve-se identificar se a solução exige o desenvolvimento de uma simples ontologia ou se deverá ser gerada uma rede de ontologias. Na metodologia não há registros de artefatos gerados nesta etapa.

Após o Estudo de Viabilidade e do Ambiente, é iniciada a atividade de Aquisição do Conhecimento, que é a segunda atividade do pré-desenvolvimento e pode incluir a atividade de Elicitação da Ontologia. Embora a aquisição do conhecimento se concentre mais fortemente no início do desenvolvimento da ontologia, ela deverá ser executada durante todo o processo de construção.

De forma simultânea à atividade de Aquisição do Conhecimento, ocorre a atividade de Especificação dos Requisitos da Ontologia, com o objetivo de gerar o Documento de Especificação dos Requisitos da Ontologia (OSRD). Este documento deve conter o propósito, o escopo e a linguagem da ontologia, os usuários da ontologia, o uso previsto da ontologia, além do conjunto de requisitos que a ontologia deverá atender, normalmente, em forma de questões de competência. A

atividade de especificação de requisitos é a única atividade do cenário “um” executada neste trabalho. Conforme ilustra a Figura 3, o cenário “um” é utilizado por todos os outros cenários em momentos específicos das atividades e processos definidos para cada um deles.

Logo após a geração do OSRD são executados os processos de Reuso de Recursos não Ontológicos, que utiliza como entrada o OSRD e gera como saída o conjunto de recursos não ontológicos que será utilizado pelo processo de Reengenharia de Recursos não Ontológicos, que culminará com a geração da ontologia. Estes dois processos pertencem ao cenário dois da metodologia NEON.

4.6.1 Especificação dos Requisitos

Conforme (SUÁREZ-FIGUEROA, 2010), o processo de especificação dos requisitos da ontologia na metodologia NEON é definido como uma coleção de requisitos que a ontologia precisa cumprir. A saída deste processo é o documento de especificação dos requisitos da ontologia – OSRD. Este documento deve conter o propósito, o escopo e a linguagem da ontologia, os usuários da ontologia, o uso previsto da ontologia, além do conjunto de requisitos que a ontologia deverá atender, normalmente, em forma de questões de competência. (GRÜNINGER, 1995) define questões de competência como um conjunto de perguntas que atribuem demanda sobre uma ontologia, ou ainda, como requisitos que deverão ser atendidos por uma ontologia, apresentados em forma de perguntas.

Oito Tarefas são executadas para a geração do OSRD, essas Tarefas são detalhadas a seguir. O fluxo de execução entre elas está ilustrado na Figura 6.

1) Identificar o propósito, escopo e o nível de formalidade : Nesta tarefa é identificado o principal objetivo da ontologia, sua cobertura e granularidade e grau de formalidade usado para codificar a ontologia, que poderá variar entre informal e formal.

2) Identificar os usuários pretendidos: Estabelece quem são os principais usuários pretendidos para a ontologia a ser criada;

3) Identificar os usos pretendidos: Identifica em quais cenários a ontologia será utilizada;

4) Identificar os requisitos: Identificação do conjunto de requisitos ou necessidades que a ontologia precisa atender. Os requisitos são escritos em

linguagem natural através de questões de competência. Como resultado, esta tarefa gera uma lista de questões de competência e suas respectivas respostas;

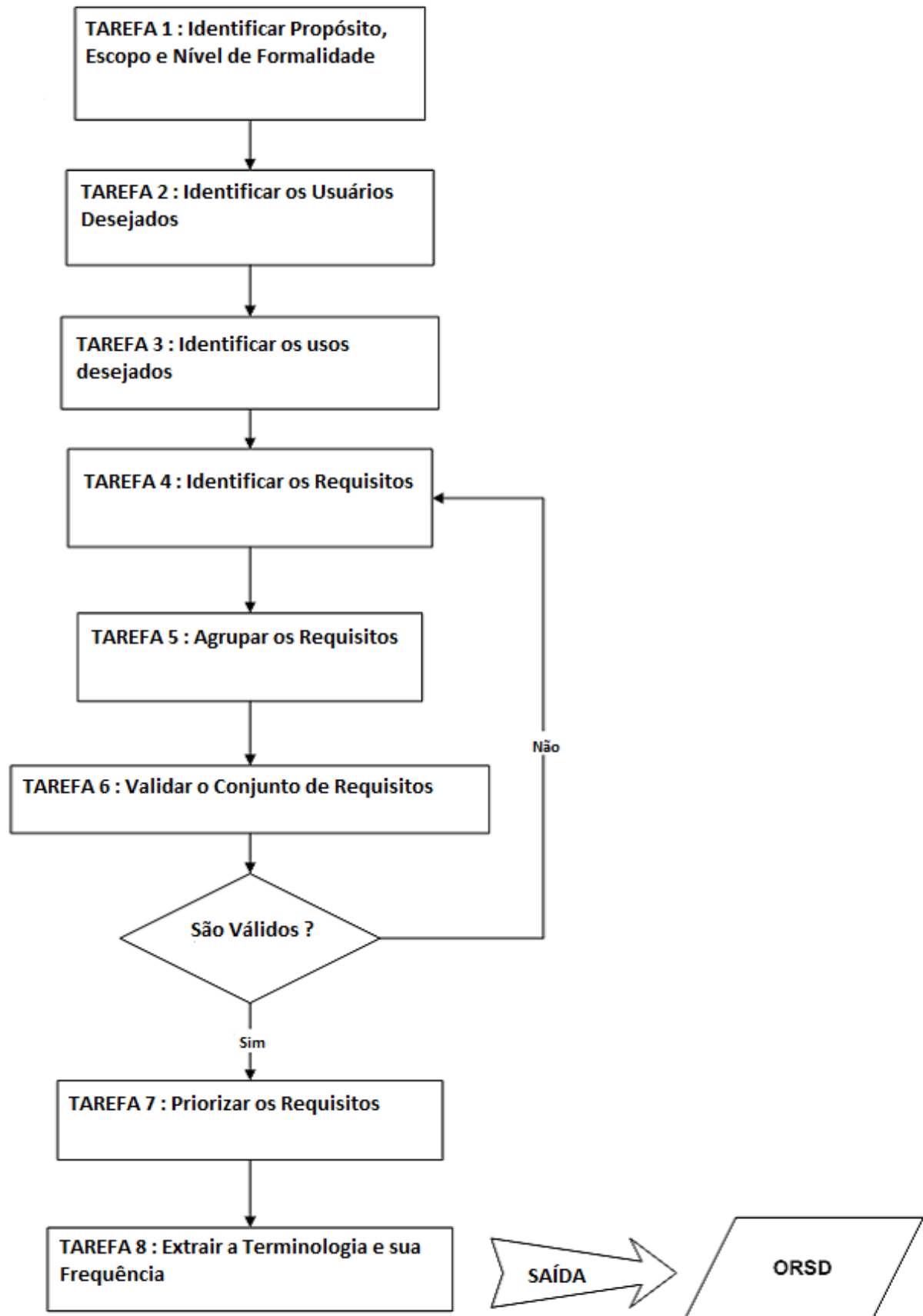
5) Agrupar os requisitos: Agrupa as questões de competência em diversas categorias;

6) Validação dos requisitos: Identifica possíveis conflitos e contradições entre as questões de competência, além de identificar uma eventual questão de competência ausente;

7) Priorizar requisitos: Esta tarefa é opcional e tem por objetivo estabelecer níveis de prioridade para cada um dos requisitos identificados;

8) Extrair a terminologia e sua frequência: O objetivo desta tarefa é extrair um glossário de termos e suas frequências a partir da lista de questões de competência e suas respostas, obtidas na atividade 4. Este glossário é dividido em três partes que são os termos das questões de competência, termos das respostas das questões de competência e termos identificados como objetos do universo do discurso.

Figura 6 - Tarefas da geração da OSRD



4.6.2 Reuso de Recursos não Ontológicos

As tarefas do processo Reuso de Recurso não Ontológico são as primeiras a serem executadas no cenário “dois” da metodologia NEON, Reuso e Reengenharia de Recursos não Ontológicos. Neste cenário os desenvolvedores da ontologia deverão analisar se existe recurso não ontológico que já possuam consenso em relação às terminologias representadas que possam ser reutilizados para a construção da ontologia. As tarefas e atividades deste processo estão dispostas na Figura 7. A busca por recursos não ontológicos é a primeira atividade do processo Reuso de Recursos não Ontológicos. O objetivo desta atividade é buscar por recursos não ontológicos, seja em web sites reconhecidamente confiáveis dentro do domínio da informação ou em documentos de dentro da própria organização. Com base nos termos de maior frequência identificados no OSRD, são buscados candidatos a recursos não ontológicos. Ao final desta atividade é gerada uma lista de candidatos a recursos não ontológicos.

A avaliação do conjunto de candidatos a recursos não ontológicos é a segunda atividade a ser realizada no processo de Reuso de Recursos não Ontológicos. O objetivo desta atividade é avaliar o conjunto de candidatos a recursos não ontológicos identificados na atividade anterior através de dois critérios mensuráveis, que são cobertura e precisão, e um critério subjetivo, que é o consenso. A atividade deverá gerar como saída um quadro de avaliação de critérios para cada um dos recursos. A atividade é composta por cinco tarefas.

A primeira tarefa é a extração de entradas léxicas de cada um dos candidatos a recursos não ontológicos. A segunda tarefa objetiva calcular o nível de precisão dos recursos não ontológicos. Tal medida representa a proporção de entradas léxicas do candidato a recurso não ontológico que estão entre os termos identificados no OSRD. A expressão algébrica que representa este cálculo é apresentada a seguir:

$$\text{Precisão} = \frac{\{ \text{NOEntradasLexicas} \} \cap \{ \text{ORSDDTerminologia} \}}{\{ \text{NOEntradasLexicas} \}}$$

Na qual, $\{ \text{NOEntradasLexicas} \}$ é o conjunto de entradas léxicas extraídas do candidato a recurso e $\{ \text{ORSDDTerminologia} \}$ é o conjunto de termos identificados no OSRD.

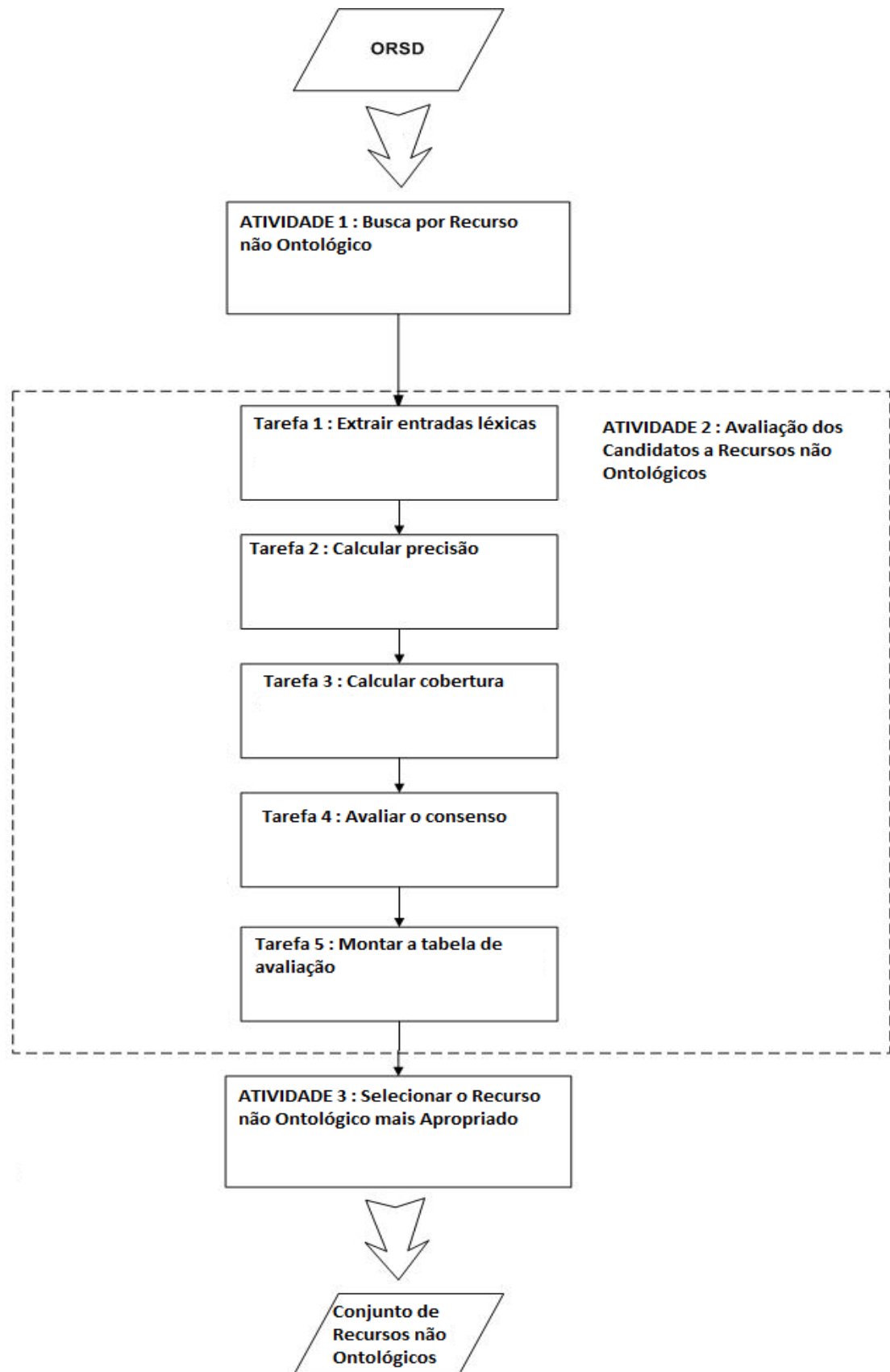
A terceira tarefa é o cálculo da cobertura do candidato a recurso não ontológico. A cobertura é obtida através da proporção dos termos identificados no ORSD que são encontrados nas entradas léxicas do candidato a recursos não ontológicos. A expressão algébrica que representa este cálculo é apresentada a seguir:

$$\text{Cobertura} = \frac{\{ \text{NOEntradasLexicas} \} \cap \{ \text{ORSDDTerminologia} \}}{\{ \text{ORSDDTerminologia} \}}$$

Na qual, {NOEntradasLexicas} é o conjunto de entradas léxicas extraídas do candidato a recurso e {ORSDDTerminologia} é o conjunto de termos identificados no ORSD. A quarta tarefa é a avaliação do consenso nas definições constantes dos candidatos a recursos não ontológicos. O objetivo desta tarefa é identificar se o conhecimento extraído do recurso é um consenso dentro do domínio da informação. Como já citado este critério é subjetivo. A quinta tarefa dentro da atividade de avaliação de recursos não ontológicos é a construção da tabela de avaliação. Esta tabela deverá conter uma coluna com o nome do candidato a recurso não ontológico, uma coluna para indicar o cálculo da precisão, outra para a cobertura e a última delas deverá indicar se há ou não consenso a respeito dos conceitos encontrados no recurso.

A partir da tabela elaborada na tarefa de construção da tabela de avaliação, deverá ser selecionado o recurso não ontológico mais apropriado para ser utilizado no processo de Reengenharia. A seleção é efetuada por meio dos critérios avaliados na tarefa de avaliação dos recursos não ontológicos. O critério de maior relevância é o Consenso. Na segunda posição está o critério Cobertura, uma vez que este critério determina a quantidade de termos do ORSD cobertos pelos candidatos a NOR. O critério menos importante é a Precisão. Então, é gerado um *ranking* de candidatos a NORS a partir dos critérios indicados em sua ordem de relevância.

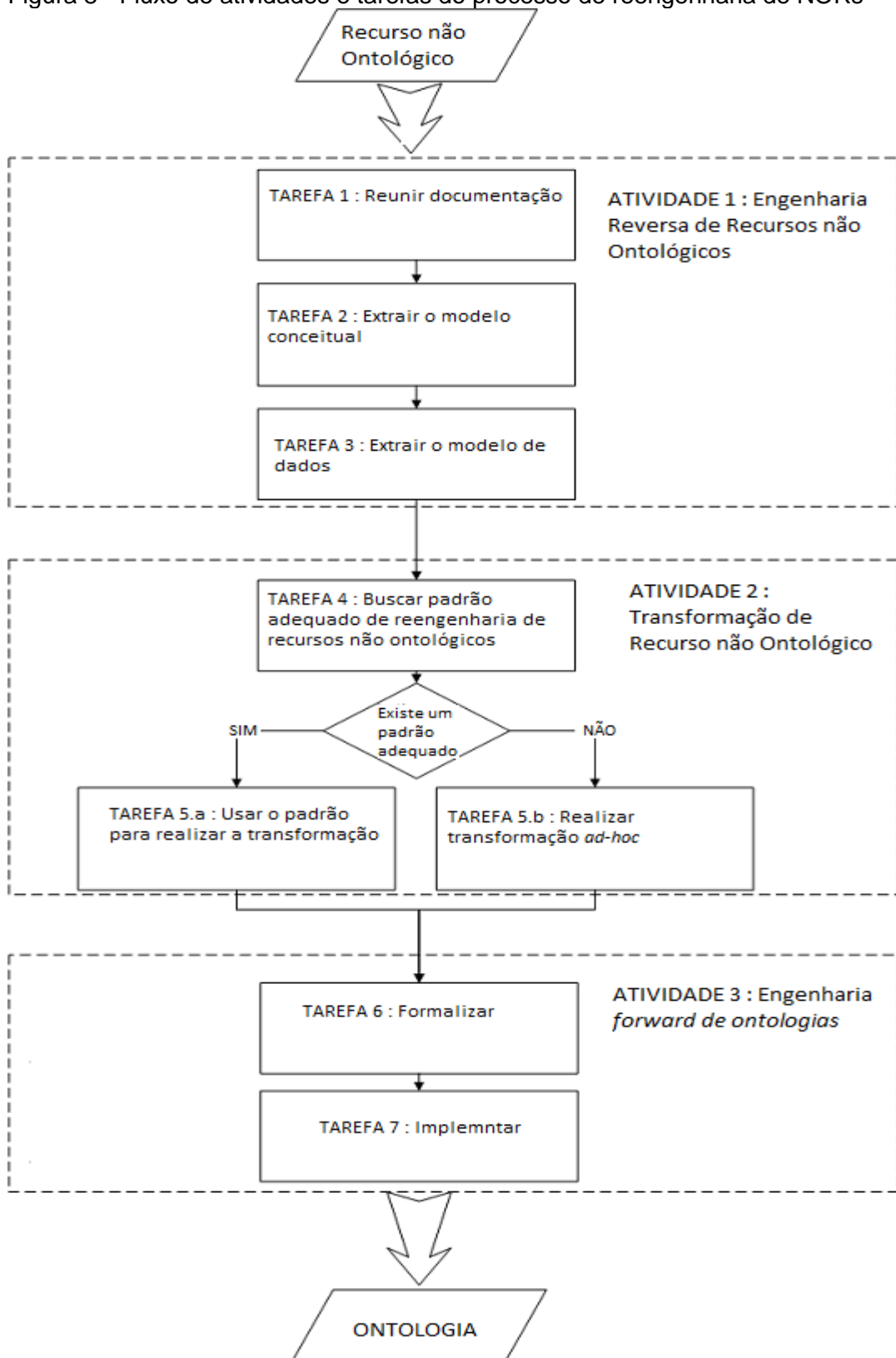
Figura 7 - Fluxo de atividades e tarefas de reuso de NORs



4.6.3 Reengenharia de Recursos não Ontológicos

Este processo tem como objetivos: (1) realizar a conversão completa do conhecimento incluído no recurso não ontológico identificado no processo de Reuso de Recursos não Ontológicos; (2) realizar a conversão do recurso para ontologia num único e simples passo; (3) aproveitar o modelo de dados do recurso não ontológico para guiar o processo de reengenharia; (4) empregar o padrão de reengenharia de recursos não ontológicos para conduzir o processo de transformação, (5) gerar ontologias num nível de conceitualização independente da linguagem utilizada para a implementação. O fluxo das atividades e tarefas executadas neste processo está mostrado na Figura 8.

Figura 8 - Fluxo de atividades e tarefas do processo de reengenharia de NORs



Esse processo, conforme apresentado na Figura 8, utiliza como entrada o recurso não ontológico selecionado no processo de reuso de recurso não ontológico

e os padrões de reengenharia de recursos não ontológicos. Ele é composto das seguintes atividades: (1) Engenharia Reversa de Recursos não Ontológicos, que correspondem às tarefas de reunir documentação, extração do esquema conceitual do recurso não ontológico e extrair o modelo de dados do recurso; (2) Transformação de Recurso não Ontológico, que correspondem às tarefas de busca pelo PR-NOR (padrão para reengenharia de recursos não ontológicos) mais adequado, utilização do padrão identificado para realizar a transformação ou realizar transformação *ad-hoc* e (3) Engenharia *Forward* de ontologia que é composta pelas tarefas de formalização e implementação.

A atividade de Engenharia Reversa de Recursos não Ontológicos tem por objetivo analisar um recurso não ontológico para identificar os seus componentes e criar uma representação do recurso num nível mais alto de abstração. As tarefas desta atividade são:

- Reunir a documentação: Consiste em reunir toda a documentação disponível a respeito do recurso como, propósito, componentes, modelo de dados e detalhes da implementação;
- Extrair o esquema do recurso não ontológico: Identifica o tipo do recurso, incluindo os componentes conceituais e suas relações. Se a documentação não apresentar um esquema estruturado, o recurso poderá ser reconstruído manualmente ou por uma ferramenta de modelagem de dados;
- Extrair modelo de dados: Objetivo desta tarefa é descobrir como o esquema do recurso não ontológico e seu conteúdo são apresentados no modelo de dados.

O objetivo da atividade de transformação de recurso não ontológico é gerar um modelo conceitual para o recurso não ontológico. A metodologia NEON preconiza a utilização de padrões para Reengenharia de Recursos não Ontológicos para guiar o processo de transformação. Padrões para reengenharia de recursos não ontológicos (PR-NOR) definem procedimentos que transformam recursos não ontológicos em representações primitivas de ontologia (VILLAZÓN-TERRAZAS, 2009). Para este propósito os PR-NORs utilizam como base o modelo de dados dos recursos não ontológicos, que definem como os diferentes componentes do recurso estão representados.

Para cada tipo de modelo de dados é possível definir um processo através de uma sequência bem definida de atividades, para extrair os componentes dos recursos não ontológicos e então, mapeá-los para um modelo conceitual de ontologia. Cada um destes processos é expresso como um padrão específico de reengenharia de NOR. As ontologias resultantes de propostas de padrões para reengenharia de recursos não ontológicos da metodologia NEON são modeladas seguindo padrões lógicos e de arquitetura definidos. Como padrões de arquitetura, são consideradas taxonomias e ontologias leves.

Nesta atividade é necessário que estejam identificados o tipo de recurso não ontológico, o modelo de dados interno do recurso não ontológico e as semânticas das relações entre as entidades do NOR. A semântica das relações entre as entidades do NOR podem ser (1) sub-classe; (2) relações *ad-hoc* ou (3) uma mistura de sub-classes e relações *ad-hoc*. Com a disponibilidade do tipo do recurso, o modelo de dados e a semântica das relações, procede-se a identificação de um PR-NOR adequado (Tarefa de Busca por um PR-NOR adequado) que servirá de guia para a tarefa de transformação do recurso em modelo conceitual.

A metodologia NEON estabelece uma forma de classificação dos recursos que leva em consideração três aspectos:

1. O tipo de recurso não ontológico;
2. O modelo de dados utilizado para representar o recurso;
3. A forma como o recurso está implementado.

De acordo com o tipo de recurso, a classificação poderá ser:

- Glossário: Dicionário de terminologias que contém designações e definições de uma ou mais áreas temáticas. (ISO 1087-1:2000, 2000);
- Dicionário: Um dicionário é uma coleção estruturada de unidades léxicas com informações linguísticas sobre cada um dos termos. (ISO 1087-1:2000, 2000);
- Léxicos: Léxico é um dicionário específico de uma pessoa, grupo ocupacional ou campo profissional;

- Esquema de Classificação: Definido em (ISSO/IEC, 1999) como informação descritiva de um arranjo ou divisão de objetos em grupos baseados em características que os objetos têm em comum;
- Thesaurus: Um thesaurus é um vocabulário controlado disposto numa ordem conhecida com o objetivo de facilitar a recuperação de informações e garantir uma indexação coerente (ANSI/NISO, 2005);
- Folksonomia: Que é o resultado da marcação livre de informações pessoais e objetos. (VANDER, 2007);

Um modelo de dados é um modelo abstrato que descreve como o dado é representado e acessado. Existem três estilos básicos de modelo de dados, são eles:

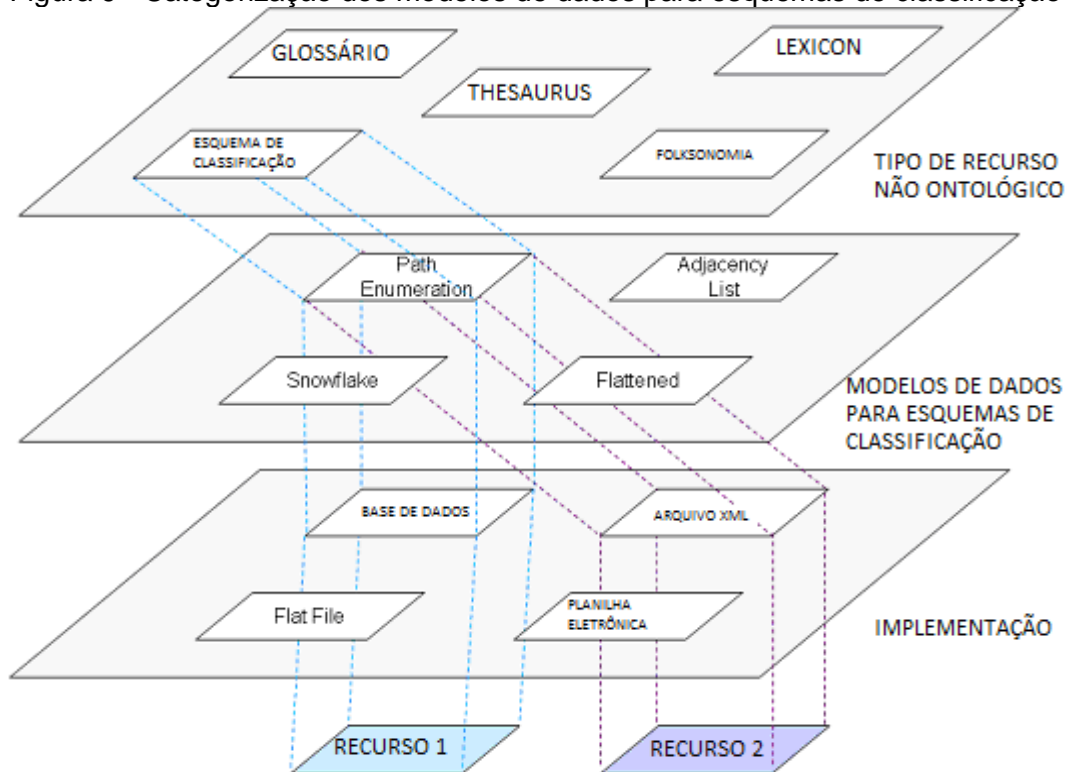
1. Modelo de dados conceitual;
2. Modelo de dados lógico; e
3. Modelo de dados físico (VILLAZÓN-TERRAZAS, 2009).

A metodologia NEON se refere ao modelo de dados físico quando estabelece o tipo de classificação do recurso não ontológico quanto ao modelo de dados. Esta classificação varia de acordo com o tipo de recurso. Quanto à implementação os recursos não ontológicos podem ser classificados em banco de dados, arquivo XML, arquivo simples ou planilha eletrônica. Para cada tipo de recurso não ontológico, há possibilidades de representações de modelo de dados e tipos de relacionamento específicos que são mapeados para padrões de reengenharia de recursos não ontológicos que deverão ser aplicados no processo de transformação do recurso em ontologia.

A Figura 9 apresenta o mapeamento entre as diversas formas de representação dos modelos de dados e formas de implementação. No primeiro plano, estão os diversos tipos de recurso não ontológico, glossário, esquema de classificação, *thesaurus*, *Lexicon* e Folksonomia. No segundo plano estão os possíveis modelos de dados para esquemas de classificação e no terceiro as possíveis formas de implementação dos recursos. No exemplo da Figura 10 estão representados dois recursos implementados em base de dados e arquivo XML.

Ambos os recursos utilizam o modelo de dados *Path Enumeration* e são do tipo esquema de classificação.

Figura 9 - Categorização dos modelos de dados para esquemas de classificação



Fonte: Suárez-Figueroa adaptado (2008).

A tarefa Engenharia *Forward* de Ontologia é a terceira atividade do processo de Reengenharia de Recursos não Ontológicos. O objetivo desta tarefa é gerar uma nova implementação da ontologia com base no modelo conceitual gerado na atividade de Transformação de Recurso não Ontológico e é composta de duas tarefas. A primeira delas é a de Formalização, cujo objetivo é transformar o modelo conceitual obtido num modelo formal. Esta tarefa é realizada de acordo com um paradigma de representação de conhecimento como lógica de descrição e lógica de primeira ordem. A segunda tarefa é a de Implementação, cujo objetivo é implementar a ontologia em uma linguagem de desenvolvimento de ontologias. Consiste em implementar o modelo formal em uma linguagem de ontologia.

4.7 CONSIDERAÇÕES FINAIS

O processo de escolha da metodologia para a construção da ontologia que é objeto deste capítulo levou em conta os critérios de avaliação de metodologias de construção de ontologias comumente utilizados na literatura, avaliando a sua adequação ao cenário no qual a ontologia seria criada. Foram estabelecidos critérios avaliativos e como resultado a escolha da metodologia NEON, que será detalhada em seus processos e atividades do cenário dois no próximo capítulo, que também descreverá o processo de construção da ontologia para o domínio dos registros acadêmicos do modelo brasileiro da educação.

A metodologia NeOn apresenta uma forma flexível e abrangente para desenvolvimento de ontologias, levando-se em conta a possibilidade de construção de redes de ontologias, realização de trabalho colaborativo, além de reutilização e reengenharia de recursos ontológicos e não ontológicos, além da gestão do ciclo de vida da ontologia. As atividades e processos propostos na metodologia são amplamente detalhados e são sugeridas diversas técnicas e ferramentas em diversos processos

5 CONSTRUÇÃO DA ONTOLOGIA DE REGISTROS ACADÊMICOS DO MODELO EDUCACIONAL BRASILEIRO – ONTO-RA

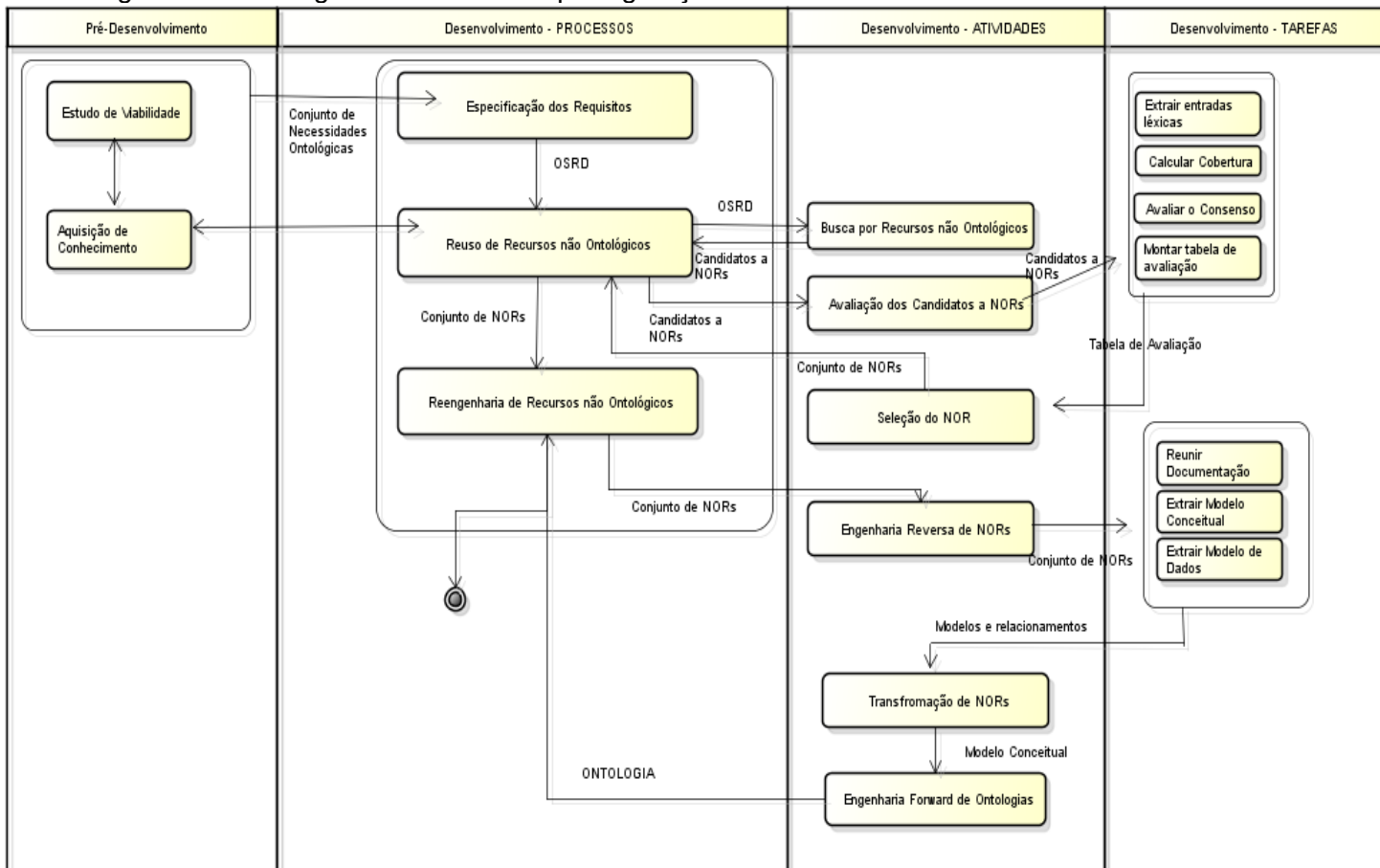
Neste capítulo será apresentada a construção da Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro – ONTO-RA. Serão detalhadas todas as etapas da aplicação da metodologia NEON, no seu cenário “dois”, no processo de construção da proposta até a apresentação da ONTO-RA, detalhando todos os passos executados no processo de construção da ONTO-RA, evidenciando, inclusive, as adaptações à metodologia realizadas durante o processo.

5.1 VISÃO GERAL DO PROCESSO DE CONSTRUÇÃO

A ONTO-RA foi construída tendo como base a utilização da metodologia de construção de ontologias NEON. A Figura 10 apresenta o fluxo de todos os processos executados durante a execução do trabalho de construção da ONTO-RA, com o devido registro dos artefatos produzidos. As atividades foram agrupadas por fases, que são Pré-desenvolvimento e Desenvolvimento. Sendo que a fase de desenvolvimento foi subdividida em três grupos que são Processos, Atividades e Tarefas.

A execução do processo de construção da ONTO-RA começa com as atividades de estudo de viabilidade e aquisição do conhecimento. Esta última perdurará por todo o processo de construção. Em seguida é iniciada a tarefa de construção do OSRD, artefato apresentado nos Quadros 5, 6, 7 e 8. A partir do OSRD é realizado o processo de reuso de NORs. A partir das tarefas executadas neste processo, é identificado o conjunto de NORs que servirá de base para o processo de Reengenharia de Recursos não Ontológicos, que inicialmente gera um modelo conceitual, apresentado no ANEXO B, e em seguida a Ontologia, apresentada parcialmente e de forma gráfica nas Figuras 12, 13, 14 e 15 e representada de forma completa no endereço <http://xbrlframework.wordpress.com/>.

Figura 10 - Fluxo geral de atividades para geração da ONTO-RA



5.2 ESTUDO DE VIABILIDADE E AQUISIÇÃO DO CONHECIMENTO

A primeira atividade executada durante a construção do ONTO-RA foi o Estudo de Viabilidade, que objetivou identificar se o cenário descrito no Capítulo 1, que evidencia um ambiente de heterogeneidade dos sistemas de informação utilizados nas instituições federais de ensino brasileiras e nos sistemas de coleta de dados utilizados pelo MEC, pode de fato ser solucionado através da construção da ONTO-RA. Tal premissa acabou sendo confirmada a partir do conhecimento acerca do uso de ontologias para promover integração entre sistemas heterogêneos e promover uniformidade semântica, tomando por base pesquisa na literatura da área como os trabalhos de (Moreira and Lopes, 2012), (Farinelli, et. al, 2013), (Plastiras, et. al, 2014) e (Park, 2014) que demonstram que a construção de uma ontologia é adequada para solucionar o problema relatado no Capítulo 1. Portanto, pode-se afirmar que no caso da construção da ONTO-RA esta atividade foi cumprida no momento da realização da revisão da literatura, que evidenciou o uso de ontologias

para resolução de problemas de interoperabilidade entre sistemas de informação baseado em semântica.

Após o Estudo de Viabilidade, foi iniciada a atividade de Aquisição do Conhecimento, que se concentrou no início do processo de desenvolvimento da ONTO-RA, mas foi executada durante todo o processo de construção, uma vez que qualquer atividade ou tarefa que tenha o propósito de obter informações ou esclarecer dúvidas a cerca do domínio dos registros acadêmicos da educação brasileira, é considerada uma atividade ou tarefa de aquisição do conhecimento. A primeira atividade executada durante a aquisição do conhecimento foi o levantamento de toda a documentação a cerca do domínio dos registros acadêmicos da educação brasileira. Assim, foram reunidos e consultados os modelos de dados, glossários e manuais dos sistemas de coletas de dados do MEC, o *thesaurus* brasileiro da educação e a documentação dos indicadores levantados pelo diagnóstico de evasão e retenção realizado pelo MEC, que garantiram o arcabouço de recursos suficientes para dar suporte à construção da ONTO-RA.

5.3 ESPECIFICAÇÃO DA ONTOLOGIA

O principal produto do processo de especificação da ONTO-RA é o documento de especificação de requisitos da ontologia – OSRD que está representado na Figura 9. As três primeiras tarefas executadas durante o processo de Especificação da Ontologia da ONTO-RA foram: (1) Identificar propósito, escopo e nível de formalidade; (2) Identificar os usuários desejados e (3) Identificar os usos desejados.

Na geração do OSRD da ONTO-RA, o propósito indicado, com base no cenário descrito nas Seções 1.1, 1.2 e 1.3, foi o de promover a uniformidade semântica no domínio dos registros acadêmicos no modelo educacional brasileiro. Devendo servir de base para o desenvolvimento de novos sistemas de gestão de registros acadêmicos, ajustes nos sistemas de gestão de registros acadêmicos já existentes, manutenção dos sistemas visando a uniformização da coleta de dados existentes no MEC e facilitar a interoperabilidade de informações entre MEC e Instituições Federais de Ensino.

O escopo da ONTO-RA, definido na Tarefa 1, foi todo o processo de registro acadêmico dentro do modelo brasileiro da educação, seus fluxos e informações

necessárias para geração de indicadores de produtividade acadêmicas e indicadores gerenciais. A granularidade da ontologia está diretamente relacionada às questões de competência e termos identificados. A linguagem de implementação de ontologia escolhida foi a *Ontology WEB Language* – OWL, que é a linguagem padrão definida pelo *World Wide Web Consortium* – W3C e pela própria metodologia NEON. Os resultados obtidos a partir da execução da Tarefa 1 podem ser vistos nos itens 1, 2 e 3 do Quadro 5.

Quadro 5 - OSRD da ONTO-RA

ONTOLOGIA DE REGISTROS ACADÊMICOS DO MODELO EDUCACIONAL BRASILEIRO – ONTO-RA
1) PROPÓSITO
Promover a uniformidade semântica no domínio dos registros acadêmicos no modelo brasileiro de educação. Devendo servir de base para o desenvolvimento de novos sistemas de gestão de registros acadêmicos, ajustes nos sistemas de gestão de registros acadêmicos já existentes, manutenção visando a uniformização dos sistemas de coleta existentes no MEC, facilitar a interoperabilidade de informações entre MEC e Instituições Federais de Ensino.
2)ESCOPO
A ontologia tem como foco todo o processo de registro acadêmicos dentro do modelo brasileiro da educação, seus fluxos e informações necessárias para geração de indicadores de produtividade acadêmica e indicadores gerenciais. A granularidade da ontologia está diretamente relacionada às questões de competência e termos identificados.
3)LINGUAGEM DE IMPLEMENTAÇÃO
OWL
4) USUÁRIOS FINAIS PRETENDIDOS
Usuário 1 : Usuário responsável por manter sistema de gestão de registros acadêmicos de Instituição Federal de Ensino; Usuário 2 : Usuário responsável por manter sistema de coleta de dados do MEC; Usuário 3 : Usuário que irá desenvolver um novo sistema de gestão de registros acadêmicos; Usuário 4: Usuário que irá desenvolver um novo sistema de coleta de dados do MEC; Usuário 5 : Gestores de Instituições Federais de Ensino; Usuário 6 : Gestores do MEC
5) UTILIZAÇÕES PREVISTAS PARA A ONTOLOGIA
Uso 1 : Referência para manutenção de sistemas de gestão de registros acadêmicos de Instituição Federal de Ensino Uso 2 : Referência para manutenção em sistemas de coleta de dados do MEC Uso 3 : Referência para desenvolvimento de novo sistema de gestão de registros acadêmicos Uso 4 : Referência para desenvolvimento de novo sistema de coleta de dados Uso 5 : Gestor de instituição Federal de Ensino em busca de atualização no modelo ou em busca de referência para ajuste em norma acadêmica institucional Uso 6 : Gestor do MEC em busca de atualização no modelo ou em busca de referência para ajuste em norma acadêmica institucional
6)GRUPO DE QUESTÕES DE COMPETÊNCIA
Disponível no Quadro 7
7)PRÉ-GLOSSÁRIO DE TERMOS
Disponível nos Quadros 8 e 9

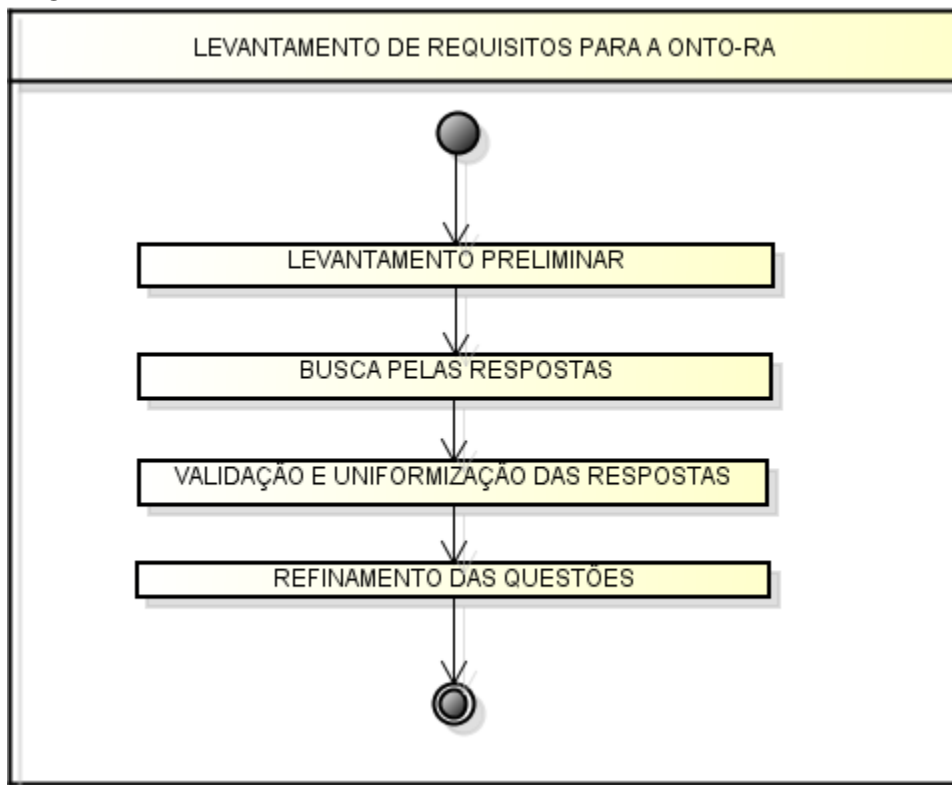
A partir dos objetivos e escopo identificados na Tarefa 1 e disponibilizados no item 1 do Quadro 5, foi possível estabelecer quais os principais usuários da ontologia, assim como, identificar os cenários em que estes usuários poderão se

beneficiar do uso da ontologia. Cada cenário corresponde a um conjunto de requisitos que a ontologia deverá atender. Os usuários identificados e os cenários de uso da ontologia, podem ser verificados no Quadro 5, nos itens 4 e 5 respectivamente.

5.3.1 Levantamento dos Requisitos em ONTO-RA

A partir dos objetivos e escopo definidos no OSRD, iniciou-se o processo de levantamento dos requisitos da ONTO-RA. Conforme previsto pela metodologia NEON, foi definido que os requisitos seriam escritos em forma de questões de competência. Partindo-se da premissa que as questões de competência representam o conjunto de perguntas que a ontologia a ser criada deverá ser capaz de apresentar respostas, e que o cenário que motivou a construção da ONTO-RA apresentado no Capítulo 1 evidencia a falta de uniformidade semântica dos sistemas de coletas de dados do MEC entre si e em relação aos sistemas de gestão acadêmica utilizados pelas instituições de ensino, dificultando assim a interoperabilidade, resolveu-se que as questões de competência seriam elaboradas a partir da documentação dos sistemas de coletas de dados, contando com o assessoramento de outros documentos como o TAMC. A LDB foi utilizada para dirimir eventuais dúvidas e atingir o consenso em relação a algumas respostas das questões levantadas. A Figura 11 apresenta o fluxo de atividades executadas durante o levantamento dos requisitos da ONTO-RA.

Figura 11 - Fluxo de Atividades Executadas no Levantamento de Requisitos da ONTO-RA



O levantamento preliminar das questões de competência, realizado Com base na análise da documentação dos sistemas de coleta de dados do MEC: EDUCACENSO, Censo do Ensino Superior e SISTEC, assim como a partir das possíveis questões levantadas para a geração dos indicadores existentes no TAMC (Termo de Acordo de Metas e Compromissos). Desta documentação foram extraídas as questões de competência iniciais, a partir da análise dos layouts dos arquivos de importação de dados, modelos de dados e descrição de indicadores. Desta maneira, foram identificadas inicialmente setenta e oito questões de competência.

A busca pelas respostas às questões levantadas preliminarmente utilizou os mesmos documentos já utilizados para a elaboração das questões, entretanto, contou com o acréscimo da utilização do *thesaurus* da educação brasileira. Nesta fase evidenciou a diferença existente entre os sistemas de coleta de dados do MEC, uma vez que algumas perguntas apresentaram respostas diferentes a partir de documentações de sistemas de coletas de dados diferentes. Um exemplo desta situação refere-se á questão de competência de número 5 presente no Quadro 6, que apresentou respostas diferenciadas extraídas da documentação do EDUCACENSO e SISTEC.

A atividade de Validação e Uniformização das Respostas teve o objetivo de tratar as divergências apresentadas entre as respostas para as questões de competência levantadas. Para tanto, foi necessário estabelecer uma estratégia para a busca do consenso em relação às respostas que se enquadrassem nesta situação. Assim, foi utilizada a Lei 9.394/96, a Lei de Diretrizes e Bases – LDB além de consultas a especialistas do domínio dos registros acadêmicos brasileiros. Os especialistas em questão foram servidores públicos lotados na Pró-Reitoria de Ensino do Instituto Federal de Educação, Ciência e Tecnologia da Bahia – IFBA. Para cada questão identificada que apresentou respostas divergentes, buscou-se no texto da LDB a resposta que fosse definitiva, e em caso de dúvida quanto à interpretação da lei ou até mesmo omissão da legislação, a consulta era feita aos especialistas do IFBA.

A atividade de Refinamento ocorre logo após a Validação e Uniformização das Respostas. A atividade consiste em avaliar cada questão de competência considerando o escopo e os objetivos já definidos no OSRD da ONTO-RA. Desta forma, após a avaliação de cada questão, foram descartadas as questões de competência que não estavam diretamente ligadas à geração de indicadores gerenciais e de desempenho acadêmicos, restando, assim, quarenta e três questões de competência. Estas questões ainda foram agrupadas de acordo com as características em: (1) Atividade Acadêmica, (2) Curso, (3) Oferta de Vagas, (4) Organização Acadêmica, (5) Pessoa, (6) Situação Acadêmica e (7) Vínculo Institucional. As questões de competência, suas respectivas respostas e grupos podem ser vistas no Quadro 6.

Quadro 6 - Lista de questões de competência por grupo

Nº	QUESTÃO DE COMPETÊNCIA	RESPOSTA	Grupo
1	Quais são os tipos de atividades extracurriculares para alunos?	Pesquisa, Bolsa para pesquisa, Extensão, Bolsa para Extensão, Monitoria, Bolsa monitoria, Estágio não obrigatório, Bolsa para atividade extracurricular	Atividade Acadêmica
2	Quais as possíveis disciplinas?	Química Física Matemática Biologia Ciências Língua /Literatura Portuguesa Língua /Literatura estrangeira - Inglês Língua /Literatura estrangeira - Espanhol Língua /Literatura estrangeira - outra Arte (Educação Artística, Teatro, Dança, Música, Artes Plásticas e outras) Educação Física História Geografia Filosofia Informática/Computação Disciplinas profissionalizantes Disciplinas voltadas ao atendimento às necessidades educacionais específicas dos alunos que são público alvo da educação especial e às práticas educacionais inclusivas. Disciplinas voltadas à diversidade sociocultural (disciplinas pedagógicas) Libras Disciplinas pedagógicas Ensino religioso Língua indígena Estudos Sociais Sociologia Língua/Literatura estrangeira - Francês Outras Disciplinas	Atividade Acadêmica
3	Quais são os possíveis turnos de estudo do aluno ?	Matutino, Vespertino, Noturno, Integral	Curso
4	Quais as modalidades de oferta de curso para alunos ?	Presencial e EAD	Curso
5	Quais os possíveis níveis de ensino de cursos ?	Educação Básica, Educação Superior, Formação Inicial e Continuada(FIC)	Curso
6	Quais as possíveis etapas de ensino da educação básica ?	Educação Infantil - CRECHE, Educação Infantil - PRÉ-ESCOLA, Ensino Fundamental, Ensino Médio, Ensino Médio - EPT(Educação profissional média de nível médio), Ensino Médio – Normal Magistério	Curso
7	Quais as formas de oferta do Ensino Médio – EPT ?	Articulada INTEGRADA, Articulada Concomitante e subseqüente	Curso
8	Quais os níveis acadêmicos dos cursos da Educação Superior ?	Sequencial, graduação, pós-graduação, extensão, Pós-graduação tecnológica	Curso

N°	QUESTÃO DE COMPETÊNCIA	RESPOSTA	Grupo
9	Quais os graus acadêmicos dos cursos da Educação Superior no nível acadêmico graduação ?	Bacharelado, Licenciatura e Tecnológico	Curso
10	Quais os programas de curso Pós-graduação e Pós-graduação tecnológica ?	Mestrado, Mestrado Profissional, Doutorado, Especialização e Aperfeiçoamento	Curso
11	Quais as possíveis modalidades de ensino ?	Educação Regular, Educação Especial e EJA (Válido para cursos da educação básica, exceto educação infantil)	Curso
12	Quais as possíveis situações dos cursos ?	Em atividade e Extinto	Curso
13	O que se entende por inscrições de candidatos em cursos ?	Número de inscrições para as vagas ofertadas, por quaisquer formas de ingresso (Sisu, Enem, vestibular, processos seletivos, sorteios, ...) por curso e campus. É importante o registro do total de inscrições para qualquer oferta da Instituição, pois será utilizada juntamente com as vagas ofertadas para o cálculo do indicador “candidato/vaga”	Oferta de vagas
14	O que se entende por vagas ofertadas para curso ?	Número de vagas ofertadas, por curso e campus, em Editais de Oferta de Vagas por meio do Sisu, Enem, vestibular, processos seletivos, sorteios e/ou outras formas de ingresso. Igualmente importante o registro do total de vagas ofertadas pela Instituição, pois será utilizada juntamente com as inscrições realizadas para o cálculo do indicador “candidato/vaga”.	Oferta de vagas
15	O que se entende por alunos ingressantes em curso?	Número de alunos ingressos por meio de vestibular, processos seletivos e outras formas de ingresso	Oferta de vagas
16	O que se entende por turma?	Grupo de alunos que estão vinculados a uma mesma disciplina ou grupo de disciplinas num mesmo horário e ministrada pelo mesmo docente.	Organização Acadêmica

N°	QUESTÃO DE COMPETÊNCIA	RESPOSTA	Grupo
17	O que se entende por ciclo de matrícula de aluno em curso?	Um ciclo de matrícula é um conceito associado à oferta de curso, no qual um conjunto de alunos irá ingressar visando a obtenção de uma mesma certificação ou diploma, tendo o curso uma mesma carga horária, data de início e data de fim prevista. O conceito de ciclo de matrícula está ligado à oferta de cursos e não de turmas. Os alunos de diferentes turmas que iniciam um curso de mesma certificação e mesma carga horária, numa mesma data podem pertencer a um mesmo ciclo de matrículas. O termo ciclo de matrículas representa uma visão relativa a dois momentos do aluno no curso, que são a sua entrada no curso (situação inicial) e a sua saída do curso (situação final). Esta última pode ser por conclusão, evasão, desligamento ou transferência.	Organização Acadêmica
18	O que se entende por turma de ingresso?	Identificação para um grupo de alunos que ingressaram no mesmo curso no mesmo período letivo	Organização Acadêmica
19	Quais são as possíveis cor / raça de alunos e docentes? (docentes e alunos)	Branca, Preta, Parda, Amarela, Indígena, Não dispõe, Não declarado.	Pessoa
20	Quais as possíveis nacionalidades de alunos e docentes? (docentes e alunos)	Brasileira, Brasileira nascido no exterior ou naturalizado, estrangeira.	Pessoa
21	Quais os possíveis tipos de deficiência físicas, intelectuais ou sensoriais de alunos e docentes?	Cegueira, Baixa visão, Surdez, Deficiência auditiva, Deficiência física, Surdo cegueira, Deficiência Múltipla, Deficiência Intelectual, Autismo, Síndrome de Asperger, Síndrome de RETT, Transtorno desintegrativo da infância, Altas habilidades / superdotação.	Pessoa
22	Quais os possíveis tipos de origem escolar do ensino médio para alunos?	Privada, Pública, Não dispõe da informação	Pessoa
23	Quais as possíveis escolaridades para os docentes ?	Fundamental incompleto, Fundamental Completo, Ensino Médio, Superior	Pessoa
24	Quais as possíveis pós-graduações dos docentes?	Especialização, Mestrado, Doutorado, Nenhum.	Pessoa
25	Quais os tipos de zonas de residência de alunos e docentes de curso?	Urbana e Rural	Pessoa
26	O que se entende por tempo de integralização de aluno no curso?	Tempo mínimo descrito em anos ou fração a duração do curso para que o estudante receba a formação pretendida	Situação Acadêmica

N°	QUESTÃO DE COMPETÊNCIA	RESPOSTA	Grupo
27	Quais os possíveis tipos de rendimento para um aluno?	Aprovado, Reprovado	Situação Acadêmica
28	O que se entende por aluno concluinte de curso?	Aluno que concluiu os créditos e está apto a colar grau.	Situação Acadêmica
29	O que se entende por aluno integralizado em curso?	Aluno que concluiu os créditos, entretanto, não fizeram o estágio curricular ou equivalente obrigatório para estar apto a colar grau	Situação Acadêmica
30	O que se entende por retenção de aluno em curso ?	Alunos que trancaram ou foram reprovados num determinado período letivo	Situação Acadêmica
31	O que se entende por evasão de aluno em curso?	Número de alunos que efetivamente matriculados no período letivo anterior que não concluíram o curso e não se encontram efetivamente matriculados no período letivo atual.	Situação Acadêmica
32	Quais as possíveis situações de vínculo do aluno?	Cursando / Em andamento, Matrícula Trancada, Desvinculado, Transferido Interno, Formado, Falecido, Mobilidade Acadêmica, Concluinte, Deixou de Frequentar / Evadido, Transferido Externo.	Vínculo Institucional
33	Quais os tipos de mobilidade acadêmica de aluno?	Nacional, Internacional	Vínculo Institucional
34	Quais os tipos de mobilidade acadêmica internacional de aluno?	Intercâmbio, Ciência sem Fronteiras.	Vínculo Institucional
35	Aluno Parfor (plano nacional de formação de professores da educação básica)?	Aluno que ingressou através do plano nacional de formação de professores da educação básica	Vínculo Institucional
36	Quais as possíveis forma de ingresso de alunos?	Vestibular, ENEM, Avaliação Seriada, Outros tipos, PEC-G(programa de estudantes-convênio de graduação), Transferência Ex-Ofício, Decisão Judicial, Transferência Interna, Transferência Externa, Portador de Diploma, Reingresso, Sem processo seletivo, sorteio, Exame de Seleção.	Vínculo Institucional
37	Quais os tipos de programa de reserva de vagas de alunos?	Etnico, Pessoas com deficiência, Estudante procedente de escola pública, Renda Familiar, Outros,	Vínculo Institucional
38	Quais os tipos de apoio social a alunos?	Alimentação, Moradia, Transporte, Material didático, bolsa trabalho, bolsa permanência.	Vínculo Institucional
39	Quais as possíveis situações de vínculo de docentes?	Em exercício, Afastado para qualificação, Afastado para exercício em outros órgãos, afastado por outros motivos, afastado para tratamento de saúde.	Vínculo Institucional
40	Quais os regimes de vínculo para docentes?	Tempo Integral com DE, Tempo Integral sem DE, Tempo Parcial, Horista.	Vínculo Institucional

N°	QUESTÃO DE COMPETÊNCIA	RESPOSTA	Grupo
41	Quais os possíveis tipos de vínculo de docentes?	Efetivo, Substituto, Visitante, Terceirizado, CLT.	Vínculo Institucional
42	Quais os possíveis tipos de atuação do docente?	Ensino em curso sequencial de formação específica, Ensino em curso de graduação presencial, Ensino de curso de graduação a distância, Ensino de pós-graduação stricto sensu presencial, Ensino de pós-graduação stricto sensu a distancia, pesquisa, extensão, gestão planejamento e avaliação, bolsa pesquisa, Ensino Médio, Ensino Fundamental.	Vínculo Institucional
43	O que é regime de internato pleno de alunos em cursos?	Regime em que os alunos residem na instituição de ensino.	Vínculo Institucional

5.3.2 Extração da Terminologia e Frequência

A atividade de Extração da Terminologia e Frequência dos termos teve por objetivo identificar os termos de maior frequência dentre as perguntas e respostas das questões de competência, que vão servir como base para os próximos processos que irão identificar os recursos não ontológicos que passarão pelo processo de reengenharia com o objetivo de gerar a ONTO-RA. Além disso, são identificados os termos que são classificados como Objetos do Universo do Discurso, que se tratam dos termos ligados a outros termos, que serão modelados como instâncias.

O conjunto de questões de competência listado no Quadro 6 gerou um total de duzentos e setenta e cinco termos identificados, entre perguntas, respostas e objetos do universo do discurso. O Quadro 7 apresenta os termos identificados e o número de ocorrências de cada um deles nas perguntas das questões de competência, além de indicar se o termo é ou não um objeto do universo do discurso.

Quadro 7 - Termos extraídos das perguntas das questões de competência

PERGUNTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Aluno	28	Não
Curso	17	Não
Docente	9	Não
Situação de vínculo	2	Não
Nível acadêmico	2	Não
Carga horária	2	Não
Cor / Raça	1	Não
Nacionalidade	1	Não
Deficiência física sensorial ou intelectual	1	Não
Turno	1	Não
Mobilidade acadêmica	1	Não
Mobilidade acadêmica internacional	1	Não
Parfor	1	Não
Origem escolar do ensino médio	1	Não
Forma de Ingresso	1	Não
Programa de reserva de vagas	1	Não
Apoio Social	1	Não
Atividades extracurriculares	1	Não
Escolaridade	1	Não
Pós-graduação	1	Não
Regime de vínculo	1	Não
Tipo de vínculo	1	Não
Tipo de atuação	1	Não
Tempo de Integralização	1	Não
Modalidade de Oferta	1	Não
Nível de ensino	1	Não
Etapas de ensino	1	Não
Educação básica	1	Não
Forma de oferta	1	Não
Ensino Médio – EPT	1	Não
Grau acadêmico	1	Não
Educação superior	1	Não
Graduação	1	Não
Programa de curso	1	Não
Pós-graduação	1	Não
Pós-graduação tecnológica	1	Não
Modalidade de ensino	1	Não
Situação de curso	1	Não
Turma	1	Não
Tipo de rendimento	1	Não
Inscrições	1	Não
Candidatos	1	Não
Vagas ofertadas	1	Não
Ingressantes	1	Não
Matrícula	1	Não
Concluinte	1	Não
Integralização	1	Não
Retenção	1	Não

PERGUNTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Evasão	1	Não
Ciclo de Matrícula	1	Não
Certificação intermediária	1	Não
Disciplina semipresencial	1	Não
Semipresencial	1	Não
Disciplina	1	Não
Zona de Residência	1	Não
Internato pleno	1	Não
Turma de Ingresso	1	Não

O Quadro 8 apresenta os termos identificados e o número de ocorrências de cada um deles nas respostas das questões de competência e é o último item do OSRD. Assim como no Quadro 8, também é apresentada a indicação se o termo é ou não um objeto do universo do discurso.

Quadro 8 - Termos extraídos das respostas das questões de competência

RESPOSTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Curso	14	Não
Aluno	9	Não
Ensino	6	Não
Vagas	6	Não
Vestibular	4	Sim
Outras formas de ingresso	4	Sim
ciclo de matrícula	4	Não
ENEM	3	Sim
Processo Seletivo	3	Não
Sorteio	3	Sim
Extensão	3	Sim
Ensino Médio	3	Não
Ofertadas	3	Não
Carga-horária	3	Não
Ingressou	2	Não
Portador de Diploma	2	Não
Pesquisa	2	Não
Bolsa para pesquisa	2	Sim
Especialização	2	Sim
Mestrado	2	Sim
Doutorado	2	Sim
Ensino fundamental	2	Sim
Estudante	2	Sim
Presencial	2	Sim
Educação Básica	2	Sim

RESPOSTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Graduação	2	Sim
Reprovado	2	Sim
Inscrições	2	Não
Campus	2	Sim
Instituição	2	Sim
Candidato	2	Não
Créditos	2	Não
Apto a colar grau	2	Sim
Certificação	2	Sim
Turma	2	Não
Horas-aula	2	Não
atividade científico-acadêmica	2	Sim
Branca	1	Sim
Parda	1	Sim
Preta	1	Sim
Amarela	1	Sim
Indígena	1	Sim
Brasileira	1	Sim
Brasileira nascido no exterior / naturalizado	1	Sim
Estrangeira	1	Sim
Cegueira	1	Sim
Baixa visão	1	Sim
Surdez	1	Sim
Deficiência auditiva	1	Sim
Deficiência física	1	Sim
Surdocegueira	1	Sim
Deficiência Múltipla	1	Sim
Deficiência Intelectual	1	Sim
Autismo	1	Sim
Síndrome de Asperger	1	Sim
Síndrome de RETT	1	Sim
Transtorno Desintegrativo da Infância	1	Sim
Altas habilidades / superdotação	1	Sim
Matutino	1	Sim
Vespertino	1	Sim
Noturno	1	Sim
Integral	1	Sim
Cursando / Em andamento	1	Sim
Matrícula Trancada	1	Sim
Desvinculado	1	Sim
Transferido interno	1	Sim
Formado	1	Sim
Falecido	1	Sim
Mobilidade Acadêmica	1	Sim
Concluinte	1	Sim
Deixou de freqüentar	1	Sim
Evadido	1	Sim
Transferido externo	1	Sim
Nacional	1	Sim

RESPOSTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Internacional	1	Sim
Intercâmbio	1	Sim
Ciência sem fronteira	1	Sim
plano nacional de formação de professores da educação básica	1	Sim
Privada	1	Sim
Pública	1	Sim
Avaliação Seriada	1	Sim
PEC-G	1	Sim
Transferência ex-ofício	1	Sim
Decisão Judicial	1	Sim
Transferência Interna	1	Sim
Transferência Externa	1	Sim
Reingresso	1	Sim
Sem Processo Seletivo	1	Sim
Exame de Seleção	1	Sim
Étnico	1	Sim
Pessoas com deficiência	1	Sim
Estudante procedente de escola pública	1	Sim
Renda familiar	1	Sim
Outros tipos de reserva de vaga	1	Sim
Alimentação	1	Sim
Moradia	1	Sim
Transporte	1	Sim
Material didático	1	Sim
bolsa trabalho	1	Sim
bolsa permanência	1	Sim
Bolsa para extensão	1	Sim
Monitoria	1	Sim
Bolsa monitoria	1	Sim
Estágio não obrigatório	1	Sim
Bolsa atividade extracurricular	1	Sim
Fundamental incompleto	1	Sim
Fundamental completo	1	Sim
Ensino Superior	1	Sim
Em exercício	1	Sim
Afastado para qualificação	1	Sim
afastado para exercício em outros órgãos	1	Sim
afastado por outros motivos	1	Sim
afastado para tratamento de saúde	1	Sim
Tempo integral com DE	1	Sim
Tempo integral sem DE	1	Sim
Tempo parcial	1	Sim
Horista	1	Sim
Efetivo	1	Sim
Substituto	1	Sim
Visitante	1	Sim
Terceirizado	1	Sim
CLT	1	Sim

RESPOSTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Seqüencial de formação específica	1	Sim
Graduação presencial	1	Sim
Graduação a distância	1	Sim
Pós-graduação stricto senso presencial	1	Sim
Pós-graduação stricto senso a distância	1	Sim
Gestão	1	Sim
Planejamento	1	Sim
Avaliação	1	Sim
bolsa pesquisa	1	Sim
Tempo mínimo	1	Sim
Formação	1	Sim
EAD	1	Sim
Educação Superior	1	Sim
Formação Inicial e Continuada(FIC)	1	Sim
Educação Infantil – Creche	1	Sim
Educação Infantil - Pré-escola	1	Sim
Ensino Médio - EPT(Educação Profissional Técnica de Nível Médio)	1	Sim
Ensino Médio - Normal Magistério	1	Sim
Articulada Integrada	1	Sim
Articulada Concomitante	1	Sim
Subseqüente	1	Sim
Seqüencial	1	Sim
Pós-graduação	1	Sim
Pós-graduação tecnológica	1	Sim
Bacharelado	1	Sim
Licenciatura	1	Sim
Tecnológico	1	Sim
mestrado profissional	1	Sim
Aperfeiçoamento	1	Sim
educação regular	1	Sim
educação especial	1	Sim
EJA	1	Sim
educação infantil	1	Sim
em atividade	1	Sim
Extinto	1	Sim
Aprovado	1	Sim
Forma de ingresso	1	Sim
total de inscrições	1	Sim
Edital	1	Sim
Oferta	1	Sim
Sisu	1	Sim
total de vagas	1	Sim
Matriculado	1	Sim
Concluiu	1	Sim
estágio curricular	1	Sim
Trancaram	1	Sim
período letivo	1	Sim
Diploma	1	Sim

RESPOSTAS

Termo	Total de ocorrências	Objetos do Universo do Discurso
Data início	1	Sim
Data fim prevista	1	Sim
Data	1	Sim
situação inicial	1	Sim
situação final	1	Sim
Conclusão	1	Sim
Evasão	1	Sim
Desligamento	1	Sim
Transferência	1	Sim
Disciplina	1	Sim
Professor	1	Sim
Semipresencial	1	Sim
MEC	1	Sim
Química Física Matemática Biologia Ciências Língua /Literatura Portuguesa Língua /Literatura estrangeira - Inglês Língua /Literatura estrangeira - Espanhol Língua /Literatura estrangeira - outra Arte (Educação Artística, Teatro, Dança, Música, Artes Plásticas e outras) Educação Física História Geografia Filosofia Informática/Computação Disciplinas profissionalizantes Disciplinas voltadas ao atendimento às necessidades educacionais específicas dos alunos que são público alvo da educação especial e às práticas educacionais inclusivas. Disciplinas voltadas à diversidade sociocultural (disciplinas pedagógicas) Libras Disciplinas pedagógicas Ensino religioso Língua indígena Estudos Sociais Sociologia Língua/Literatura estrangeira - Francês Outras Disciplinas	1	Sim
Urbana	1	Sim
Rural	1	Sim
Residem	1	Sim

5.4 BUSCA POR RECURSOS NÃO ONTOLÓGICOS PARA A ONTO-RA

Com base no OSRD da ONTO-RA, foi iniciada a tarefa de Busca por Recurso não Ontológico. Cujo objetivo foi identificar os potenciais candidatos a recurso não ontológico que poderá ser submetido à avaliação e seleção. Durante a busca por recursos não ontológicos para o desenvolvimento da ontologia de registros acadêmicos do modelo educacional brasileiro foram levantadas informações a cerca dos sistemas de coletas de dados utilizados pelo MEC e outros levantamentos realizados por órgãos de controle ou pelo próprio MEC com o intuito de gerar indicadores gerencias ou de desempenho acadêmico das instituições federais de ensino, além do levantamento da legislação que regula a educação e os registros acadêmicos no Brasil. A seguir estão listados todos os candidatos a recursos não ontológicos identificados:

- Documentação do Censo do Ensino Superior, composta de Glossário, Manual e Leiaute de migração;
- Diagnóstico de Evasão e Retenção do MEC;
- Documentação do EDUCACENSO, composta de Manual e Leiaute;
- Lei nº 9.394, de 20 de dezembro de 1996(LDB);
- Manual do SISTEC;
- Termo de Acordo de Metas e Compromissos - TAMC.

Tais recursos passaram a ser considerados candidatos a recursos não ontológicos na medida em que foram identificadas em cada um deles, respostas aos principais termos identificados no OSRD. Estes principais termos são os de maior frequência, que são os listados nos Quadros 7 e 8, acrescidos de termos que foram considerados relevantes, mesmo com frequência mais baixa, como “Nível de Ensino”, “Etapas de Ensino”, “Forma de Oferta” e “Modalidade de Ensino”. A relevância atribuída a estes termos menos frequentes se deveu ao fato dos mesmos serem importantes para a definição da classificação taxonômica dos cursos, fruto de divergências conhecidas entre os diversos sistemas de coleta de dados do MEC.

Dentro desta lógica, a utilização do Thesaurus Brasileiro da Educação foi descartada, uma vez que este recurso, apesar de apresentar respostas para alguns dos termos de maior frequência identificados no OSRD listados nos Quadros 7 e 8, como “Aluno”, “Curso” e “Docente”, não apresentou respostas para conceitos de maior frequência, como “Situação de Vínculo” e “Nível Acadêmico” e para conceitos

menos frequentes, mas classificados como relevantes, como “Nível de Ensino”, “Etapas de Ensino”, “Forma de Oferta” e “Modalidade de Ensino”.

O Censo do Ensino Superior, o EDUCACENSO e o SISTEC, são sistemas de coleta de dados utilizados pelo MEC para obter informações a cerca das instituições de ensino brasileiras. O SISTEC é ainda utilizado para definição de matriz orçamentária e gerenciamento de aplicação de recursos e oferta de vagas para cursos através de programas custeados diretamente pelo governo federal como o PROFUNCIÁRIO e o PRONATEC.

O Diagnóstico de Evasão e Retenção realizado pelo MEC foi motivado pelo relatório de auditoria operacional nº TC 026.062/2011-9 que originou o acórdão nº 506/2013 que recomenda a adoção de medidas e a implementação de ações conjuntas entre a SETEC / MEC e as instituições da Rede Federal de Educação Profissional e Tecnológica. Uma das ações originadas do acórdão 506/2013 é o diagnóstico de evasão e retenção a ser realizado pelo MEC nas instituições da Rede Federal de Educação Profissional e Tecnológica, que se baseia em indicadores especificados pelo próprio acórdão para realizar o levantamento da situação de evasão e retenção de alunos naquelas instituições.

A lei nº 9.394, de 20 de dezembro de 1996, conhecida como Lei de Diretrizes e Bases da Educação estabelece os princípios e finalidades, além de determinar a forma como a organização nacional está organizada em seus diversos níveis, modalidades.

5.5 AVALIAÇÃO DO CONJUNTO DE CANDIDATOS A RECURSOS NÃO ONTOLÓGICOS

A atividade de avaliação de candidatos a recursos não ontológicos utilizou como entrada o conjunto de candidatos a recursos não ontológicos - NORS identificados na atividade de busca de recursos não ontológicos. Tais recursos foram avaliados segundo dois critérios estabelecidos pela metodologia NEON, que foram Precisão e Consenso. O cálculo do critério precisão foi feito mediante a aplicação da fórmula estabelecida pela metodologia NEON, conforme apresentado na Seção 4.6.2. A aferição do consenso, que é um critério subjetivo, foi verificada a partir de consulta à LDB e aos profissionais especialistas do domínio do IFBA.

O Quadro 9 mostra a tabela de avaliação dos candidatos a recursos não ontológicos identificados na atividade de busca por recursos não ontológicos realizada para a construção da ontologia para o domínio dos registros acadêmicos do modelo brasileiro de educação.

Quadro 9 - Tabela de avaliação dos candidatos a NORs

Candidato a NOR	Cobertura	Consenso
Documentação do Censo do Ensino Superior	0,2	Não
Diagnóstico de Retenção e Evasão	0,1	Não
Documentação do EDUCACENSO	0,2	Não
LDB	0,2	Sim
Manual do SISTEC	0,1	Não
TAMC	0,2	Não

A metodologia NEON prevê o cálculo do nível de precisão dos recursos não ontológicos para compor a tabela de avaliação dos candidatos a NORs. Tal medida representa a proporção de entradas léxicas do candidato a recurso não ontológico que estão entre os termos identificados no OSRD. A tabela de avaliação dos candidatos a NORs da ONTO-RA foi construída sem o cálculo deste critério, que visa posicionar o recurso não ontológico em relação à quantidade de termos úteis identificados. Como os candidatos a recursos não ontológicos identificados neste trabalho estão desestruturados, é natural que muitos dos termos encontrados nos textos não estejam presentes no OSRD. Além disso, alguns documentos analisados como NORs, são específicos de um subdomínio da educação, como por exemplo, o Censo do Ensino Superior trata de termos utilizados no universo do ensino superior. Desta forma avaliou-se que seria suficiente a análise do consenso aliada com a cobertura que é a proporção de respostas que o candidato a NOR foi capaz de oferecer, ou seja, determina quanto o recurso atende aos requisitos levantados.

5.6 SELEÇÃO DO RECURSO NÃO ONTOLÓGICO MAIS APROPRIADO

A partir da análise dos resultados apresentados no Quadro 9, percebe-se que o nível de cobertura de todos os candidatos a NORs não ultrapassou 20 % dos

termos identificados no ORSD. É possível verificar, ainda, que apenas os termos identificados na LDB possuem consenso em relação ao domínio dos registros acadêmicos brasileiros. Outra característica dos candidatos a NORS é que apenas o glossário do Censo do Ensino Superior se encaixa em um dos tipos de recursos não ontológicos utilizados pela metodologia NEON, entretanto, tal documento consiste apenas em uma parte de um candidato a NOR. Segundo (SUÁREZ-FIGUEROA, 2008), durante o processo de Reengenharia de Recursos não Ontológicos, um recurso poderá ser reconstruído, caso não seja possível extrair o seu modelo de dados. Nas atividades de Transformação de Recursos não Ontológicos, são utilizados padrões de transformação que são identificados a partir do tipo de recurso não ontológico, o seu modelo de dados e as semânticas das relações entre os termos.

Neste contexto, com o objetivo de buscar o nível de cobertura de 100 % dos termos da ORSD, estruturar os recursos não ontológicos de acordo com o preconizado pela metodologia e garantir um consenso sobre os recursos, decidiu-se fundir os candidatos de recursos não ontológicos, identificando todos os termos cobertos pela ORSD e ao mesmo tempo utilizando a LDB para dirimir eventuais conflitos semânticos existentes entre os diversos recursos. Analisando os tipos de recursos não ontológicos definidos pela metodologia NEON e o relacionamento entre os termos representados nestes recursos, optou-se por reestruturar os recursos não ontológicos utilizando uma estrutura de esquema de classificação, que segundo (HEPP, 2007) são a contribuição mais valiosa para a criação de ontologias em diversos domínios. Os esquemas de classificação contêm uma grande variedade de definições de categoria, assim como uma hierarquia, além de refletir um alto grau de consenso.

A metodologia NEON utiliza diferentes formas de representação do modelo de dados para os mais diversos tipos de recurso não ontológico. Valendo a mesma situação para a implementação dos recursos. Para o caso específico dos esquemas de classificação esta representação poderá ser quanto ao modelo de dados:

1. *Path Enumeration* (BRANDON, 2005), que é uma estrutura recursiva para representação de hierarquia definida como um modelo que armazena para cada nó, o caminho para o nó raiz. *Adjacency List* (BRANDON, 2005), que também é uma estrutura recursiva de

representação hierárquica, compreendida por uma lista de nós com uma coluna com *links* para os nós pais.

2. *Snowflake*, que é uma estrutura normalizada para representação hierárquica. Para cada nível hierárquico, uma tabela é criada. Neste modelo, cada nó hierárquico possui uma coluna que faz o *link* para o seu nó pai.
3. *Flattened*, que é uma estrutura desnormalizada de representação hierárquica. A hierarquia é representada usando uma tabela onde cada nível hierárquico é armazenado em uma coluna diferente.

De acordo com a implementação, os esquemas de classificação podem ser base de dados relacionais, arquivos XML e planilhas eletrônicas. Optou-se por utilizar o modelo de dados *snowflake* e implementar em uma planilha eletrônica. Tal escolha se deveu ao fato de que o modelo de dados *snowflake*, por ser normalizado apresenta uma estruturação mais simples e de mais fácil compreensão e visualização quando comparadas com as demais possibilidades de modelo de dados. As tabelas do esquema de classificação resultante da fusão dos candidatos a recursos não ontológicos são mostrados no ANEXO A.

5.7 TRANSFORMAÇÃO DE RECURSO NÃO ONTOLÓGICO EM ONTO-RA

A atividade de Transformação de Recurso não Ontológico consistiu na busca por um padrão de reengenharia de recurso não ontológico que fosse adequado para realizar a transformação do esquema de classificação disponível no ANEXO A em ontologia leve. O esquema de classificação citado foi representado com o modelo de dados *snowflake*. Além dos relacionamentos subclasse representados no esquema de classificação, foram identificados relacionamentos *ad-hoc*, portanto, a semântica dos relacionamentos é misto. A partir das características do NOR, foi identificado um padrão de projeto de reengenharia adequado a estas características.

A metodologia NEON possui padrões específicos de reengenharia para cada conjunto de características de tipo de dados, modelo de dados, implementação e tipos de relacionamento dos recursos não ontológicos. O padrão identificado como mais adequado ao desenvolvimento da ONTO-RA, foi o PR-NOR-CLLO-03 (VILLAZÓN-TERRAZAS, 2009) que é um padrão para realizar a reengenharia de um esquema de classificação cujo modelo de dados é *snowflake* e cujas relações entre

os objetos do recurso são do tipo mista. O padrão de reengenharia PR-NOR-CLLO-03 determina a execução dos seguintes passos para transformar o esquema de classificação em ontologia leve:

1. Identificar os diferentes grupos de G_i , e para cada grupo criar classes para cada entidade C_i ;
2. Se existir um relacionamento entre as entidades, criá-los como uma relação binária ad-hoc, $r_i \in R$;
3. Se existir uma superclasse para a nova entidade relacionada, ela deverá ser criada e uma relação subclasse deverá ser criada entre a classe e a superclasse.

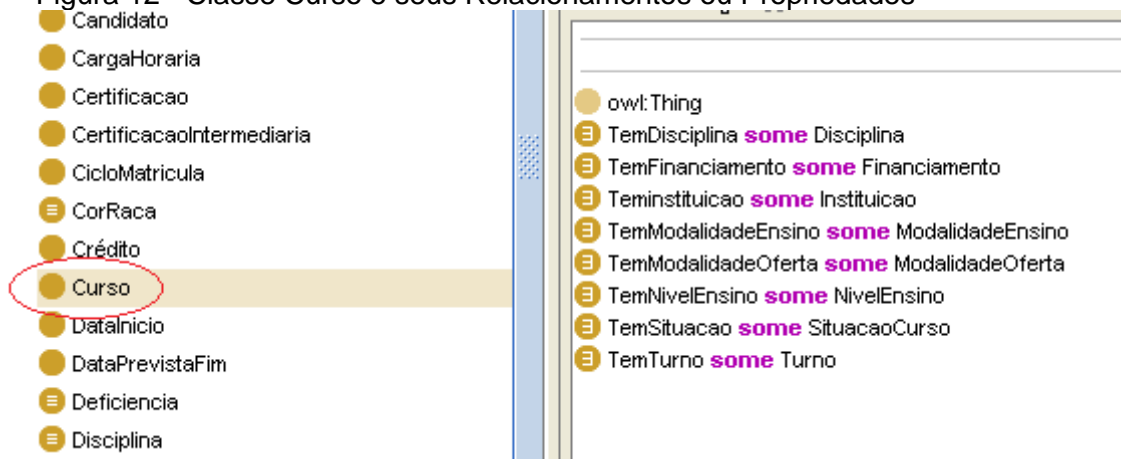
Após a execução dos passos citados anteriormente, o esquema de classificação definido a partir da junção dos candidatos a recursos não ontológicos identificados no processo de construção da ontologia para o domínio dos registros acadêmicos do modelo brasileiro de educação, foi transformado para ontologia leve. A ontologia leve gerada está disponível no ANEXO B.

5.8 ENGENHARIA *FORWARD* DE ONTOLOGIA EM ONTO-RA

O intuito da tarefa de Engenharia *Forward* na construção da ONTO-RA foi gerar uma nova implementação da ontologia leve disponível no ANEXO B, portanto, com base naquele conceitual gerado foi utilizada a lógica de descrição, através da linguagem OWL DL. Para a execução da tarefa de Formalização e a Implementação foi feita utilizando a linguagem OWL. Em ambas as tarefas utilizou-se a ferramenta Protegè.

O produto gerado pelo processo apresentado neste trabalho é a Ontologia de Registro Acadêmico do Modelo Brasileiro de Educação - ONTO-RA. Esta ontologia possui 65 classes e 117 instâncias. Tais classes e suas restrições e relacionamentos representam a taxonomia do domínio dos registros acadêmicos e está disponível na sua totalidade em <http://xbrlframework.wordpress.com/>. A Figura 12 apresenta a captura de parte da tela do editor Protegé que ilustra a classe Curso e seus relacionamentos que também são conhecidos como propriedades.

Figura 12 - Classe Curso e seus Relacionamentos ou Propriedades



As propriedades ou relacionamentos apresentados na coluna do lado direito da Figura 12, são do tipo Restrição Existencial, o que significa dizer que para que um indivíduo possa ser considerado instância da classe curso, é necessário que ele possua ao menos um relacionamento com indivíduos das classes Disciplina, Financiamento, Instituição, ModalidadeEnsino, ModalidadeOferta, NivelEnsino, SituacaoCurso e Turno.

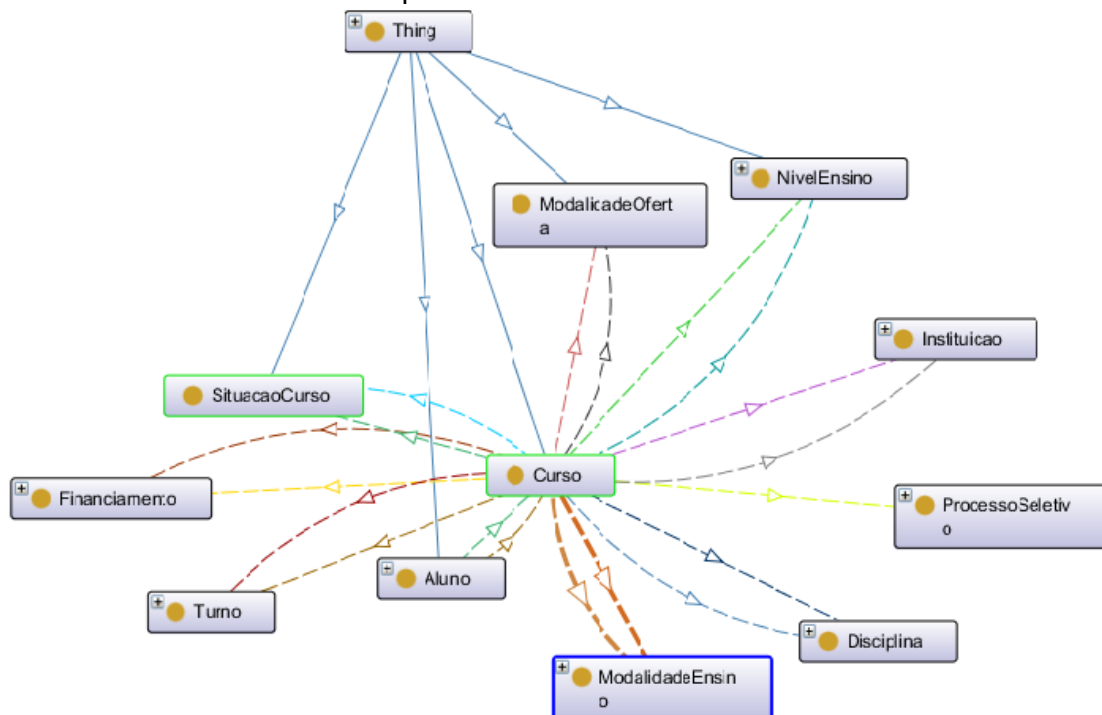
Como citado na Seção 5.4, termos como Nível de Ensino, Etapa de Ensino, Forma de Oferta, e Modalidade de Ensino, apesar de não apresentarem uma frequência significativa entre os termos extraídos do OSRD, foram considerados relevantes, uma vez que tratavam de definições da estrutura taxonômica dos cursos na educação brasileira, que é objeto de inconsistência entre os diversos sistemas de coleta de dados do MEC. A Figura 13 apresenta a definição destas classes através da captura de um fragmento da tela do editor Protegé.

Figura 13 - Estrutura Hierárquica dos Cursos em ONTO-RA



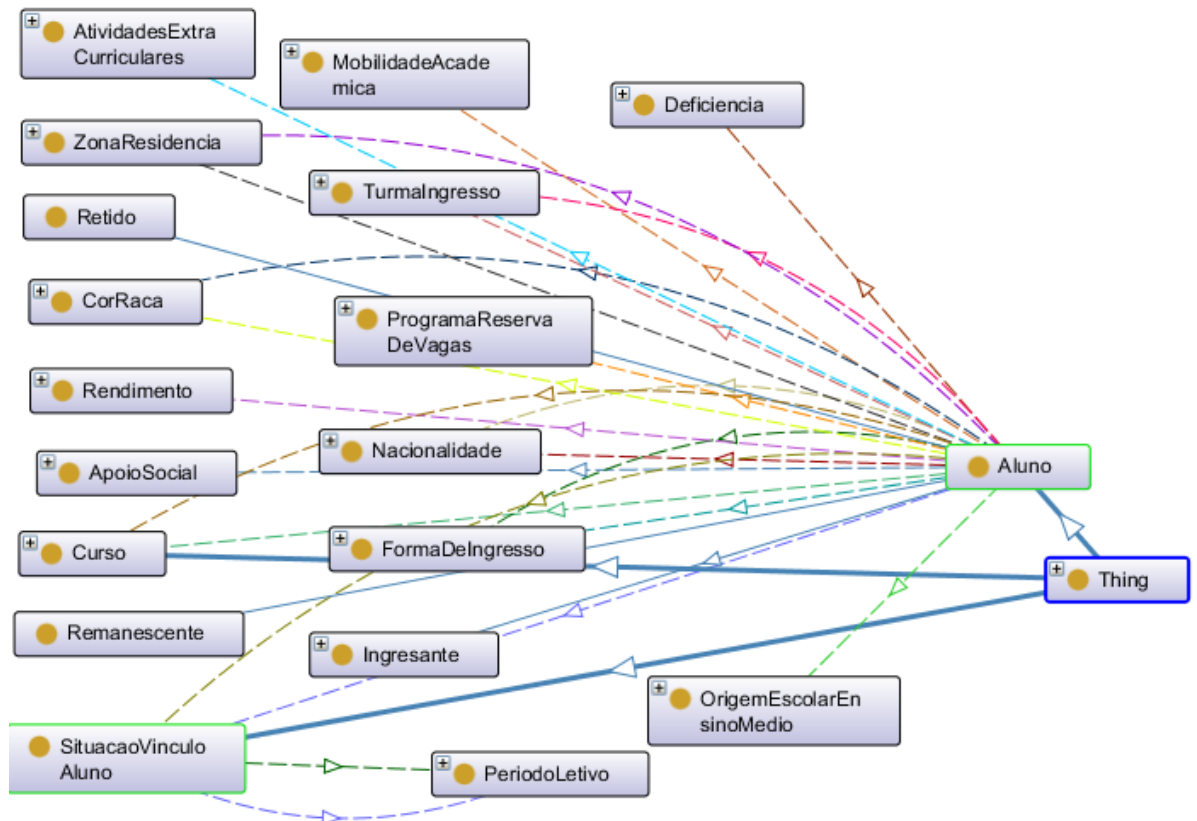
A estrutura hierárquica das classes representadas na Figura 13 poderá permitir o ajuste dos sistemas de coleta de dados do MEC, assim como os sistemas de gestão acadêmica das diversas instituições de ensino. Na Figura 14 está a representação gráfica da ontologia, contendo apenas as principais classes que se relacionam com a classe curso.

Figura 14 - Classes da ONTO-RA que se relacionam com a classe curso



Já na Figura 15, está a representação gráfica das classes que se relacionam com a classe Aluno. As Classes Aluno e Curso são as principais classes da ONTO-RA, que está representada de forma completa na linguagem OWL em <http://xbrlframework.wordpress.com/>.

Figura 15 - Classes da ONTO-RA que se relacionam com a classe aluno



5.9 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os processos, atividades e tarefas que foram executados para a geração da ontologia para o domínio dos registros acadêmicos do modelo brasileiro de educação ONTO-RA.

A construção da ONTO-RA foi feita utilizando a metodologia NEON, que sofreu pequenas alterações para promover uma melhor adequação aos recursos não ontológicos disponíveis. A primeira delas ocorreu durante a avaliação dos candidatos a NORs, quando não foi utilizado o critério precisão. A segunda adaptação foi a fusão dos candidatos a NORs em um único documento estruturado, garantindo assim 100% de cobertura dos requisitos e a aplicabilidade direta da

metodologia pela compatibilidade do tipo de recurso e modelo de dados com padrões de projeto de reengenharia de NORs do repositório da metodologia NEON.

A ONTO-RA foi desenvolvida utilizando a ferramenta Protege e a linguagem de construção de ontologias utilizada foi a OWL. A versão final da ontologia apresenta as diversas classes e relacionamentos do tipo Subclasse extraídos do esquema de classificação gerado na Atividade de Seleção de Recurso não Ontológico mais Adequado e os relacionamentos *Ad-hoc* entre as classes.

A ONTO-RA cobre apenas os requisitos responsáveis por geração de indicadores de desempenho acadêmico e de gestão, portanto, características básicas de algumas classes não foram representadas neste trabalho. Conseguiu-se, assim, através da ontologia e do glossário de termos gerado durante a fase de especificação de requisitos, o mapeamento semântico dos principais termos responsáveis por gerar os indicadores do TAMC, assim como possibilitar a uniformização dos sistemas de coleta de dados responsáveis desde a geração de estatísticas oficiais do MEC, até a definição de matriz orçamentária dos Institutos Federais.

6 AVALIAÇÃO DA ADESÃO DOS SISTEMAS EM RELAÇÃO À ONTO-RA

Neste capítulo será apresentada uma avaliação dos sistemas de coleta de dados do MEC e dos sistemas de gestão acadêmica utilizados no IFBA com o objetivo de evidenciar o nível de inadequação destes sistemas em relação ao modelo ontológico construído neste trabalho e conseqüentemente avaliar a efetividade da proposta, uma vez que a ONTO-RA se propõe a servir de base para ajustes nos sistemas existentes e para a construção de novos sistemas.

O estudo foi realizado tomando como base a ontologia desenvolvida, incluindo os artefatos gerados durante o processo de desenvolvimento, como o glossário de termos que compõe o OSRD, o próprio OSRD, o modelo conceitual gerado no processo de Reengenharia de Recursos não Ontológicos, os principais sistemas de coletas de dados utilizados pelo MEC, que são o EDUCACENSO, o CENSO DO ENSINO SUPERIOR e o SISTEC e os sistemas de Gestão Acadêmica utilizados pelo IFBA que são o SICAD e o SIGA-EPCT, além dos documentos como o TAMC e o Diagnóstico de Retenção e Evasão do MEC.

O objetivo da avaliação é, através da ontologia gerada, identificar todas as possíveis necessidades de alteração nos sistemas de coleta de dados do MEC e dos sistemas de gestão acadêmica utilizados no IFBA. Além disso, para cada inadequação identificada, deverão ser identificados os indicadores presentes no TAMC ou no Diagnóstico de Retenção e Evasão do MEC impactados. Então, as devidas alterações nos modelos de dados dos sistemas de gestão acadêmica ou nos layouts dos arquivos dos sistemas de coleta de dados.

Inicialmente, será feita uma breve apresentação dos sistemas de coletas de dados do MEC, assim como dos sistemas de gestão acadêmica do IFBA. Em seguida a execução da aplicação da ONTO-RA será apresentada. Por fim, será apresentado um quadro que apresenta o nível de inadequação de cada sistema avaliado.

6.1 SISTEMAS DE COLETA DO MEC

Conforme mencionado na Seção 1.1 a autonomia administrativa das autarquias federais de ensino, que propicia a possibilidade de definição de normas acadêmicas distintas entre as diversas instituições federais de ensino, tem como

conseqüência a estruturação dos registros acadêmicos de forma heterogênea. Tal característica, fez com que houvesse dificuldade por parte do MEC em desenvolver e disponibilizar para as instituições federais de ensino um sistema informatizado de gestão acadêmica centralizado, o que fez com que cada instituição buscasse soluções próprias de sistema.

Este contexto, fez com que o MEC desenvolvesse os chamados sistemas de coleta de dados para que fosse possível obter informações acerca das ações desenvolvidas nas instituições federais de ensino e a partir daí gerar indicadores gerenciais, indicadores de desempenho acadêmico, acompanhar metas estabelecidas, mensurar o alcance das políticas públicas implementadas, gerar matriz orçamentária das instituições vinculadas ao MEC, entre outras possibilidades. Dentre os sistemas desenvolvidos com esta finalidade, atualmente existem em funcionamento o EDUCACENSO, o CENSO DA EDUCAÇÃO SUPERIOR e o SISTEC.

O EDUCACENSO (<http://portal.inep.gov.br/web/educacenso/educacenso>) é um sistema de coleta de dados da educação básica das instituições das diversas esferas de governo (Municipais, Estaduais e Federais) e também as instituições da rede particular de ensino. Os dados coletados pelo EDUCACENSO abrangem desde a caracterização das instituições, através de características físicas, administrativas e de disponibilidade de recursos diversos, passando pela matrícula dos alunos em cursos, atuação dos docentes chegando até a movimentação e rendimento escolar.

As informações levantadas através do EDUCACENSO são utilizadas para traçar um panorama nacional da educação básica e servem de referência para a formulação de políticas públicas e execução de programas na área da educação, incluindo os de transferência de recursos públicos como merenda e transporte escolar, distribuição de livros e uniformes, implantação de bibliotecas, instalação de energia elétrica, Dinheiro Direto na Escola e Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb). Além disso, os resultados obtidos no EDUCACENSO sobre o rendimento (aprovação e reprovação) e movimento (abandono) escolar dos alunos do ensino Fundamental e Médio, juntamente com outras avaliações do INEP (Saeb e Prova Brasil), são utilizados para o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB), indicador que serve de referência para as metas do Plano de Desenvolvimento da Educação (PDE), do MEC.

O CENSO DA EDUCAÇÃO SUPERIOR (<http://portal.inep.gov.br/web/censo-da-educacao-superior/censo-da-educacao-superior>) é um sistema de coleta de dados destinado às instituições que possuem atuação no ensino superior nas esferas municipal, estadual e federal e na rede pública e particular. Assim como o EDUCACENSO, o CENSO DA EDUCAÇÃO SUPERIOR parte da caracterização da instituição, embora que num nível de detalhamento inferior ao EDUCACENSO, passando pelo cadastro dos cursos de graduação presenciais e a distância, oferta de vagas em vestibulares e outras formas de oferta, matrícula, passando também por informações sobre os docentes que atuam no curso.

O Censo da Educação Superior, é realizado anualmente pelo INEP, é o instrumento de pesquisa mais completo sobre as instituições de educação superior (IES) que ofertam cursos de graduação e sequenciais de formação específica, e seus alunos e docentes. As estatísticas educacionais produzidas oferecem aos usuários informações sobre ingressos, matrículas, concluintes, vagas, dados de financiamento estudantil, recursos de tecnologia assistiva disponíveis às pessoas com deficiência, entre outros. O Censo ainda subsidia o planejamento e a avaliação de políticas públicas, além de contribuir no cálculo de indicadores de qualidade como o Cálculo Preliminar de Curso (CPC) e Índice Geral de Cursos (IGC).

6.2 SISTEMAS DE GESTÃO ACADÊMICA DO IFBA

O Sistema de Informações da Educação Profissional e Tecnológica – SISTEC é destinado à educação básica, educação superior e à formação inicial e continuada. O objetivo do SISTEC é promover mecanismos de registro e controle dos dados da educação profissional e tecnológica do país. Através do SISTEC é possível realizar o cadastro dos cursos vinculados às instituições, inclusão de alunos agrupados por ciclos de matrícula, que servem para caracterizar o momento da entrada do aluno na instituição e a sua previsão de saída, bem como o registro da situação final do aluno. O SISTEC monitora apenas o momento do início do vínculo do aluno e o momento em que cada aluno possui alguma movimentação que represente uma situação final do aluno no curso.

O Sistema de Controle Acadêmico – SICAD é o sistema de gestão acadêmica utilizado pelo IFBA desde 1999 e é responsável pela gestão acadêmica dos campi com maior número de alunos do IFBA, como Salvador, Simões Filho, Vitória da

Conquista, Santo Amaro e Eunápolis. O SICAD realiza o controle acadêmico desde o cadastro dos cursos com suas respectivas matrizes curriculares até a gestão dos rendimentos escolares dos alunos em todos os níveis de ensino e formas de oferta. Apesar de ser desenvolvido em Delphi 6, o SICAD conta com módulos disponibilizados na WEB que são o portal do aluno, desenvolvido em ASP, que permite a realização de renovação de matrícula a cada novo período letivo e emissão de alguns documentos acadêmicos, como boletim e histórico escolar.

O Sistema Integrado de Gestão Acadêmica da Educação Profissional, Científica e Tecnológica – SIGA-EPCT, é fruto de uma iniciativa dos Institutos Federais de Educação Tecnológica que desenvolveu e mantém, através de um projeto de desenvolvimento colaborativo e distribuído, um sistema que visa atender às demandas de gestão acadêmica para instituições de educação profissional e tecnológica. Este sistema, desenvolvido em Java, está em fase de implantação no IFBA.

6.3 ANÁLISE POR QUESTÃO DE COMPETÊNCIA

Nesta seção será verificada cada questão de competência, doravante tratada como QC, identificada na ORSD e atendida pela ONTO-RA, em relação ao EDUCACENSO, SISTEC e CENSO DO ENSINO SUPERIOR. A avaliação começa com a análise do atendimento das questões de competência pelos sistemas de coleta de dados do MEC, EDUCACENSO, SISTEC e CENSO DO ENSINO SUPERIOR. Para cada questão de competência identificada na geração do ORSD e listadas no Quadro 6, será verificado o atendimento pelos sistemas e em caso de não atendimento, serão identificados os possíveis indicadores impactados e as ações necessárias para adaptação à ONTO-RA. Em seguida, será apresentado um quadro comparativo entre os diversos sistemas com relação ao nível de atendimento às questões de competência

Para cada questão de competência, será apresentado o trecho da ontologia que atende à QC e será utilizada uma tabela que indicará o atendimento à QC pelos sistemas analisados. Em caso de não atendimento, as ações necessárias em cada sistema será indicada. Em seguida serão identificados os indicadores do TAMC e do Diagnóstico de Evasão e Retenção relacionados à questão de competência e a justificativa para a utilização da QC, em caso de não existência de indicadores

relacionados identificados. A seguir, estão listadas as questões de competência e suas respectivas avaliações

- QC-01: Quais são os tipos de atividades extracurriculares para alunos ?

As atividades extracurriculares foram representadas na ONTO-RA, como uma classe chamada “AtividadesComplementares” que se relaciona com a classe “Aluno” através do relacionamento “TemAtividadesComplementares”.

Quadro 10 - Análise da QC-01

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Inclusão de atividades extracurriculares no layout de migração de dados
SISTEC	NÃO	Como o SISTEC tem o objetivo de identificar a situação inicial e final dos alunos, não há necessidade de atendimento deste requisito.
CENSUP	SIM	-
SICAD	NÃO	Ajuste no Modelo de Dados e na Aplicação
SIGA-EPCT	NÃO	Ajuste no Modelo de Dados e na Aplicação

Indicadores Impactados: Uma das metas pactuadas entre IFs e MEC, indica que deverá haver um número mínimo de projetos de pesquisa e estabelece o nível de crescimento desta atividade. Entretanto, não há impacto direto no levantamento para o Diagnóstico de Evasão e retenção do MEC. Entretanto, o EDUCACENSO e o CENSO DO ENSINO SUPERIOR requisitam tal informação e como estes sistemas de coleta geram estatísticas oficiais da educação brasileira, a inexistência do tratamento deste requisito deve ocasionar falha na geração deste indicador estatístico específico.

- QC-02: Quais são as possíveis disciplinas?

Em ONTO-RA, Disciplina é uma classe que possui como indivíduos ou instâncias o elenco de possíveis disciplinas listadas como respostas à QC-02. Em ONTO-RA a classe Disciplina se relaciona com as classes Docente e Curso através dos relacionamentos “TemDocente” e “TemCurso” respectivamente. O Quadro 11 apresenta o resultado da análise do atendimento ao requisito QC-02.

Quadro 11 - Análise da QC-02

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	SIM	-
SISTEC	NÃO	Como o SISTEC tem o objetivo de identificar a situação inicial e final dos alunos, não há necessidade de atendimento deste requisito.
CENSUP	NÃO	A lista de disciplinas elencadas na resposta da QC-02 são disciplinas voltadas ao atendimento do ensino do nível básico, portanto, não há necessidade de ajuste em relação ao CENSUP.
SICAD	NÃO	Ajuste no Modelo de Dados
SIGA-EPCT	NÃO	Ajuste no Modelo de Dados

Indicadores Impactados: Não há impacto direto em indicadores do TAMC, tão pouco em indicadores do Diagnóstico de Evasão e retenção do MEC. Entretanto, o EDUCACENSO e o CENSO DO ENSINO SUPERIOR requisitam tal informação e como estes sistemas de coleta geram estatísticas oficiais da educação brasileira, a inexistência do tratamento deste requisito deve ocasionar falha na geração deste indicador estatístico específico. Embora não haja indicadores baseados em tipo de disciplina, este requisito é importante para a identificação de evasão relacionada a dificuldade específica em determinado conteúdo.

- QC-03: Quais os possíveis turnos de estudo do aluno?

Em ONTO-RA, Turno é uma classe que possui como indivíduos ou instâncias matutino, vespertino, noturno e integral. A classe Turno se relaciona com a classe Curso através da Propriedade “TemTurno”.

Quadro 12 - Análise da QC-03

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	SIM	-
SISTEC	SIM	-
CENSUP	SIM	-
SICAD	SIM	-
SIGA-EPCT	SIM	-

Conforme pode ser verificado no Quadro 12, todos os sistemas analisados atendem à QC-03.

- QC-04: Quais são as modalidades de oferta de cursos para alunos?

Em ONTO-RA, ModalidadeOferta é uma classe que possui como indivíduos ou instâncias o “Presencial” e “A Distancia - EAD”. Em ONTO-RA ModalidadeOferta se relaciona com a classes Curso através do relacionamento “TemModalidadeOferta”. O Quadro 13 apresenta o resultado da análise do atendimento ao requisito QC-04.

Quadro 13 - Análise da QC 04

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Inclusão no leiaute do arquivo de migração e inclusão no sistema de cadastro.
SISTEC	NÃO	Modificação no sistema para permitir a indicação desta característica para o curso, adequando assim o sistema à LDB.
CENSUP	NÃO	Não há necessidade de atualização, uma vez que tal informação consta no E-MEC, que é um sistema de coleta de informações do MEC que é responsável por validar os cursos superiores.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	SIM	Ajuste no Modelo de Dados

Indicadores Impactados: Uma das metas do TAMC indica que deverão ser implementados pelos IFs cursos cuja Modalidade de Oferta seja “A distância”. Além disso, nos indicadores de eficácia do mesmo documento também não deverão ser considerados os cursos cuja forma de oferta seja “A distância”. A mesma recomendação vale para o cálculo da relação aluno/professor estabelecida pelo TAMC.

O Diagnóstico de Evasão e retenção do MEC necessita desta informação, que permite comparar os índices de evasão e retenção entre cursos presenciais e a distância.

- QC-05: Quais são os possíveis níveis de ensino de cursos?

Em ONTO-RA, NívelEnsino é uma classe que se relaciona com a classe Curso através da propriedade “TemNívelEnsino” e possui como subclasses “EducacaoBasica”, “EducacaoInicialeContinuadaFic” e “EducacaoSuperior”. O Quadro 15 apresenta o resultado da análise do atendimento ao requisito QC-05.

Quadro 14 - Análise da QC-05

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Não são necessários ajustes, uma vez que o EDUCACENSO destina-se a coletar dados de cursos apenas da educação básica.
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB.
CENSUP	NÃO	Não são necessários ajustes uma vez que todos os cursos que são objetos das coletas realizadas pelo CENSUP são do nível da educação superior.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	SIM	-

Indicadores Impactados: Não há indicadores específicos do TAMC e do Diagnóstico para Retenção e Evasão do MEC, mas qualquer indicador estatístico que utilize esta classificação como agrupador, terá dificuldade de ser gerado em função das inadequações encontradas no SICAD e SISTEC.

- QC-06: Quais são as possíveis etapas de ensino da educação básica?

Em ONTO-RA, EtapaEnsino é uma Subclasse da Classe EducacaoBasica. EtapaEnsino tem como subclasse EnsinoMedioEPT e possíveis instâncias Educação Infantil–CRECHE, Educação Infantil – PRÉ-ESCOLA, Ensino Fundamental, Ensino Médio e Ensino Médio–Normal Magistério.

Quadro 15 - Análise da QC-06

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Embora o EDUCACENSO mapeie todas as possibilidades de etapas de ensino, há uma inadequação semântica neste sistema. O EDUCACENSO mistura as classificações de etapas de ensino, tratado na QC-06, com formas de oferta do ensino médio – EPT, tratados pela QC-07.
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB.
CENSUP	NÃO	Não são necessários ajustes uma vez que todos os cursos que são objetos das coletas realizadas pelo CENSUP são do nível da educação superior.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	SIM	-

Indicadores Impactados: Não há indicadores específicos do TAMC e do Diagnóstico para Retenção e Evasão do MEC, mas qualquer indicador estatístico que utilize esta classificação como agrupador, terá dificuldade de ser gerado em função das inadequações encontradas no SICAD e SISTEC.

- QC-07: Quais as formas de oferta do Ensino Médio - EPT?

Em ONTO-RA, EnsinoMedioEPT é uma Subclasse de EtapaEnsino. EnsinoMedioEPT tem como subclasse FormaOferta, que possui como instâncias Articulada Integrada, Articulada Concomitante e Subsequente.

Quadro 16 - Análise da QC-07

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Embora o EDUCACENSO mapeie todas as possibilidades de etapas de ensino, há uma inadequação semântica neste sistema. O EDUCACENSO mistura as classificações de etapas de ensino, tratado na QC-06, com formas de oferta do ensino médio – EPT, tratados pela QC-07.
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB
CENSUP	NÃO	Não são necessários ajustes uma vez que todos os cursos que são objetos das coletas realizadas pelo CENSUP são do nível da educação superior.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	SIM	-

Indicadores Impactados: Não há indicadores específicos do TAMC e do Diagnóstico para Retenção e Evasão do MEC, mas qualquer indicador estatístico

que utilize esta classificação como agrupador, terá dificuldade de ser gerado em função das inadequações encontradas no SICAD e SISTEC. A inadequação semântica identificada no EDUCACENSO pode gerar a falha na interpretação de resultados de dados estatísticos gerados.

- QC-08: Quais os níveis acadêmicos dos cursos cujo nível é Educação Superior?

Em ONTO-RA, NivelAcademico é uma Subclasse de EducacaoSuperior. NivelAcademico tem como subclasses Extensao, Graduacao, PosGraduacao e PosGraduacaoTecnologica.

Quadro 17 - Análise da QC-08

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Não são necessários ajustes uma vez que todos os cursos que são objetos das coletas realizadas pelo CENSUP são do nível da educação básica.
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB.
CENSUP	NÃO	Não há necessidade de ajustes, uma vez que o cadastro de curso é realizado no sistema E-MEC.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	NÃO	O SIGA-EPCT mapeia todas as possibilidades de níveis acadêmicos dos cursos da Educação Superior, entretanto, sem utilizar a mesma semântica utilizada na LDB.

Indicadores Impactados: Embora não existam indicadores específicos que utilizem as características dos Níveis Acadêmicos tratados na QC-08, os índices de

eficácia existentes no TAMC são gerados segmentados pela classificação hierárquica dos cursos. A indefinição semântica detectada neste requisito pode gerar informações equivocadas assim como interpretações indevidas dos resultados de geração de indicadores.

- QC-09: Quais os graus acadêmicos dos cursos da Educação Superior no nível acadêmico graduação?

Em ONTO-RA, GrauAcademico é uma Subclasse de Graduação e possui como instâncias Bacharelado, Licenciatura e Tecnológico.

Quadro 18 - Análise da QC-09

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Não são necessários ajustes uma vez que todos os cursos que são objetos das coletas realizadas pelo CENSUP são do nível da educação básica.
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB.
CENSUP	NÃO	Não há necessidade de ajustes, uma vez que o cadastro de curso é realizado no sistema E-MEC.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	NÃO	O SIGA-EPCT mapeia todas as possibilidades de níveis acadêmicos dos cursos da Educação Superior, entretanto, sem utilizar a mesma semântica utilizada na LDB.

Indicadores Impactados: Embora não existam indicadores específicos que utilizem as características dos Graus Acadêmicos tratados na QC-09, os índices de eficácia existentes no TAMC são gerados segmentados pela classificação

hierárquica dos cursos. A indefinição semântica detectada neste requisito pode gerar informações equivocadas assim como interpretações indevidas dos resultados de geração de indicadores.

- QC-10: Quais os Programas de curso dos níveis acadêmicos de Pós-graduação e Pós-graduação Tecnológica?

Em ONTO-RA, PosGraduacao e PosGraduacaoTecnologica são subclasses de NivelAcademico, e cada uma delas possui como suas subclasses Programas e ProgramasPosGraduacaoTecnologica. Como instâncias têm-se Mestrado, Mestrado Profissional, Doutorado, Especialização e Aperfeiçoamento.

Quadro 19 - Análise da QC - 10

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Não são necessários ajustes uma vez que todos os cursos que são objetos das coletas realizadas pelo CENSUP são do nível da educação básica.
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB.
CENSUP	NÃO	Não há necessidade de ajustes, uma vez que o cadastro de curso é realizado no sistema E-MEC.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	NÃO	O SIGA-EPCT mapeia todas as possibilidades de níveis acadêmicos dos cursos da Educação Superior, entretanto, sem utilizar a mesma semântica utilizada na LDB.

Indicadores Impactados: Embora não existam indicadores específicos que utilizem as características dos Níveis Acadêmicos tratados na QC-10, os índices de

eficácia existentes no TAMC são gerados segmentados pela classificação hierárquica dos cursos. A indefinição semântica detectada neste requisito pode gerar informações equivocadas assim como interpretações indevidas dos resultados de geração de indicadores.

- QC-11: Quais as possíveis modalidades de ensino?

Em ONTO-RA, ModalidadeEnsino é uma classe que se relaciona com a classe Curso através da propriedade TemModalidadeEnsino. A classe ModalidadeEnsino possui como indivíduos ou instâncias Educação Regular, Educação Especial e de Jovens e Adultos (EJA), esta última que pode se relacionar apenas com cursos da educação básica.

Quadro 20 - Análise da QC-11

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	SIM	-
SISTEC	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SISTEC, que não segue o modelo definido na LDB.
CENSUP	NÃO	Não há necessidade de ajustes, uma vez que o CENSUP não coleta dados de cursos da educação básica.
SICAD	NÃO	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	NÃO	O SIGA-EPCT mapeia todas as possibilidades de níveis acadêmicos dos cursos da Educação Superior, entretanto, sem utilizar a mesma semântica utilizada na LDB.

Indicadores Impactados: O Termo de Acordo de Metas e Compromissos – TAMC, tem entre os compromissos dos IFs a ofertar cursos da Modalidade de Ensino EJA, no nível de ensino de Formação Inicial e Continuada e Educação Básica, forma de oferta Ensino Médio – EPT.

- QC-12: Quais as possíveis situações de curso?

Este requisito está representado em ONTO-RA, pela classe SituacaoCurso que se relaciona com a classe Curso através da propriedade TemSituacaoCurso. A classe SituacaoCurso possui as instâncias Em Atividade e Extinto.

Quadro 21 - Análise da QC-12

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Adequação do leiaute do arquivo de migração e do sistema de carga de dados.
SISTEC	NÃO	Adequação do sistema para registrar a situação do curso.
CENSUP	NÃO	Não há necessidade de ajustes, uma vez que o CENSUP não coleta dados de cursos da educação básica.
SICAD	SIM	A estruturação hierárquica das características dos cursos deverá ser completamente refeita no SICAD, que não segue o modelo definido na LDB.
SIGA-EPCT	SIM	O SIGA-EPCT mapeia todas as possibilidades de níveis acadêmicos dos cursos da Educação Superior, entretanto, sem utilizar a mesma semântica utilizada na LDB.

Indicadores Impactados: Todos os indicadores do TAMC, do diagnóstico de Retenção e Evasão e qualquer indicador estatístico gerado levando em conta informações de ingressantes em cursos sem verificar se o mesmo está extinto ou em atividade, poderá gerar indicadores incorretos.

- QC-13: O que se entende por inscrições de candidatos em cursos?
- QC-14: O que se entende por vagas ofertadas para cursos ?
- QC-15: O que se entende por alunos ingressantes em cursos ?

A QC-13 está representado em ONTO-RA, pela propriedade TemInscricoesEmProcessoSeletivo, que determina o número de inscritos em instâncias da classe ProcessoSeletivo. Já a QC-14 é representada pela propriedade TemVagas e QC-15 define Ingressante como uma subclasse de Aluno.

Quadro 22 - Análise da QC-13, QC-14 e QC-15

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	SIM	-
SISTEC	SIM	-
CENSUP	SIM	-
SICAD	SIM	-
SIGA-EPCT	SIM	-

Na análise dos sistemas, conforme pode ser visto no Quadro 22, não foram identificadas inadequações. Estes requisitos são utilizados no cálculo de eficiência e eficácia definidos no TAMC. Além disso, estes requisitos são importantes para a definição dos termos ciclo de matrícula e turma de ingresso.

- Requisitos do Grupo Organização Acadêmica

As questões de competência agrupadas como organização acadêmica são:

- QC-16: O que se entende por turma?
- QC-17: O que se entende por ciclo de matrícula de aluno em curso?
- QC-18: O que se entende por turma de ingresso?

A presença da QC-16 entre as questões de competência tem o propósito de diferenciar os conceitos de turma e turma de ingresso. O primeiro representa um agrupamento de alunos que estão vinculados a uma mesma disciplina ou grupo de disciplinas no mesmo horário e ministrado pelos mesmos docentes. Já a turma de ingresso, definida QC-18, representa o conjunto de alunos que ingressaram no mesmo curso no mesmo período letivo. Já o ciclo de matrícula, definido na QC-17, define que é um conjunto de alunos que ingressam num curso num mesmo período letivo e, portanto, possuem a mesma data de início e data de fim prevista.

Quadro 23 - Análise da QC-16, QC-17 e QC-18

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Não há necessidade de ajustes, uma vez que os conceitos de turma de ingresso e ciclo de matrícula são utilizados pelo SISTEC.
SISTEC	SIM	-
CENSUP	SIM	Não há necessidade de ajustes, uma vez que os conceitos de turma de ingresso e ciclo de matrícula são utilizados pelo SISTEC.
SICAD	NÃO	Ajuste no modelo de dados e interface de cadastro para permitir a definição de turma de ingresso e ciclo de matrícula.
SIGA-EPCT	NÃO	Ajuste no modelo de dados e interface de cadastro para permitir a definição de turma de ingresso e ciclo de matrícula.

Estes requisitos precisam ser implementados apenas nos sistemas de gestão acadêmica, uma vez que há necessidade apenas de atender a um requisito do sistema de coleta SISTEC.

- Requisitos do grupo Pessoa

Os requisitos agrupados com o tipo pessoa são:

- QC-19: Quais as possíveis cor/raça de alunos e docentes?
- QC-20: Quais as possíveis nacionalidades de alunos e docentes?
- QC-21: Quais os possíveis tipos de deficiência físicas, intelectuais ou sensoriais de alunos e docentes?
- QC-22: Quais os possíveis tipos de origem escolar do ensino médio para alunos?
- QC-23: Quais as possíveis escolaridades para docentes?
- QC-24: Quais as possíveis pós-graduações para os docentes?
- QC-25: Quais os tipos de zona de residência de alunos e docentes?

Quadro 24 - Análise das QCs do grupo pessoa

Sistema Analisado	Atendimento	QC não atendidas	Ações Necessárias
EDUCACENSO	SIM	-	-
SISTEC	NÃO	19, 20, 21, 22, 23, 24 e 25	Não há necessidade de ajustes, uma vez que o foco do SISTEC é na situação inicial e final dos alunos e não leva em consideração dados pessoais de alunos e docentes.
CENSUP	SIM	-	-
SICAD	NÃO	21	Ajuste no modelo de dados e interface de cadastro de docentes.
SIGA-EPCT	SIM	-	-

Este grupo de requisitos é utilizado principalmente para geração de estatísticas sobre o nível de qualificação dos docentes e alunos, bem como, pode ser possível correlacionar tais informações com situações como evasão e retenção, visando identificar as possíveis causas destas ocorrências.

- QC-26: O que se entende por tempo mínimo de integralização?

ONTO-RA trata o requisito QC-26 como um relacionamento chamado TemTempoIntegralizacao que define o tempo típico necessário para a integralização do aluno no curso.

Quadro 25 - Análise da QC-26

Sistema Analisado	Atendimento	Ações Necessárias
EDUCACENSO	NÃO	Ajuste no leiaute e no sistema para permitir o cadastro de curso.
SISTEC	NÃO	O SISTEC utiliza o critério de ciclo de matrícula para identificar a previsão de término dos alunos, portanto não há necessidade de ajustes.
CENSUP	SIM	Ajuste no leiaute e no sistema para permitir o cadastro de curso.
SICAD	SIM	-
SIGA-EPCT	SIM	-

Este requisito é importante para a definição do cálculo de retenção, portanto, tem um forte impacto no levantamento para o diagnóstico de evasão e retenção de alunos.

- Requisitos do Grupo Situação Acadêmica

Os requisitos agrupados com o tipo Situação Acadêmica são:

- QC-27: Quais são os possíveis tipo de rendimento de alunos?
- QC-28: O que se entende por aluno concluinte em curso?
- QC-29: O que se entende por aluno integralizado em curso?
- QC-30: O que se entende por retenção de aluno em curso?
- QC-31: O que se entende por evasão de aluno em curso?

Quadro 26 - Análise das QCs do grupo situação acadêmica

Sistema Analisado	Atendimento	QC não atendidas	Ações Necessárias
EDUCACENSO	SIM	-	-
SISTEC	SIM	-	-
CENSUP	SIM	-	-
SICAD	SIM	-	-
SIGA-EPCT	SIM	-	-

Todos os sistemas de coleta e de gestão acadêmica analisados possuem em seus modelos de dados informações que permitem auferir as informações relatadas

nas QCs agrupadas como Situação Acadêmica. A grande contribuição que ONTORA apresenta ao atender este grupo de requisitos é garantir a definição semântica para termos como retenção e evasão que são informações obtidas a partir da análise de dados de rendimento dos alunos. Além disso, os índices de eficiência e eficácia definidos no TAMC utilizam o número de alunos concluintes, o que evidencia a importância de tais definições.

- Requisitos do Grupo Vínculo Institucional para alunos

Um dos grupos utilizados nas questões de competência é Vínculo Institucional, entretanto, para efeito deste estudo de caso, tais requisitos serão separados em Vínculo Institucional para Alunos e Vínculo institucional para Docentes. Nesta seção serão analisados o atendimento aos requisitos de Vínculo Institucional para Alunos. Tais requisitos são representados pelas QCs listadas a seguir:

- QC-32: Quais são as possíveis situações de vínculo dos alunos?
- QC-33: Quais os tipos de mobilidade acadêmica de aluno?
- QC-34: Quais os tipos de mobilidade acadêmica internacional de aluno?
- QC-35: O que significa aluno PARFOR?
- QC-36: Quais são as possíveis formas de ingresso de alunos?
- QC-37: Quais os tipos de programas de reserva de vagas de alunos?
- QC-38: Quais os tipos de apoio social a alunos?

Quadro 27 – Análise das QCs do grupo vínculo institucional

Sistema Analisado	Atendimento	QC não atendidas	Ações Necessárias
EDUCACENSO	SIM	-	-
SISTEC	NÃO	32,33,34,35,37 e 38	Ajuste nas interfaces do sistema que atualizam a situação de vínculo dos alunos.
CENSUP	SIM	-	-
SICAD	NÃO	32 e 38	Ajuste no modelo de dados e cadastros do sistema.
SIGA-EPCT	NÃO	32 e 38	Ajuste no modelo de dados e cadastros do sistema.

O cálculo de evasão e retenção, assim como a medida da eficiência e eficácia definidas no TAMC, são bastante dependentes dos requisitos de vínculo institucional, que precisam, inclusive, estar unificados para que os parâmetros de comparação possam ser estabelecidos sem possibilidade de comparações inadequadas. A própria definição de alunos efetivamente matriculados depende destes conceitos estarem unificados e bem definidos.

6.3.1 Requisitos do Grupo Vínculo Institucional para docentes

Nesta seção serão analisados o atendimento aos requisitos de Vínculo Institucional para Docentes. Tais requisitos, são representados pelas QCs listadas a seguir:

- QC-39: Quais são as possíveis situações de vínculo de docentes?
- QC-40: Quais são os possíveis regimes de vínculo para docentes?
- QC-41: Quais são os possíveis tipos de vínculo para docentes?
- QC-42: Quais são os possíveis tipos de atuação para docentes?

Quadro 28 – Análise das QCs do grupo vínculo institucional para docentes

Sistema Analisado	Atendimento	QC não atendidas	Ações Necessárias
EDUCACENSO	SIM	-	-
SISTEC	NÃO	39, 40, 41 e 42	Não são necessários ajustes, visto que o objetivo do SISTEC é apenas voltado para situações inicial e final de alunos
CENSUP	SIM	-	-
SICAD	NÃO	39 e 42	Ajuste no modelo de dados e telas de cadastro.
SIGA-EPCT	SIM	-	-

Os requisitos relacionados ao Vínculo Institucional de docentes são utilizados no TAMC para definir a relação entre alunos matriculados e a força de trabalho. Além disso, servem para a definição de estatísticas a cerca dos docentes presentes na instituição e correlacionar com outros indicadores.

6.4 CONSIDERAÇÕES FINAIS

Neste capítulo foi realizado uma avaliação que visou identificar os possíveis ajustes necessários nos sistemas analisados a partir de suas respectivas inadequações em relação ao modelo ontológico proposto pela ONTO-RA. Deste estudo, pôde-se inferir que as inadequações em relação ao modelo ontológico dos sistemas avaliados, geram consequências importantes na geração de indicadores de desempenho acadêmico e gerenciais do MEC e das Instituições de Ensino.

Ao evidenciar tais inadequações e propor um modelo de organização hierárquica e semântica dos conceitos, a Ontologia gerada neste trabalho, permite que os ajustes aos citados sistemas sejam realizados pelas equipes de desenvolvimento próprias de cada instituição assim como pela equipe de Tecnologia da Informação do MEC.

O fato do modelo proposto não ser apenas um modelo de dados ou outra forma de representação, garante que os desenvolvedores que irão trabalhar com desenvolvimento no domínio dos registros acadêmicos brasileiro poderão utilizar as

definições dos termos modelados na ontologia para uniformizar o tratamento dos diversos conceitos do domínio nos sistemas desenvolvidos e assim facilitar a interoperabilidade.

A partir da análise realizada foi possível inferir o percentual de inadequações identificadas a partir da quantidade de questões de competência não atendidas por cada sistema analisado. O Quadro 29 apresenta o percentual de QCs não atendidas por sistema analisado. Sendo assim a aplicação da ONTO-RA permite inferir que os sistemas de gestão acadêmica utilizados no IFBA apresentam distorções importantes no que diz respeito ao padrão do domínio dos registros acadêmicos estabelecido por lei. Esta situação ainda permite a geração de indicadores pouco confiáveis, uma vez que nem todos os itens necessários para a geração destes indicadores, estão representados nos sistemas de forma fidedigna. Pode-se ainda perceber, que os próprios sistemas de coleta de dados do MEC apresentam um grau alto de distorção. No próximo capítulo serão apresentadas as considerações finais desta Dissertação.

Quadro 29 - Percentual das QCs não atendidas por sistema analisado

SISTEMA	% DE QCs NÃO ATENDIDAS
EDUCACENSO	13 %
SISTEC	30 %
CENSUP	2 %
SICAD	37 %
SIGA	23 %

7 CONCLUSÃO

Neste capítulo serão apresentadas as considerações finais do trabalho de desenvolvimento da proposta de ontologia do domínio dos registros acadêmicos para o modelo brasileiro da educação. A Seção 7.1 apresenta as considerações finais sobre o trabalho, ressaltando os principais aspectos do processo de desenvolvimento, a Seção 7.2 apresenta as principais contribuições advindas do trabalho e, por fim, a Seção 7.3 elenca as possibilidades de trabalhos futuros. E a Seção 7.4 mostra os artigos produzidos a partir do trabalho desenvolvido nessa publicação.

7.1 CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma ontologia para o domínio dos registros acadêmicos para o modelo brasileiro da educação. Este modelo ontológico visa superar uma lacuna semântica existente neste domínio, e assim promover a possibilidade de uma melhor integração entre os sistemas acadêmicos das instituições de ensino e o MEC, bem como uma maior qualidade na geração de indicadores de gestão e de desempenho acadêmico.

Um dos objetivos específicos deste trabalho era realizar uma pesquisa para identificar a possibilidade de existência de outros trabalhos que realizassem o mapeamento semântico do domínio dos registros acadêmicos. Nesta busca não foram encontradas abordagens que fossem específicas do objeto da busca, entretanto, verificou-se que o uso de ontologias vem sendo trabalhado na comunidade científica como base para integração de sistemas e padronização semântica entre os mesmos.

Outro objetivo alcançado foi a busca e conseqüente escolha de uma metodologia que fosse mais adequada ao cenário para o qual a ontologia foi desenvolvida, para tanto, foi realizado um processo de levantamento de informações, avaliação e escolha da metodologia de construção de ontologias, que resultou na adoção da metodologia NEON, mais especificamente no seu cenário 2, que preconiza a reutilização de recursos não ontológicos para o desenvolvimento da ontologia. Nesta busca, foi possível estabelecer, através da utilização de critérios de avaliação já utilizados na literatura, mas que se adequassem ao cenário específico

deste trabalho, um painel comparativo entre as diversas metodologias de desenvolvimento de ontologias.

Assim, atendendo ao principal objetivo deste trabalho, a metodologia NEON foi aplicada e a Ontologia de Registros Acadêmicos do Modelo Educacional Brasileiro – ONTO-RA foi construída, gerando assim uma definição de todos os termos importantes para a geração dos principais indicadores de desempenho acadêmico, além de estabelecer os principais relacionamentos entre eles.

Por fim, foi executado uma avaliação dos sistemas de coleta de dados do MEC e dois sistemas acadêmicos utilizados no âmbito do IFBA a partir da ONTO-RA, que demonstrou a existência de uma lacuna semântica importante no domínio dos registros acadêmicos, responsáveis por gerar indicadores que não refletem a realidade do cenário educacional brasileiro. Com isso, espera-se que os sistemas de coleta e sistemas acadêmicos já existentes e os que virão a ser desenvolvidos, utilizem este modelo semântico como base para as adaptações ou construções necessárias.

7.2 PRINCIPAIS CONTRIBUIÇÕES

A principal contribuição deste trabalho foi a geração da Ontologia para o Domínio dos Registros Acadêmicos do Modelo Brasileiro de Educação, e sua capacidade de propiciar que os sistemas de informação, seja no âmbito do MEC, seja no âmbito das Instituições de Ensino tenham um nível maior de interoperabilidade através da uniformidade semântica proposta através da ONTO-RA.

Outra importante contribuição deste trabalho é a revisão e avaliação do cenário atual da engenharia de ontologias, mais especificamente no que se refere às metodologias de construção de ontologias. Foram avaliadas as principais abordagens para construção de ontologias identificando os principais aspectos de cada proposta. Desta maneira, tem-se o cenário atual das metodologias de construção de ontologias. A avaliação realizada leva a concluir que ainda não foi construído um consenso a respeito de uma metodologia específica para construção de ontologias, mas que cada abordagem poderá ser utilizada em situações específicas. Outra contribuição importante foi a validação da metodologia NEON,

com as devidas adaptações ao contexto do processo de construção, para a reutilização de recursos não ontológicos.

Outro aspecto que pode ser destacado como contribuição é a identificação das principais lacunas semânticas do domínio dos registros acadêmicos do modelo educacional brasileiro, de maneira que se permita utilizar a ontologia gerada para adaptar os sistemas de informação já existentes e construir novos sistemas.

No processo de construção da ontologia ficou evidente que alguns sistemas de coleta de dados e sistemas acadêmicos não estavam de acordo com o que está definido estruturalmente na Lei de Diretrizes e Bases da Educação - LDB. Então, esta é outra contribuição deste trabalho, pois permite identificar que os sistemas de informação, responsáveis por gerir os dados acadêmicos brasileiros, não estão de acordo com o que determina a LDB.

7.3 TRABALHOS FUTUROS

Apesar de apresentar contribuições, este trabalho também deixou algumas lacunas que poderão vir a ser preenchidas através de trabalhos futuros. A primeira lacuna identificada é o fato do OSRD gerado no processo de construção da ontologia levar em consideração apenas requisitos necessários para a geração de indicadores gerenciais e de desempenho acadêmicos, desprezando, portanto, requisitos ligados a classes e propriedades básicas do domínio dos registros acadêmicos da educação brasileira, tais como, nome, documentação e outras informações ligadas aos alunos. Assim, a ontologia apresentada poderá vir a ser estendida para alcançar todo o domínio da educação e servir como artefato para a construção do documento de especificação de requisitos de softwares de gestão acadêmica.

Outro aspecto que poderá ser alcançado diz respeito ao ciclo de vida da ontologia e a sua manutenção. Entende-se que um modelo conceitual semântico como este deverá ser usado e eventualmente sofrer modificações e contribuições de diversas instituições de ensino e do próprio MEC, seja motivado por modificações nas legislações e normas vigentes ou até mesmo por ajustes necessários ao modelo apresentado. Desta forma, é importante estabelecer um processo de manutenção e de busca de consenso dos conceitos que eventualmente precisem ser incluídos à ontologia.

Por fim, outra possibilidade de trabalho futuro é a construção de uma linguagem de integração, baseada em XML, que pode servir para que a coleta de dados realizada pelo MEC junto às instituições de ensino seja realizada de forma automática, a partir de um modelo de dados e de uma sintaxe específica para a representação de cada item.

7.4 ARTIGOS PUBLICADOS

Durante o desenvolvimento deste trabalho foi aceito para publicação na *13th International Conference WWW/INTERNET*, o artigo *ASSESSMENT OF METHODOLOGIES FOR THE CONSTRUCTION OF THE ONTOLOGY FOR BRAZILIAN ACADEMIC RECORDS*. Este artigo discute o processo de escolha da metodologia para a execução deste trabalho que foi apresentado no Capítulo 4.

REFERÊNCIAS

- ANSI/NISO. **Z39.19-2005. Guidelines for the construction, format, and management of monolingual controlled vocabularies**. [S. l.]: [s.n.], 2005.
- ANGELETOU, S.; LEWEN, Holger; VILLAZÓN-TERRAZAS, B. **NeOn Deliverable D2. 2.4 Final version of methods for re-engineering and evaluation**. Technical report, NeOn, 2009.
- BRAY, Tim et al. **Extensible markup language (XML)**. World Wide Web Consortium Recommendation REC-xml-19980210. 1998. Disponível em: <http://www.w3.org/TR/1998/REC-xml-19980210> Acesso em: 12 jan. 2014.
- BERNARAS, Amaia; LARESGOITI, Inaki; CORERA, J. Building and Reusing Ontologies for Electrical Network Applications'. In: ECAI. Pitman, 1996. p. 298-302.
- BORTOLETO, Silvio. **Metodologia para construção de modelos conceituais para aplicação multirrelacional com auxílio de ontologia**. 2010. 153 f. Tese (Doutorado) - Universidade Federal do Rio de Janeiro - UFRJ, 2010.
- BRANDON, Daniel. Recursive database structures. **Journal of Computing Sciences in Colleges**, v. 21, n. 2, p. 295-304, 2005.
- BUCOS, Marian; DRAGULESCU, Bogdan; VELTAN, Marius. Designing a semantic web ontology for E-learning in higher education. In: ELECTRONICS AND TELECOMMUNICATIONS (ISETC), INTERNATIONAL SYMPOSIUM ON. IEEE, 2010. 9., 2010. **Proceedings...** 2010. p. 415-418.
- CANTELE, Regina Cláudia. **Construindo ontologias a partir de recursos existentes: uma prova de conceito no domínio da educação**. 2009. Tese (Doutorado)- Universidade de São Paulo – USP, 2009.
- CHUNG, Í. S.; KIM, Jung-Min. Learning Ontology Design for Supporting Adaptive Learning in e-Learning Environment. In: INTERNATIONAL COMPUTER SCIENCE & INFORMATION TECHNOLOGY, 27, 2012. **Proceedings...** 2012.
- DA SILVA, Daniela Lucas; SOUZA, Renato Rocha; ALMEIDA, Maurício Barcellos. Ontologias e vocabulários controlados: comparação de metodologias para construção. **Ci. Inf**, Brasília, v. 37, n. 3, p. 60-75, dez. 2008.
- DE ALMEIDA FALBO, Ricardo; DE MENEZES, Crédine Silva; DA ROCHA, Ana Regina C. A systematic approach for building ontologies. In: PROGRESS IN ARTIFICIAL INTELLIGENCE—IBERAMIA 98. Springer Berlin Heidelberg, 1998. **Proceedings...** 1998. p. 349-360.
- EUZENAT, Jérôme. Corporate memory through cooperative creation of knowledge bases and hyper-documents. In: KAW, 10., 1996. **Proceedings...**1996.
- FERNÁNDEZ-LÓPEZ, Mariano; GÓMEZ-PÉREZ, Asunción; JURISTO, Natalia. **Methontology: from ontological art towards ontological engineering**. [S. l.]: [s.n.], 1997.
- FERNÁNDEZ-LOPEZ, M. Fernandez. Overview of methodologies for building ontologies. In: OF THE IJCAI-99 WORKSHOP ON ONTOLOGIES AND PROBLEMSOLVING METHODS (KRR5). 1999. **Proceedings...** p. 4.1-4.13, 1999.

- FERNÁNDEZ-LÓPEZ, Mariano; GÓMEZ-PÉREZ, Asunción. Overview and analysis of methodologies for building ontologies. **The Knowledge Engineering Review**, v. 17, n. 2, p. 129-156, 2002.
- GANGEMI, Aldo; PISANELLI, Domenico; STEVE, Geri. Ontology integration: Experiences with medical terminologies. In: **FORMAL ontology in information systems**. IOS Press, Amsterdam, AM, 1998. p. 98-94
- GARCIA-SILVA, Andres et al. A pattern based approach for re-engineering nonontological resources into ontologies. In: **THE SEMANTIC Web**. Springer Berlin Heidelberg, 2008. p. 167-181.
- GENESERETH, Michael R. et al. **Knowledge interchange format-version 3.0: reference manual**. [S. l.]: [s.n.], 1992.
- GÓMEZ-PÉREZ, Asunción; FERNÁNDEZ, Mariano; VICENTE, A. de. **Towards a method to conceptualize domain ontologies**. [S. l.]: [s. n.], 1996.
- GÓMEZ-PÉREZ, Asunción; CORCHO, Oscar; FERNANDEZ-LOPEZ, Mariano. **Ontological engineering: with examples from the areas of knowledge management, e-Commerce and the Semantic Web**. (advanced information and knowledge processing). [S. l.]: [s.n.], 2004a.
- GOMEZ-PEREZ, Asuncion; FERNÁNDEZ-LÓPEZ, Mariano; CORCHO, Oscar. **Ontological engineering**. Heidelberg: Springer, 2004b.
- GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing?. **International journal of human-computer studies**, v. 43, n. 5, p. 907-928, 1995.
- GRÜNINGER, Michael; FOX, Mark S. **Methodology for the design and evaluation of ontologies**. [S. l.]: [s.n.], 1995.
- GUARINO, Nicola. Formal Ontology and information systems. In: **PROCEEDINGS OF THE FOIS'98 – FORMAL ONTOLOGY IN INFORMATION SYSTEMS**, 1998, Trento. **Proceedings ... 1998**
- GUARINO, Nicola; WELTY, Christopher. Evaluating ontological decisions with OntoClean. **Communications of the ACM**, v. 45, n. 2, p. 61-65, 2002. 117
- GUARINO, N.; OBERLE, D.; STAAB, S. What is an Ontology? In: STAAB, S.; STUDER, R. (Eds.). **Handbook on Ontologies**. International handbooks on information systems. 2. ed. Berlin: Springer Verlag, 2009. p. 201-221.
- HEPP, Martin; DE BRUIJN, Jos. GenTax: A generic methodology for deriving OWL and RDF-S ontologies from hierarchical classifications, thesauri, and inconsistent taxonomies. In: **THE SEMANTIC Web: Research and Applications**. Springer Berlin Heidelberg, 2007. p. 129-144.
- HODGE, Gail. **Systems of knowledge organization for digital libraries: beyond traditional authority files**. Washington: Digital Library Federation, Council on Library and Information Resources, 2000.
- HORRIDGE, M. **A Practical guide to building owl ontologies using protégé 4 and CO-ODE Tools**. [S. l.]: [s.n.], 2009.

IEEE Standard for Developing Software Life Cycle Processes. IEEE Std. 1074-1997 (Revision of IEEE Std. 1074-1995; Replaces IEEE Std. 1074.1-1995). [S. 1.]: 1995.

ISO/IEC. **ISO 1087-1:2000. Terminology work. Vocabulary. Part 1: Theory and application.** [S. 1.]: 2000.

ISO/IEC. **International Standard 11179, Part 1, Framework for the specification and standardization of data elements.** [S. 1.]: 1999.

JIMÉNEZ-RUIZ, Ernesto; BERLANGA, R. A view-based methodology for collaborative ontology engineering: an approach for complex applications (VIMethCOE). In: ENABLING TECHNOLOGIES: INFRASTRUCTURE FOR COLLABORATIVE ENTERPRISES, 2006. WETICE'06. IEEE INTERNATIONAL WORKSHOPS ON. IEEE, 15., 2006. **Proceedings...** 2006. p. 376-381.

LENAT, Douglas B.; GUHA, Ramanathan V. **Building large knowledge-based systems; representation and inference in the Cyc project.** [S. 1.]: Addison-Wesley Longman Publishing Co., Inc., 1989.

LOPEZ, M. Fernandez. Overview of methodologies for building ontologies. In: IJCAI- 99 WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING METHODS (KRR5). 1999, Seattle. **Proceedings ...** 1999. p. 4.1-4.13.

Maedche, A. Staab , S. **Ontology Learning for the Semantic Web.** IEEE Intelligent Systems. [S.1.], 2001

MALVIYA, Naveen et al. Developing University Ontology using Protégé OWL Tool: Process and Reasoning. **International Journal of Scientific & Engineering Research**, v. 2, n. 9, p. 1-8, 2011.

NOY, Natalya. Order from Chaos. **ACM Queue**, v. 3 , n. 8, p. 42-49, 2005

NOY, Natalya F.; DEBORAH L. **McGuinness. Ontology development 101: a guide to creating your first ontology.** Stanford Knowledge Systems Laboratory Technical 118 Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001- 0880. [S.1.]: Alto, 2001.

PANETTO, Hervé; DASSISTI, Michele; TURSI, Angela. ONTO-PDM: Product-driven ONTOlogy for Product Data Management interoperability within manufacturing process environment. **Advanced Engineering Informatics**, v. 26, n. 2, p. 334-348, 2012

RAUTENBERG, Sandro; TODESCO, José L.; GAUTHIER, Fernando A. O. Processo de desenvolvimento de ontologias: uma proposta e uma ferramenta. **Rev. Tecnol**, v. 30, p. 133-144, 2009.

REED, Stephen L.; LENAT, Douglas B. Mapping ontologies into Cyc. In: AAAI 2002 CONFERENCE WORKSHOP ONTOLOGIES FOR THE SEMANTIC WEB. 2002, Edmonton, Canadá. **Proceedings ...** 2002. p. 1-6.

RICHARD BENJAMINS, V. et al. (KA)²: building ontologies for the Internet: a midterm report. **International Journal of Human-Computer Studies**, v. 51, n. 3, p. 687- 712, 1999.

SABOU, Marta. D2. **2.1 Methods for selection and integration of reusable components from formal or informal user specifications.** [S.1.]: [s.n.], 2007.

SANTOS, Eluzai Souza dos. **Uma proposta de integração de sistemas computacionais utilizando ontologias.** [S.l.]: [s.n.], 2006.

SANTOS, Neide; CAMPOS, Fernanda CA; VILLELA, R. M. M. B. Ontologia para o domínio da educação mediada pela Web. **Taller Intenacional de Software Educativo**, v. 8, 2003.

SCATALON, Lilian Passos; GARCIA, Rogério Eduardo; CORREIA, Ronaldo Celso Messias. Introdução a ontologias e suas aplicações. **Omnia Exatas**, v.3, n.2, p.19- 28, 2010

STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge Engineering: Principles and Methods. **Data and Knowledge Engineering**, v. 25, n. 1-2, p. 161-197, 1998.

SUÁREZ-FIGUEROA, Mari Carmen. **NeOn Methodology for building ontology networks: specification, scheduling and reuse.** 2010. 268 f.. Tese (Doutorado) – Universidade Politécnica de Madrid, 2010.

SUÁREZ-FIGUEROA, Mari Carmen. D5.4.1. **NeOn Methodology for Building Contextualized Ontology Networks.** [S.l.]: [s.n.], 2008.

SUÁREZ-FIGUEROA, Mari Carmen. 2007. **D5. 3.1 NeOn Development Process and Ontology Life Cycle.** [S.l.]: [s.n.], 2007.

SUN, L. J.; DING, Bo. Ontology-based semantic interoperability among heterogeneous CAD systems. **Information Technology Journal**, v. 9, n. 8, p. 1635- 1640, 2010. 119

SURE, York et al. **OntoEdit: Collaborative ontology development for the semantic web.** Berlin : Springer Berlin Heidelberg, 2002.

SURE, York; STAAB, Steffen; STUDER, Rudi. Methodology for development and employment of ontology based knowledge management applications. **ACM SIGMOD Record**, v. 31, n. 4, p. 18-23, 2002.

SWARTOUT, Bill et al. Toward distributed use of large-scale ontologies. In: WORKSHOP ON KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS, 10., 1996. **Proceedings ...** 1996.

VANDER WAL. **Folksonomy coinage and definition.** 2007. Disponível em: <<http://vanderwal.net/folksonomy.html>> Acesso em: 10 jan. 2014.

USCHOLD, Michael; KING, Martin. **Towards a methodology for building ontologies.** Artificial Intelligence Applications Institute. Edinburgh: University of Edinburgh, 1995.

USCHOLD, Mike; GRUNINGER, Michael. Ontologies: Principles, methods and applications. **The knowledge engineering review**, v. 11, n. 2, p. 93-136, 1996.

VILLAZÓN-TERRAZAS, Boris. **D2. 2.2 Methods and Tools Supporting Reengineering.** [S.l.]: [s.n.], 2009.

VRANDECIC, Denny et al. The DILIGENT knowledge processes. **Journal of Knowledge Management**, v. 9, n. 5, p. 85-96, 2005.

**ANEXO A - Tabelas do esquema de classificação resultante da fusão dos
candidatos a recursos não ontológicos**

Código	Nome
1	Aluno

Código	Código Nível 1	Nome
12	1	Cor/Raça
13	1	Nacionalidade
14	1	Deficiência Física Sensorial ou Intelectual
16	1	Mobilidade Acadêmica
17	1	PARFOR
18	1	Origem escolar ensino médio
19	1	Forma de ingresso
110	1	Programa de reserva de vagas
111	1	Apoio social
112	1	Atividades extra-curriculares
113	1	Turma
114	1	Tipo de rendimento
115	1	Zona de residência

Código	Código Nível 2	Nome
161	16	Tipo Mobilidade Acadêmica

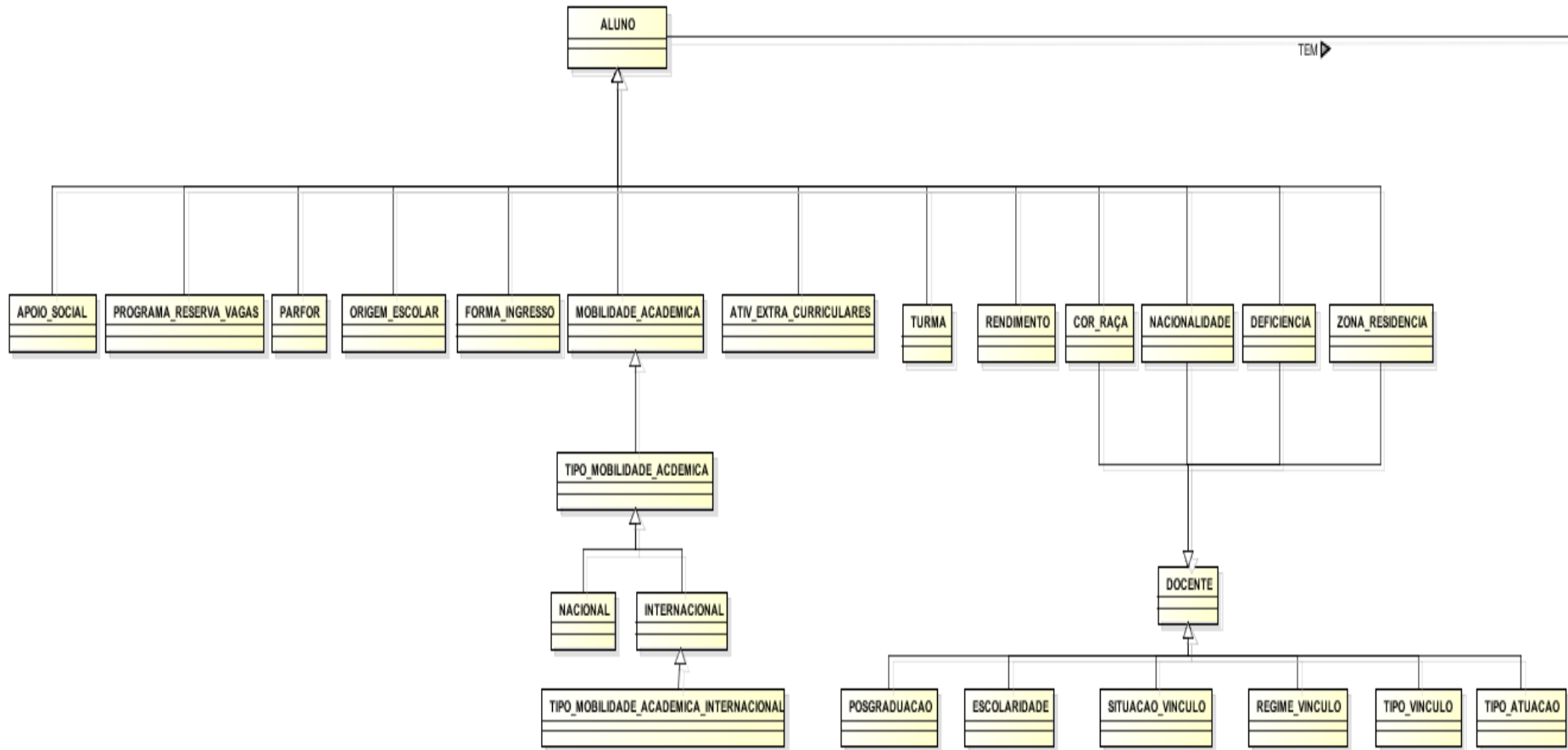
Código	Código Nível 3	Nome
1611	161	Nacional
1612	161	Internacional

Código	Código Nível 4	Nome
16121	1612	Tipo Mobilidade Acadêmica Internacional
	Código	Nome
	2	Docente
Código	Código Nível 1	Nome
22	2	Cor/Raça
23	2	Nacionalidade
24	2	Deficiência Física Sensorial ou Intelectual
25	2	Situação de Vínculo
26	2	Escolaridade
27	2	Pós-graduação
28	2	Regime de vínculo
29	2	Tipo de vínculo
210	2	Tipo de atuação
211	2	Zona de residência
	Código	Nome
	3	Processo Seletivo
Código	Código Nível 1	Nome
31	3	Inscrições

32	3	Candidato
33	3	Oferta de vaga em curso
	Código	Nome
	14	AlunoXCurso
Código	Código Nível 1	Nome
141	14	Situação de Vínculo
142	14	Ingressante
143	14	Matricula
144	14	Concluente
145	14	Integralizado
146	14	Retido
147	14	Evadido
148	14	ciclo de matrícula
149	14	turma de ingresso
	Código	Nome
	4	Curso
Código	Código Nível 1	Nome
41	4	Tempo de integralização
42	4	Modalidade de Oferta
43	4	Nível de Ensino
44	4	Modalidade de Ensino
45	4	Situação
46	4	Certificação intermediária
47	4	Disciplina semipresencial
48	4	Disciplina
49	4	Internato pleno
Código	Código Nível 2	Nome
431	43	Educação Básica
432	43	Educação Superior
433	43	Formação Inicial e Continuada - FIC
471	47	Carga horária semipresencial
472	48	Carga horária
Código	Código Nível 3	Nome
4311	431	Etapas de Ensino
4321	432	Níveis Acadêmicos
Código	Código Nível 4	Nome
43111	4311	Educação Infantil – creche
43112	4311	Educação Infantil - pré-escola
43113	4311	Ensino Fundamental
43114	4311	Ensino Médio
43115	4311	Ensino Médio – EPT
43116	4311	Ensino Médio - Normal Magistério
43211	4321	Seqüencial
43212	4321	Graduação
43213	4321	Pós-graduação

43214	4321	Extensão
43215	4321	Pós-graduação tecnológica
Código	Código Nível 5	Nome
431151	43115	Forma de Oferta
432121	43212	Grau Acadêmico
432131	43213	Programas - Pós-Graduação
432151	43215	Programas - Pós-Graduação Tecnológica
Código	Código Nível 6	Nome
4311511	431151	Articulada INTEGRADA
4311512	431151	Articulada CONCOMITANTE
4311513	431151	SUBSEQUENTE
4321211	432121	Bacharelado
4321212	432121	Licenciatura
4321213	432121	Tecnológico
4321311	432131	Mestrado
4321312	432131	Mestrado Profissional
4321313	432131	Doutorado
4321314	432131	Especialização
4321315	432131	Aperfeiçoamento
4321511	432151	Mestrado
4321512	432151	Mestrado Profissional
4321513	432151	Doutorado
4321514	432151	Especialização
4321515	432151	Aperfeiçoamento

ANEXO B - Modelo Conceitual



ANEXO B - Modelo Conceitual - Continuação

