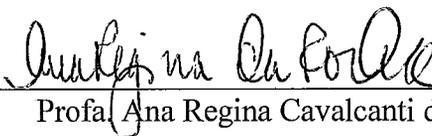


APOIO À CONSTRUÇÃO DE BASE DE DADOS DE PESQUISA EM AMBIENTES
DE DESENVOLVIMENTO DE SOFTWARE ORIENTADOS A DOMÍNIO

Ana Mirtes Maciel Fouro

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM
ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

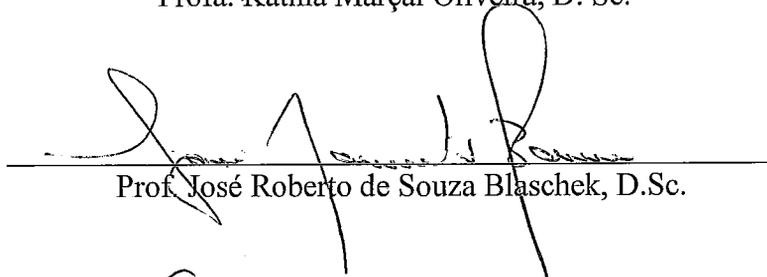
Aprovada por:



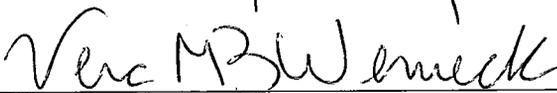
Profª Ana Regina Cavalcanti da Rocha, D. Sc.



Profª. Káthia Marçal Oliveira, D. Sc.



Prof. José Roberto de Souza Blaschek, D.Sc.



Profª. Vera Maria Benjamim Werneck, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

MARÇO DE 2002

FOURO, ANA MIRTES MACIEL

Apoio à Construção de Bases de Dados de
Pesquisa em Ambientes de Desenvolvimento de
Software Orientados a Domínio [Rio de Janeiro] 2002

xi, 157p. 29,7 cm (COPPE/UFRJ, M.Sc.,
Engenharia de Sistemas, 2002)

Tese - Universidade Federal do Rio de
Janeiro, COPPE

1. Ambiente de Desenvolvimento de Software
2. Orientação a Domínio
3. Ontologia
4. Modelagem de Dados
5. Abordagem Entidade Relacionamento

I. COPPE/UFRJ II. Título (série)

Para aprender a sabedoria (Eclesiástico 6 18:37)

Meu filho, aceita a instrução desde teus jovens anos; ganharás uma sabedoria que durará até à velhice.

Vai ao encontro dela, como aquele que lava e semeia; espera pacientemente seus excelentes frutos;

terás alguma pena em cultivá-la, mas, em breve, comerás os seus frutos.

Quanto à sabedoria é amarga para os ignorantes! O insensato não permanecerá junto a ela.

Ela lhes será como uma pesada pedra de provação, e ele não tardará a desfazer-se dela.

Pois a sabedoria que instrui justifica o seu nome, não se manifesta a muitos; mas, naqueles que a conhecem, persevera, até (tê-los levado) à presença de Deus.

Escuta, meu filho, recebe um sábio conselho, não rejeites minha advertência.

Mete os teus pés nos seus grilhões, e teu pescoço em suas correntes.

Abaixa teu ombro para carregá-la, não sejas impaciente em suportar seus liames.

Vem a ela com todo o teu coração. Guarda seus caminhos com todas as tuas forças.

Segue-lhe os passos e ela se dará a conhecer; quando a tiveres abraçado, não a deixes.

Pois acharás finalmente nela o teu repouso. E ela transformar-se-á para ti em um motivo de alegria.

Seus grilhões ser-te-ão uma proteção, um firme apoio; suas correntes te serão um adorno glorioso,

pois nela há uma beleza que dá vida, e seus liames são ligaduras que curam.

Com ela te revestirás como de uma vestimenta de glória, e a porás sobre ti como uma coroa de júbilo.

Meu filho, se me ouvires com atenção, serás instruído; se submeteres o teu espírito, tornar-te-ás sábio.

Se me deres ouvido, receberás a doutrina. Se gostares de ouvir, adquirirás a sabedoria.

Permaneça na companhia dos doutos anciãos, une-te de coração à sua sabedoria, a fim de que possas ouvir o que dizem de Deus, e não te escapem suas louváveis máximas.

Se vires um homem sensato, madruga para ir ter com ele, desgaste o teu pé o limiar de sua porta.

Concentra teu pensamento nos preceitos de Deus, sê assíduo à meditação de seus mandamentos.

E ele próprio te dará um coração, e ser-te-á concedida a sabedoria que desejas.

A todos, que compartilharam sua sabedoria e conhecimento contribuindo para a realização deste trabalho.

Agradecimentos

À professora Ana Regina, por ter dedicado seu tempo e conhecimento na orientação deste trabalho e mais ainda pelo ensinamento que será eterno.

À Káthia Oliveira por mostrar a cada momento compreensão, luta, paciência, inteligência, dedicação e incentivo para a realização deste trabalho.

Aos professores José Roberto Blaschek e Vera Werneck pela honra de tê-los em minha banca examinadora.

À Embrapa por tornar possível este trabalho e acreditar no seu resultado.

Aos amigos da Embrapa Afonso Cardoso, Michell Olívio, Rubenise Gato, Lucilda, Márcia Maués, Antônio de Brito e Lidáurea que empenharam seu trabalho me auxiliando durante os momentos em que precisei.

Aos professores e funcionários do Programa, especialmente, Cláudia Werner, Solange, Cláudia, Sueli, Mercedes, Lúcia, Marli, Frederico, Sonia e Tia Gercina, sempre solícitos tornando essa experiência acadêmica mais fácil e alegre.

À Ana Paula, por sua presença sempre simpática e disponível, principalmente pelo apoio remoto.

Aos pesquisadores Miguel Moné (Museu Nacional/UFRJ) e Inocêncio Gorayeb (Museu Paraense Emílio Goeldi), que com uma carinhosa recepção e competência muito me auxiliaram na execução deste trabalho.

Aos companheiros e novos amigos Selma Créspio, Alexandre Soares, e Moisés Júnior, por terem tornado menos árdua, mais alegre e divertida essa caminhada.

À minha mãe Maria pelo carinho, dedicação, torcida e por ser a principal responsável pela minha formação acadêmica e pessoal.

Ao Gleison pelas inúmeras vezes que me auxiliou e ao Sômulo pelo suporte prestado sempre que solicitado.

Ao Professor Ricardo Falbo e a amiga Cátia Galotta pelo apoio.

À pessoa amiga, carinhosa, paciente, incentivadora, compreensiva e muito generosa, Valdemir, que Deus colocou no meu caminho para tornar toda a vida mais feliz.

E, principalmente, a Deus pela saúde e força, permitindo que eu continue na busca pela sabedoria.

Résumé da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

APOIO À CONSTRUÇÃO DE BASE DE DADOS DE PESQUISA EM AMBIENTES DE DESENVOLVIMENTO DE SOFTWARE ORIENTADOS A DOMÍNIO

Ana Mirtes Maciel Fouro

Março/2002

Orientadores: Ana Regina Cavalcanti da Rocha

Káthia Marçal de Oliveira

Programa: Engenharia de Sistemas e Computação

Os dados resultantes da execução de projetos de pesquisa e necessários ao planejamento e aperfeiçoamento de novos projetos, possuem um valor agregado muito alto; pois o processo utilizado para sua obtenção, além de demorado, é muito caro. Estes dados necessitam ser preservados de forma apropriada, pois constituem um dos componentes essenciais do patrimônio de instituições com atividade em ciência e tecnologia. Nesse sentido, bases de dados que atendam às reais necessidades de informação de um determinado ambiente, são largamente utilizadas para armazenamento desse importante patrimônio das instituições. Visando apoiar e homogeneizar o desenvolvimento de bases de dados para projetos de pesquisa específicos em um domínio, definimos um processo padrão a ser utilizado na construção de bases de dados de pesquisa. Organizamos ainda, um procedimento apoiado pelo conhecimento do domínio, modelado sob a forma de ontologia, visando apoiar a modelagem conceitual de dados. Finalmente, apresentamos, como exemplo desta abordagem, **INSECTA**, um Ambiente de Desenvolvimento de Software Orientado ao Domínio de Entomologia, particularizado para a construção de bases de dados de pesquisa.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

SUPPORT TO THE DEVELOPMENT OF RESEARCH DATA BASE IN DOMAIN-ORIENTED SOFTWARE DEVELOPMENT ENVIRONMENTS

Ana Mirtes Maciel Fouro

March/2002

Advisors: Ana Regina Cavalcanti da Rocha

Káthia Marçal de Oliveira

Department: Computer Science and System Engineering

The collect data from research projects necessary for planning and improvement of new projects has a high value for institutions, considering the process applied for obtaining it, which is quite time consuming and expensive as well. Those data consist in one of the essential components of the property of institutions which focus on Science and Technology, so they need to be preserved in an appropriate manner. In that sense, data bases which fullfill the actual demands for information in a given environment are widely used for storage of this important institution's property. With the goal of support the development of data bases related to research projects which are specific to a given domain, we defined a standard process to be applied in the development of research data bases. We also organized a procedure that uses the knowledge of the domain, organized with domain ontology, to support the conceptual data modeling. Finally, we present, as an example of this approach, *INSECTA*, a Domain-Oriented Software Development Environment for Entomology, restricted to the development of research data bases.

Índice

Capítulo 1 – Introdução.....	1
1.1 – Motivação.....	1
1.2 – Contexto e Objetivo da Tese.....	2
1.3 – Contribuições.....	4
1.4 – Organização da Tese.....	5
Capítulo 2 – Ontologia e Modelagem de Dados.....	7
2.1 – Ontologias.....	7
2.1.1 – Definição e Classificação.....	8
2.1.2 – Componentes e Engenharia de Ontologias.....	10
2.1.3 – Utilização e Reutilização de Ontologias de Domínio.....	12
2.2 – Modelagem de Dados.....	13
2.2.1 – Projeto Conceitual.....	15
2.2.2 – Projeto Lógico.....	17
2.2.3 – Projeto Físico.....	17
2.3 – Usos e Benefícios da Modelagem de Dados.....	17
2.4 – Relação entre Ontologias de Domínio e Modelo de Dados.....	18
2.5 – Considerações Finais.....	20
Capítulo 3 – Processo de Construção de Bases de Dados de Pesquisa.....	22
3.1 – Processo de Software.....	22
3.2 – Processo para Construção de Bases de Dados de Pesquisa.....	24
3.3 – Considerações Finais.....	40
Capítulo 4 – Definição de Bases de Dados com Apoio de Ontologias de Domínio.....	41
4.1 – Introdução.....	41
4.2 – Uso de ontologias de domínio nas etapas do processo de construção de bases de dados.....	42
4.3 – Projeto Conceitual com base em Ontologias de Domínio.....	46
4.3.1 – Procedimento para Modelagem Conceitual ER.....	46
4.4 – Mapeamento dos construtores da ontologia para os construtores da modelagem conceitual.....	56
4.5 – Considerações Finais.....	64

Capítulo 5 – INSECTA: Um Ambiente de Desenvolvimento de Software Orientado ao Domínio de Entomologia.....66

5.1 – A Estação TABA e Ambientes de Desenvolvimento de Software Orientados a Domínio.....	66
5.1.1 – Ambiente de Desenvolvimento de Software.....	66
5.1.2 – A Estação TABA.....	68
5.2 – Teoria do Domínio para Entomologia.....	71
5.2.1 – Identificação do Propósito.....	72
5.2.2 – Definição.....	73
5.2.2.1 – Subteoria Anatomia de Insetos.....	75
5.2.2.2 – Sub-teoria Biologia de Insetos.....	79
5.2.2.3 – Sub-teoria Ecologia de Insetos.....	81
5.2.2.4 – Sub-teoria Classificação de Insetos.....	85
5.2.2.5 – Sub-teoria Processo de Trabalho do Entomólogo.....	87
5.2.3 – Formalização e Codificação.....	91
5.2.4 – Atividades de Apoio.....	92
5.3 – Definição e Instanciação do Ambiente Insecta.....	98
5.3.1 – Especialização do Processo para o Ambiente Insecta.....	101
5.3.2 – Edição da Teoria do Domínio para Entomologia.....	103
5.3.3 – Instanciação do ADSOD Insecta.....	106
5.4 – Exemplo da Execução da Atividade Projeto Conceitual para uma Base de Dados no Domínio de Entomologia.....	108
5.4.1 – Especificação de Requisitos de Dados.....	108
5.4.2 – Projeto Conceitual.....	110
5.5 – Considerações Finais.....	119

Capítulo 6 – Conclusões e Perspectivas Futuras.....120

Referências Bibliográficas.....123

Anexos.....132

Anexo 1 – Descrição de Conceitos da Teoria do Domínio de Entomologia.....	132
Anexo 2 – Avaliação da Teoria do Domínio de Entomologia.....	152
Anexo 3 – LINGO: Linguagem Gráfica para descrever Ontologias.....	156

Índice de Figuras

Figura 2.1 – Classificação de ontologias segundo GUARINO (1998).....	9
Figura 2.2 – Engenharia de ontologias.....	10
Figura 2.3 – Etapas (níveis de abstração) na construção de bases de dados.....	15
Figura 3.1 – Fases do Ciclo de Vida do Processo para Construção de Bases de Dados de Pesquisa.....	25
Figura 3.2 – Atividades do Processo Para Construção de Bases de Dados de Pesquisa.....	26
Figura 3.3 – Estrutura hierárquica da abordagem GQM.....	27
Figura 3.4 – Estrutura hierárquica da abordagem OQD.....	27
Figura 4.1 – Procedimento para Modelagem Conceitual.....	47
Figura 4.2 – Equivalência entre relação binária ontológica e relacionamento binário do DER.....	57
Figura 4.3 – Equivalência entre relação ternária ontológica e estrutura de agregação do DER.....	58
Figura 4.4 – Equivalência da relação entre instâncias ontológicas e o auto-relacionamento do DER.....	59
Figura 4.5 – Equivalência da estrutura todo-parte ontológica e relacionamentos binários do DER.....	60
Figura 4.6 – Equivalência da estrutura subtipo-de/supertipo-de ontológica e a estrutura de generalização/especialização do DER.....	61
Figura 4.7 – Equivalência da estrutura condicionante Ou-exclusivo ontológica e a estrutura de relacionamentos mutuamente exclusivos do DER.....	62
Figura 4.8 – Equivalência da estrutura condicionante de Obrigatoriedade ontológica e a estrutura de relacionamentos Contingentes do DER.....	62
Figura 5.1 – Processo de Desenvolvimento da Teoria do Domínio.....	68
Figura 5.2 – Sub-teorias da Teoria de Domínio de Entomologia.....	74
Figura 5.3 – Conceitos da Sub-teoria da Anatomia de Insetos.....	78
Figura 5.4 – Conceitos da Sub-teoria Biologia de Insetos.....	80
Figura 5.5 – Conceitos da Sub-teoria Ecologia de Insetos.....	83
Figura 5.6 – Conceitos da Sub-teoria Classificação de insetos.....	87
Figura 5.7 – Conceitos da Sub-teoria Processo de trabalho do entomólogo.....	90
Figura 5.8 – Avaliação da Teoria do Domínio de Entomologia por perfil dos avaliadores.....	94

Figura 5.9 – Sugestões de utilização da Teoria do Domínio de Entomologia	95
Figura 5.10 – Visão integrada das sub-teorias que compõem a Teoria do Domínio de Entomologia.....	97
Figura 5.11 – Possíveis formas de definição, especialização e instanciação de processo de software para um ADSOD na Estação TABA.....	100
Figura 5.12 – Identificação das atividades do Processo Padrão para Construção de Bases de Dados de Pesquisa, onde se incluiu a sub-atividade Investigação do Domínio.....	102
Figura 5.13 – Edição da atividade Projeto Conceitual através do uso da ferramenta EDIT-PRO.....	103
Figura 5.14 (a) – Opções do menu Inserir da ferramenta EDITED.....	104
Figura 5.14 (b) – Inserção de subteorias na ferramenta EDITED.....	104
Figura 5.14 (c) – Inserção de Conceitos na ferramenta EDITED.....	105
Figura 5.14(d) – Inserção de Relações na ferramenta EDITED.....	105
Figura 5.15 (a) – Opções da Ferramenta para Instanciação de ADS.....	106
Figura 5.15 (b) – Definição do ambiente <i>Insecta</i>	106
Figura 5.15 (c) – Instanciação do ambiente <i>Insecta</i>	107
Figura 5.16 – Ambiente <i>Insecta</i>	108
Figura 5.17 – Definição dos Dados que devem responder às Questões.....	109
Figura 5.18 – Identificação das Entidades que devem responder às Questões.....	111
Figura 5.19 – Dicionarização da Entidade Inseto que comporá o MER.....	112
Figura 5.20 – Identificação dos Atributos da Entidade Inseto.....	113
Figura 5.21 – Dicionarização dos Atributos da Entidade Localidade.....	113
Figura 5.22 – Dicionarização de uma estrutura de Generalização/Especialização.....	114
Figura 5.23 – Identificação da estrutura de agregação.....	117
Figura 5.24 – Dicionarização de relacionamentos e estruturas de agregação.....	117
Figura 5.25 – Diagrama ER inicial da base de dados.....	118

Índice de Tabelas

Tabela 2.1 – Critérios para o projeto e avaliação de ontologias.....	11
Tabela 2.2 – Mapeamento entre conceitos ontológicos e modelos de dados.....	19
Tabela 4.1 – Tipos de Restrições de Integridade para a abordagem relacional.....	43
Tabela 4.2 – Realização do procedimento para modelagem conceitual com base no conteúdo de uma ontologia de domínio.....	55
Tabela 4.3 – Equivalência entre Construtores Ontológicos e Construtores do MER.....	63
Tabela 4.4 – Uso dos Axiomas Formais da Ontologia na Modelagem Conceitual.....	64
Tabela 5.1 – Correspondência entre tarefas e sub-teorias.....	91
Tabela 5.2 – Definição dos Objetivos e Questões.....	109
Tabela 5.3 – Relação entre os Dados da Abordagem OQD e as entidades identificadas.....	111
Tabela 5.4 – Relacionamentos que correspondem às questões estabelecidas.....	116

Capítulo 1

Introdução

Este capítulo apresenta, os fatores que motivaram a realização deste trabalho, o contexto e os objetivos da tese, as contribuições que este trabalho se propõe a fornecer e sua forma de organização.

1.1 – Motivação

Instituições de pesquisa agropecuária necessitam contar com dados ambientais em sentido amplo, como por exemplo: dados de clima, de solos, de recursos hídricos, de flora, de fauna, econômicos, sociais e demográficos, uma vez que as condições ambientais exercem enorme influência sobre a produção e produtividade agrícola e pecuária.

As bases de dados constituem uma forma de armazenar dados ambientais, além de ser uma importante fonte de informação que permite gerar conhecimento, possibilitando o aperfeiçoamento do planejamento, traduzido em decisão racional, orientando a ação da pesquisa. As mesmas deverão estar dentro de padrões estabelecidos, de forma que dêem apoio ao ambiente que integram.

Um dos mais importantes centros de pesquisa da Embrapa¹, onde desenvolvemos nossas atividades há doze anos, é responsável pela pesquisa na Amazônia Oriental e tem como principal programa de pesquisa a avaliação dos recursos naturais e socioeconômicos da região, tendo na informação/conhecimento um recurso dos mais importantes e estratégicos, constituindo-se ainda em um de seus produtos finais.

Desta forma, os dados gerados pela pesquisa, ao longo de sua trajetória institucional de sessenta e um anos, constituem um dos componentes essenciais do seu patrimônio, possuindo um valor agregado muito alto, pois o processo utilizado para sua obtenção é muito caro, incluindo dentre outros, a experimentação no campo, a simulação e as pesquisas em laboratórios.

¹ Empresa Brasileira de Pesquisa Agropecuária vinculada ao Ministério da Agricultura e do Abastecimento e criada em 26 de abril de 1973. Sua missão é viabilizar soluções para o desenvolvimento sustentável do agronegócio brasileiro por meio de geração, adaptação e transferência de conhecimentos e tecnologias, em benefício da sociedade.

No intuito de planejar, gerir e utilizar, da melhor forma possível este patrimônio e ao mesmo tempo recurso vital, é necessário que estes dados sejam preservados na forma apropriada, já que numerosos processos de pesquisa e de apoio à pesquisa exigem a manipulação de grande quantidade de dados/informação de forma rotineira.

Com base neste contexto, os profissionais responsáveis pela estruturação e validação de bases de dados institucionais, defrontam-se frequentemente com diversos assuntos (domínios) específicos, alguns deles com poucos especialistas que detém o conhecimento relativo aos dados resultantes das pesquisas realizadas, e necessários às pesquisas a realizar.

Além disso, na Embrapa, como em qualquer outra organização, existe muita pressão à produção rápida de software de qualidade a um custo razoável, desafiando a equipe de desenvolvimento de software, composta de poucos profissionais, a buscar formas de produzir software para vários domínios de aplicação, ferramentas para manipular software, além da utilização de um processo de desenvolvimento, que facilite tanto o desenvolvimento rápido, quanto as adaptações futuras e sua integração a novos contextos.

Assim sendo, desenvolvemos este trabalho, visando organizar o conhecimento de domínios específicos com os quais os desenvolvedores de bases de dados se defrontam com frequência, permitindo sua utilização como apoio à execução das atividades que compõem um processo de desenvolvimento de software adequado a este tipo específico de problema.

1.2 – Contexto e Objetivo da Tese

A Embrapa Amazônia Oriental, como uma instituição com atividade em ciência e tecnologia (C&T) atuante na Amazônia, tem que considerar a informação como uma força impulsionadora, fundamental para destacá-la pelos ganhos de qualidade e produtividade na realização de sua missão institucional.

Os dados gerados por suas pesquisas, entram no sistema de pesquisa na forma de informação nos diferentes níveis de descentralização onde é processada e sua saída deve proporcionar impactos sociais, econômicos e ambientais positivos de forma desconcentrada. Isso acontece devido à variedade de condições ecológicas, de características sócio-econômicas e de sistemas de produção.

Desta forma, existe uma grande necessidade pela disponibilização rápida de bases de dados de pesquisa, para grupos pequenos de especialistas numa área específica, e os recursos humanos, materiais e financeiros disponíveis para a construção dessas bases são muito restritos, devido os recursos públicos disponíveis para a execução de projetos de pesquisa estarem sendo cada vez mais reduzidos.

Além disso, no seu contexto organizacional não há qualquer tipo de processo de desenvolvimento de software definido e padronizado para uso na construção de bases de dados e os desenvolvedores trabalham com poucas estimativas precisas de tempo e de esforços para o desenvolvimento.

Outro cenário motivador para a realização deste trabalho é a existência, no âmbito da Embrapa, de um programa de pesquisa no SEP (Sistema Embrapa de Planejamento) identificado como Programa 14 - *Intercâmbio e Produção de Informação em Apoio às Ações de Pesquisa e Desenvolvimento*, onde está inserido o projeto do SEI (Sistema Embrapa de Informação). Um dos seus objetivos é a estruturação (implementação e implantação), validação e gerenciamento de bases de dados de pesquisa nas unidades da Embrapa.

Neste contexto, a modelagem do conhecimento específico dos diversos domínios e dos dados relacionados a esses domínios, passa a ter grande importância na construção de bases de dados de pesquisa, o que nos proporciona a motivação e dois objetivos para a realização deste trabalho: a definição de um processo padrão para desenvolvimento de bases de dados de pesquisa e a modelagem do conhecimento específico de um domínio, de forma que possa ser utilizado ao longo do processo de desenvolvimento, permitindo ainda o apoio à modelagem conceitual de dados de pesquisa e para pesquisa.

A existência de um importante projeto de pesquisa no Programa de Engenharia de Sistemas e Computação da COPPE/UFRJ nos chamou a atenção. O Projeto TABA (ROCHA *et al.*, 1990) tem como produto principal a Estação TABA, que se constitui de um meta-ambiente capaz de instanciar Ambientes de Desenvolvimento de Software (ADS) que sejam adequados a um projeto específico em um determinado domínio de aplicação, tornando operacional o ambiente definido (TRAVASSOS, 1994).

Uma das características da Estação TABA que nos motivou a utilizá-la, foi a proposta de OLIVEIRA (1999) para construção de Ambientes de Desenvolvimento de Software Orientados a Domínio (ADSOD), definindo a organização do conhecimento e o seu uso ao longo de todo o processo de desenvolvimento para um domínio específico.

Isto significa, que podemos utilizar a infra-estrutura da Estação TABA para criar uma ferramenta capaz de apoiar de forma semi-automatizada o uso do processo definido e do conhecimento modelado, na construção de bases de dados de pesquisa.

Diante de todo este cenário envolvendo as necessidades: pela definição de processos de desenvolvimento adequados à construção de bases de dados de pesquisa; pela construção de bases de dados temáticas (diferentes domínios) como um dos produtos principais das instituições de C&T, implicando na necessidade de uso do conhecimento do domínio ao longo de sua estruturação e validação; pela construção e utilização de ambientes de desenvolvimento de software que considerem o conhecimento do domínio, ocorreu a motivação para a realização deste trabalho.

O objetivo desse trabalho é, portanto, definir um processo de desenvolvimento de software para construção de bases de dados de pesquisa, com ênfase na modelagem conceitual de dados baseada na abordagem Entidade-Relacionamento (CHEN, 1976). Além disso, tem-se o objetivo de verificar como ontologias podem ser utilizadas ao longo de todo o processo de desenvolvimento de software definido e auxiliando de forma mais específica a execução da modelagem conceitual de dados.

Foi desta forma, construído, *Insecta*, um ADSOD específico para o domínio de Entomologia. O conhecimento do domínio foi elicitado e organizado na forma de ontologias (CHANDRASEKARAN *et al.*, 1999), o processo de desenvolvimento específico para construção de bases de dados foi definido, um procedimento de modelagem conceitual de dados foi organizado e foi verificado ainda, como a Teoria do Domínio (OLIVEIRA, 1999) construída, pode apoiar a execução da atividade projeto conceitual no processo definido.

1.3 – Contribuições

Organizar informação implica, também, em promover o acesso real aos seus conteúdos. Informações pontuais e desorganizadas nunca foram suficientes para propiciar as possíveis respostas desejadas pelos consumidores de informação. Este é o problema historicamente vivido no atendimento de demandas por informação, substancialmente agravado nos dias correntes com a disponibilização de informações em redes de computadores e comunicação de dados.

Muitas das informações e conhecimento da Embrapa encontram-se dispersos em dados isolados e não estruturados ou em documentos impressos. Visando estruturar e validar estes dados, de forma a utilizar métodos e técnicas já propostos com vistas à obtenção da qualidade e produtividade no desenvolvimento de software, é que este trabalho foi realizado.

Podemos destacar como contribuições deste trabalho:

- a) a definição de um processo de desenvolvimento de bases de dados de pesquisa que possa ser utilizado, com os devidos ajustes, por qualquer instituição de C&T;
- b) a organização de um procedimento para modelagem conceitual de dados, com base na abordagem entidade-relacionamento, para ser utilizado na execução da atividade Projeto Conceitual, componente do processo definido;
- c) o relacionamento das diversas atividades do procedimento organizado, com o conhecimento modelado por meio de ontologias de domínio e disponíveis na forma de Teoria do Domínio;
- d) o conhecimento básico de Entomologia, modelado e organizado na forma de Teoria do Domínio, necessário e de grande utilidade para uma boa construção de bases de dados para este domínio;
- e) a construção do ADSOD *Insecta*.

A construção de um ADSOD implica na necessidade de execução de duas etapas: a construção de uma Teoria do Domínio, com base em ontologias de domínio, e a definição de um processo de desenvolvimento de software. E, essa construção é feita utilizando-se a infra-estrutura da Estação TABA. Todos estes aspectos serão detalhados nos capítulos deste trabalho.

1.4 – Organização da Tese

Este trabalho é composto de mais cinco capítulos além desta introdução e mais três anexos.

O capítulo 2 aborda o tema de ontologias tratando-se da definição e classificação adotadas neste trabalho, seus componentes e sua engenharia, bem como a utilização e reutilização de ontologias de domínio. Apresenta ainda, a modelagem conceitual de dados, mostrando como a abordagem Entidade-Relacionamento (ER) é utilizada na modelagem conceitual de dados; para que é utilizada esta modelagem e quais os benefícios de sua utilização.

Com base nestas duas visões conceituais, ainda no capítulo 2, é feita a relação entre ontologias de domínio e modelagem conceitual de dados, através da análise das características apresentadas por alguns trabalhos existentes na literatura que as relacionam.

No capítulo 3 apresentamos o processo padrão para construção de bases de dados de pesquisa, que será inserido na Estação TABA, como parte dos requisitos para a construção do ADSOD *Insecta*.

O Capítulo 4 mostra como ontologias de domínio podem apoiar as principais etapas da definição de bases de dados, apresentando inicialmente, na forma de um procedimento, uma das principais atividades do processo definido no capítulo 3, o projeto conceitual. Mostra também, um mapeamento entre os construtores ontológicos e da modelagem ER.

O capítulo 5 apresenta, inicialmente, as principais características da Estação TABA e de Ambientes de Desenvolvimento de Software Orientados a Domínio (ADSOD). Apresenta em seguida, a definição da Teoria do Domínio para Entomologia, tendo ontologias como base, focalizando cada uma das atividades desenvolvidas e sua utilização dentro de um ADSOD.

Apresentamos também, *Insecta*, um ADSOD para o domínio de Entomologia, definido e instanciado através da Estação TABA, sendo descritos todos os aspectos necessários para a sua construção, inclusive a inserção do processo definido no capítulo 3 e da Teoria do Domínio construída.

Também é mostrado um exemplo da execução da atividade Projeto Conceitual com base na Teoria do Domínio de Entomologia para uma base de dados de pesquisa neste domínio.

No capítulo 6, são apresentadas as conclusões e contribuições deste trabalho, bem como as perspectivas para futuras pesquisas.

O Anexo 1, apresenta a descrição de conceitos da Teoria do Domínio de Entomologia, e que se constitui de dois dos principais documentos gerados na construção de uma Teoria do Domínio: o dicionário de conceitos e o dicionário de propriedades de instâncias.

O Anexo 2 apresenta o questionário utilizado para efetivar a avaliação da Teoria do Domínio de Entomologia.

E, finalizando, o Anexo 3 apresenta um resumo da notação de LINGO, a LINGuagem Gráfica para descrever Ontologias que é utilizada no contexto da Estação TABA.

Capítulo 2

Ontologia e Modelagem de Dados

Este capítulo trata de ontologias abordando a definição e classificação, adotadas neste trabalho; os seus componentes e sua engenharia; bem como a utilização e reutilização de ontologias de domínio. Também é apresentada, a modelagem de dados, tendo como base, as etapas de construção de bases de dados e a utilização da abordagem Entidade-Relacionamento (ER) na sua execução; mostramos os objetivos desta modelagem e quais os benefícios de sua utilização. Com base nestas duas visões conceituais, é mostrada a relação entre ontologias de domínio e modelagem de dados, através de alguns trabalhos identificados na literatura.

2.1 – Ontologias

A aquisição e representação de conhecimento têm sido alvo de muitas pesquisas nos últimos anos (GUARINO, 1997; FALBO, 1998; STAAB e MAEDCHE, 2000; KNUBLAUCH e ROSE, 2000), devido tratar-se de uma das atividades mais importantes, mais trabalhosas, mais demoradas e de custo mais alto em qualquer contexto onde seja necessária sua utilização. Para modelar e tornar explícito o conhecimento, é necessário utilizar formas coerentes de organização do conhecimento disponível. Somente após esta organização será possível a utilização do conhecimento no desenvolvimento de software.

Visando tornar explícito o conhecimento e utilizá-lo no desenvolvimento de software, pesquisadores da área de engenharia de software e especificamente do Projeto TABA, buscaram uma abordagem de representação do conhecimento, cujas raízes são da filosofia, e que se chama Ontologia (OLIVEIRA, 1999).

As pesquisas em ontologia estão tornando-se cada vez mais difundidas na comunidade de ciência da computação (GUARINO, 1998; MILTON e KAZMIERCZAK, 1999; SMITH, 2000). O termo ontologia tem origem na filosofia e vem obtendo papel específico na Inteligência Artificial, Linguística Computacional, e Teoria de Bases de Dados.

A importância do uso de ontologias vem sendo reconhecida em diversas áreas de pesquisa tais como: engenharia e representação do conhecimento, modelagem qualitativa, engenharia da linguagem, projeto de bases de dados, modelagem e integração de informação, análise orientada a objeto, recuperação e extração de informação, gerência e organização do conhecimento, projeto de sistemas baseado em agentes, e modelagem de domínio (FOURO e WERNER, 1999; FOURO e WERNER, 2000; SMITH, 2000). Também, vem sendo utilizada em diferentes áreas de aplicação, como por exemplo, integração de empresas, medicina, comércio eletrônico e sistemas de informação biológica (GUARINO, 1998).

2.1.1 – Definição e Classificação

Existem várias visões e definições de ontologia na literatura, havendo atualmente, um grande interesse na criação, divulgação e compartilhamento de ontologias sobre os mais variados domínios (USCHOLD e GRUNINGER, 1996; ROSENBERG, 1997; GUARINO, 1998; CHANDRASEKARAN *et al.*, 1999; GOMÉS-PÉREZ e BENJAMINS, 1999; LÓPEZ *et al.*, 1999; OLIVEIRA, 1999; VALENTE *et al.*, 1999; SILVA, 2000; KNUBLAUCH e ROSE, 2000).

Qualquer que seja a sua definição, a construção de uma ontologia sempre resultará num vocabulário, que representará o entendimento consensual de um grupo de pessoas que atuam sobre um mesmo domínio. A definição adotada no presente trabalho é a mesma adotada por OLIVEIRA (1999), que tomou como base, a definição de FIKES e FARQUHAR (1999), onde, ontologia é uma teoria sobre um domínio que especifica um vocabulário de entidades, classes, propriedades, predicados e funções; e um conjunto de relações que necessariamente interliga esse vocabulário.

Apesar de adotar essa definição, OLIVEIRA (1999) considera ainda, que uma ontologia é uma descrição parcial projetada para ser compartilhada dentro de uma comunidade que concorda com a sua definição, mas com o objetivo específico de desenvolvimento de software em um dado domínio.

Nesta definição, os conceitos que compõem a ontologia, são definidos e representados claramente, de forma que possam ser compartilhados, e possuem restrições expressas através de axiomas que, explicitamente, restringem os significados dos termos.

As relações entre conceitos são representadas por axiomas, que podem ser classificados em três tipos (FALBO, 1998): (i) axiomas epistemológicos, que são derivados simplesmente da estrutura dos conceitos e não de seus significados particulares; (ii) axiomas de consolidação, que têm por objetivo verificar a coerência das informações existentes e não derivam novas informações e (iii) axiomas ontológicos, que representam restrições entre os conceitos que compõem a ontologia.

Da mesma forma que a definição, existem diferentes modos de se classificar uma ontologia (USCHOLD e GRUNINGER, 1996; GUARINO, 1998). Neste trabalho, será utilizada a classificação de GUARINO (1998), já que ela é uma classificação bastante aceita, e apresenta de maneira explícita, uma separação clara entre ontologia de domínio e de tarefa conforme a figura 2.1.

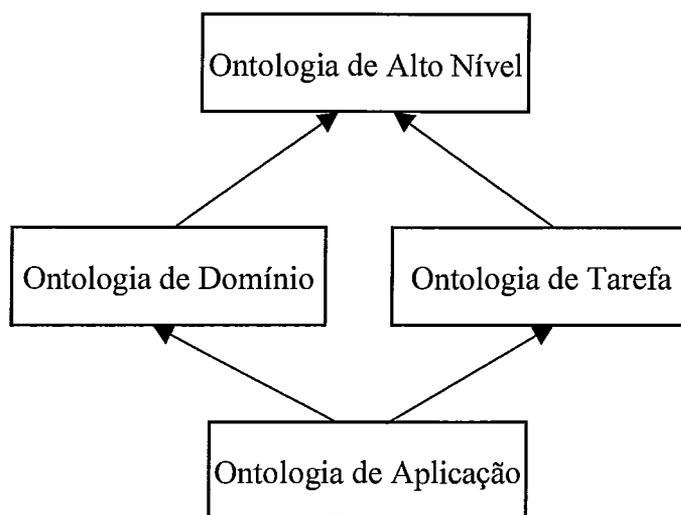


Figura 2.1 – Classificação de ontologias segundo GUARINO (1998)

Esta classificação baseia-se no conteúdo das ontologias e propõe que ontologias sejam construídas segundo seu nível de generalidade. As ontologias de alto nível são aquelas que descrevem conceitos bastante gerais como espaço, tempo, matéria, objeto, evento, ação, que são independentes de um problema ou domínio particular.

As ontologias de domínio e as ontologias de tarefa descrevem, respectivamente, o vocabulário relacionado a um domínio genérico como medicina, ou uma tarefa ou atividade genérica como, diagnose ou venda. Elas especializam os termos introduzidos na ontologia de alto nível.

As ontologias de aplicação descrevem conceitos dependentes do domínio e da tarefa particulares e são especializações de ambas as ontologias, de domínio e de tarefa. Estes conceitos freqüentemente correspondem a papéis desempenhados por entidades do domínio quando da realização de uma certa atividade.

GUARINO (1998) diferencia ainda, ontologia de aplicação de base de conhecimento, considerando o propósito da ontologia. A ontologia de aplicação é uma base de conhecimento particular, que descreve os fatos assumidos como sendo sempre verdadeiros para uma comunidade de usuários que concordam com o significado do vocabulário usado. Já a base de conhecimento genérica, é composta pela ontologia, que contém as informações independentes de estado; e pelo “núcleo” da base de conhecimento, que contém as informações dependentes de estado.

2.1.2 – Componentes e Engenharia de Ontologias

Existem diversos métodos para a construção de ontologias (USCHOLD e GRUNINGER, 1996; FALBO, 1998; CRANFIELD e PURVIS, 1999; GOMÉZ-PÉREZ e BENJAMINS, 1999; KNUBLAUCH e ROSE, 2000). No método de construção proposto por FALBO (1998), por exemplo, é ressaltado que o processo de desenvolvimento da ontologia deve ser visto como um processo fortemente iterativo, e não como passos seqüenciais. Este processo está evidenciado na figura 2.2.

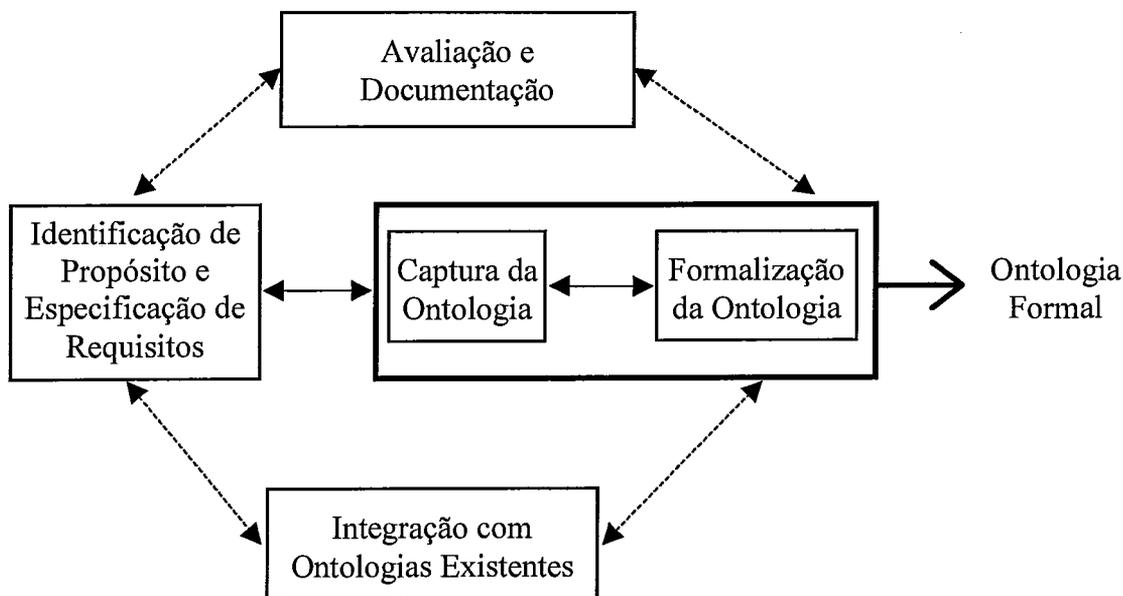


Figura 2.2 – Engenharia de ontologias (FALBO, 1998)

De maneira geral, qualquer que seja o domínio, a complexidade envolvida na construção de ontologias é grande e, portanto, algum mecanismo de decomposição deve ser usado para facilitar o processo de construção (FALBO, 1998).

Basicamente, o processo de construção de ontologias segue as etapas de: identificação do propósito e especificação dos requisitos da ontologia; captura, formalização e integração com ontologias existentes; avaliação e documentação. Isto implica que os componentes básicos da ontologia são: uma especificação de requisitos; uma rede semântica ou árvore de conceitos; uma base de conhecimento contendo o significado dos conceitos e os axiomas que restringem a interpretação dos conceitos (FALBO, 1998).

Segundo OLIVEIRA (1999), semelhante ao desenvolvimento de sistemas, a construção de ontologias deve considerar critérios específicos e um processo de avaliação que envolva verificação e validação. A Tabela 2.1 apresenta alguns dos critérios mais importantes citados na literatura para o projeto e avaliação de ontologias.

Tabela 2.1 – Critérios para o projeto e avaliação de ontologias (OLIVEIRA, 1999)

Critério	Descrição
Clareza	Comunicar efetivamente o significado planejado dos termos definidos (definições objetivas, documentadas em linguagem natural e independentes do contexto social ou computacional) através de formalismos.
Consistência	Garantir consistência com a sua definição, tanto no que se refere à especificação dos axiomas lógicos quanto aos conceitos informais, como os comentários em linguagem natural. Uma ontologia é, semanticamente, consistente se e somente se suas definições são semanticamente consistentes e uma dada definição é semanticamente consistente, se e somente se (i) os significados tanto das definições formais quanto informais são consistentes com o mundo real e consistentes umas com as outras e, (ii) se as definições não são sentenças contraditórias que podem ser inferidas usando outras definições e axiomas que pertençam ou não à mesma ontologia.
Extensibilidade	Ser extensível para englobar novos conceitos com base no vocabulário existente. Garantir a capacidade de extensão refere-se ao esforço necessário para se incluir novas definições sem alterar o conjunto de propriedades já avaliadas na ontologia.
Mínimo “bias” de codificação	Uma conceituação deve ser especificada sem a dependência de uma codificação simbólica específica, procurando não escolher uma representação somente por causa de uma determinada implementação ou codificação.
Conjunto mínimo de compromissos ontológicos	Uma ontologia deve ser suficiente para apoiar as atividades pretendidas de compartilhamento do conhecimento. Isto é conseguido através de uma especificação de uma teoria, definição somente de termos que são essenciais para comunicação do conhecimento consistente com esta teoria, e definição do mínimo de restrições possíveis sobre o mundo que está sendo modelado.
Generalidade	Ser capaz de ser compartilhada entre diferentes atividades como projeto e análise.
Suporte	Apoiar o desenvolvimento de grandes aplicações.
Completude	Garantir a completude (dependendo do nível de granularidade da ontologia) de uma definição implica em garantir a completude tanto da definição formal quanto da informal. Para garantir que uma definição formal é completa, deve-se determinar se a definição está de acordo com seu critério estrutural para uma definição completa, se o domínio e um grupo de relações e funções estão precisamente e exatamente bem delimitados; se a generalização/especialização de uma determinada classe representa exata e precisamente uma determinada classe do mundo real; e, finalmente, verificar a existência de um conjunto completo de atributos em cada definição. Uma definição informal escrita em linguagem natural é completa se e somente se expressa o mesmo conhecimento pretendido em uma definição formal.
Concisão	Capturar apenas a informação necessária com definições concisas. Uma definição em uma ontologia é concisa se são evitadas redundâncias formais e informais.
Robustez	Ser robusta de forma que pequenas mudanças não afetem o conjunto de definições já avaliadas.

Para se utilizar efetivamente uma ontologia, é necessário que ela seja expressa através de uma linguagem de representação do conhecimento formal, ou mesmo informal. Desta forma, um conjunto de linguagens tem sido utilizado para implementar ontologias (OLIVEIRA, 1999; SILVA, 2000).

CORCHO e GÓMEZ-PÉREZ (2000) fazem um levantamento geral das linguagens mais representativas e em uso. A Ontolinguagem (FARQUHAR *et al.*, 1996) é a mais representativa de todas elas, e é considerada como um padrão pelos pesquisadores da área. Mas, outras linguagens também têm sido utilizadas para a especificação de ontologias: lógica de primeira ordem (VALENTE, 1995); KIF (Knowledge Interchange Format) (GENESERETH e FIKES, 1992); FLogic (KIFER *et al.*, 1995); LINGO (Linguagem Gráfica para Ontologias) (FALBO, 1998) e OCML (MOTTA, 1999).

Novas linguagens de especificação de ontologias baseadas nas linguagens padrão da Web - XML (BRAY *et al.*, 2000) e RDF (LASSILA e SWICK, 1999) - têm surgido mais recentemente. Tem-se como exemplo: SHOE (LUKE e HEFLIN, 2000), XOL (KARP *et al.*, 1999), OIL (FENSEL *et al.*, 2000), o esquema RDF (BRICKLEY e GUHA, 2000) e o esquema XML (THOMPSON *et al.*, 2001).

O papel de novas linguagens neste cenário, de constante mudança de contexto, é duplo: podem ser utilizadas para prover as semânticas da informação contida nos documentos eletrônicos, como também, podem ser usadas para o intercâmbio de ontologias através da Web (CORCHO e GÓMEZ-PÉREZ, 2000).

2.1.3 – Utilização e Reutilização de Ontologias de Domínio

Uma ontologia pode ser utilizada com muitos objetivos e pode proporcionar muitos benefícios (VALENTE *et al.*, 1999; LÓPEZ *et al.*, 1999; MELLO e HEUSER, 2000). Um dos principais objetivos é tornar viável o entendimento comum, bem como possibilitar a interoperabilidade de sistemas.

A utilização de ontologias tem um forte impacto na aquisição do conhecimento, pois o conhecimento geral do domínio é capturado, organizado e especificado na forma de um modelo de conhecimento (FALBO, 1998). Além disso, ontologias possibilitam o compartilhamento (CHANDRASEKARAN *et al.*, 1999) e o reuso do conhecimento por diferentes aplicações (USCHOLD *et al.*, 1998).

Ontologias podem ser utilizadas como esquemas de objetos para sistemas OO, esquemas conceituais para bases de dados, glossários estruturados para a colaboração entre pessoas e para construção de bases de conhecimento (FIKES e FAGUHAR, 1999).

Do ponto de vista da reutilização de ontologias, existe a possibilidade de incorporar ontologias já existentes na construção de uma nova aplicação. Mas, pode ter ainda outros significados para o reuso, tais como os apresentados em PINTO *et al.* (1999): *integração e fusão* de ontologias, que têm como produto uma outra ontologia.

O desenvolvimento de ontologias em diferentes domínios e com diferentes propósitos (LÓPEZ *et al.*, 1999; VALENTE *et al.*, 1999; OLIVEIRA *et al.*, 1999b; GALLOTA *et al.*, 1999), tem contribuído para a constatação de que o uso de uma ontologia em determinada área melhora a comunicação e, conseqüentemente, proporciona maior possibilidade de reutilização, interoperabilidade e produção de sistemas de software de maior confiabilidade.

OLIVEIRA (1999) considera que a existência de uma ontologia e, conseqüentemente, de uma Teoria do Domínio, pode, principalmente em engenharia de software, apoiar o desenvolvedor de software no entendimento do problema e ao longo do processo de desenvolvimento.

Sob o ponto de vista do desenvolvimento de sistemas de informação, cada componente pode utilizar uma ontologia de uma forma própria e específica. Mas, o uso mais óbvio de uma ontologia é na conexão com o componente base de dados (GUARINO, 1998).

USCHOLD e GRUNINGER (1996), descrevem algumas vantagens do uso de ontologias no desenvolvimento de software, tais como:

- permitir um entendimento comum e a efetiva comunicação entre pessoas com diferentes necessidades e pontos de vista particulares em um determinado contexto;
- servir como uma linguagem intermediária que apóia a tradução entre diferentes linguagens e representações;
- apoiar tarefas específicas, através da integração de ontologias em diferentes domínios ou mesmo diferentes ontologias dentro de um mesmo domínio;
- apoiar a especificação e reutilização durante a engenharia de software.

2.2 – Modelagem de Dados

Nesta seção será abordada a criação de modelos de dados e sua evolução/transformação ao longo do processo geral de construção de bases de dados.

O objetivo da modelagem de dados é desenvolver um modelo ou uma representação gráfica das informações necessárias aos usuários em uma organização, que possa ser utilizado em um conjunto de aplicações definidas (McLELLAN, 1995). Outro objetivo identificado é possibilitar a apresentação de uma visão única, não redundante e resumida dos dados de uma aplicação, durante o projeto de bases de dados (HEUSER, 2001).

A modelagem de dados ou modelagem semântica está sempre associada à atividade de projeto de banco de dados (COUGO, 1997), uma vez que, provê técnicas para a identificação, análise, especificação e documentação dos requisitos para uma base de dados, permitindo ainda, envolver o usuário na especificação do banco de dados.

Originalmente, a modelagem de dados, centrada somente sob o enfoque da abordagem Entidade-Relacionamento proposta por Peter Chen (CHEN, 1976), cuja base é o modelo Entidade-Relacionamento (MER), introduziu uma técnica de diagramação conhecida como Diagrama Entidade-Relacionamento (DER) (DATE, 2000).

Durante sua evolução ao longo dos anos, têm sido agregados novos elementos ao MER, tornando a técnica mais rica em semântica e, portanto, aplicável a novas finalidades. Atualmente, é considerado um padrão para a modelagem de dados (HEUSER, 2001).

A abordagem Entidade-Relacionamento defende, basicamente, a elaboração de um modelo que represente as entidades observadas e seus relacionamentos, independentemente de preocupações com a implementação (MACHADO e ABREU, 1996).

O processo de construção de uma base de dados é composto de três grandes etapas (KOLP e ZIMÁNYI, 2000), que constituem na verdade níveis de abstrações diferentes dos dados que devem atender à especificação de requisitos da base de dados. A figura 2.3 mostra estas etapas: Projeto Conceitual, Projeto Lógico e Projeto Físico.

A partir da observação dos requisitos de dados, o analista organiza as idéias sobre o domínio de interesse, definindo o Projeto Conceitual, que por sua vez descreve o domínio de interesse. Cada etapa gera como produto um modelo ou esquema, que constitui uma evolução do modelo da etapa anterior; além de estar associada a uma técnica de representação gráfica e métodos de especificação de esquemas. A seguir, detalharemos cada uma destas etapas (MACHADO e ABREU, 1996; COUGO, 1997).

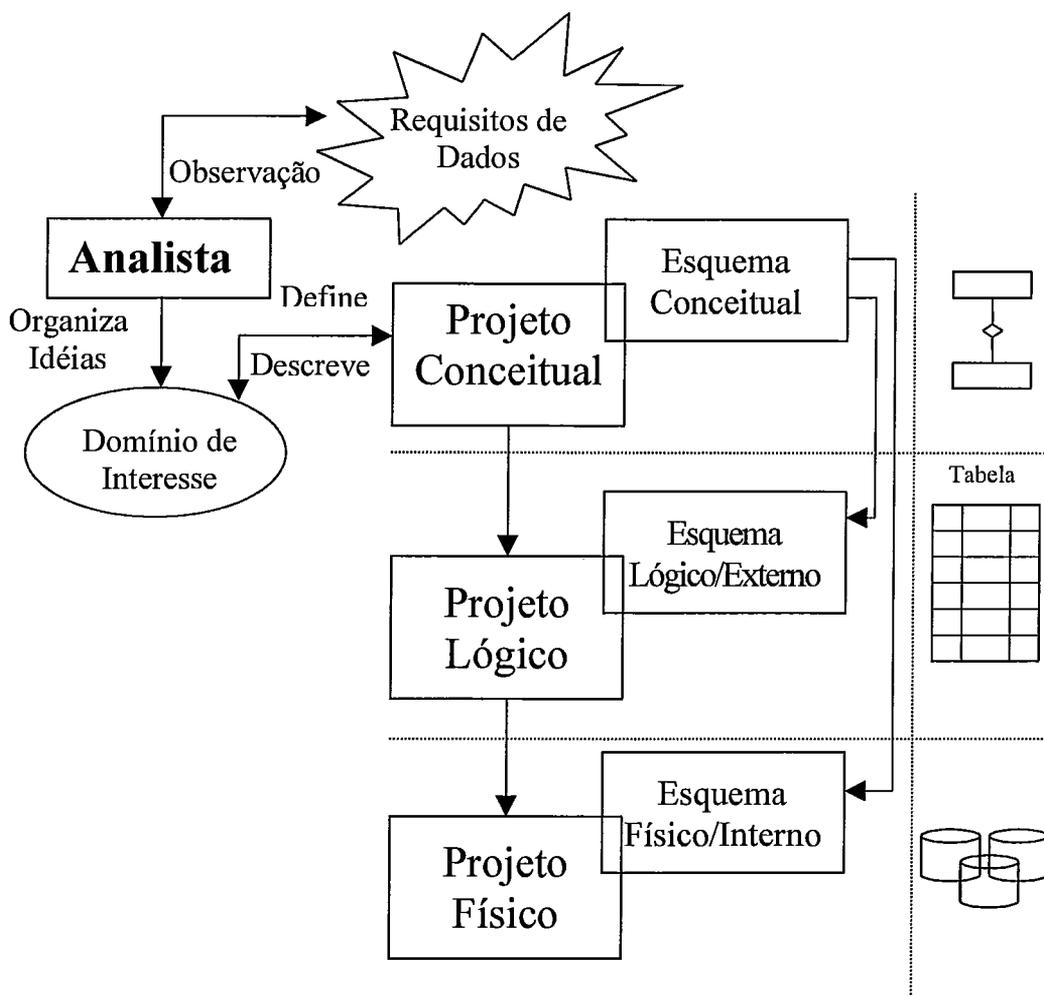


Figura 2.3 – Etapas (níveis de abstração) na construção de bases de dados (MACHADO e ABREU, 1996; COUGO, 1997)

2.2.1 – Projeto Conceitual

O projeto conceitual é a primeira etapa do projeto de uma base de dados (MACHADO e ABREU, 1996; COUGO, 1997; HEUSER, 2001), onde é construído o MER. Este modelo captura as necessidades da organização em termos de armazenamento de dados de forma independente de implementação.

O modelo gerado nesta etapa (MER) é aquele em que as entidades, suas características e relacionamentos têm a representação o mais fiel possível ao ambiente observado, incorporando algumas das informações semânticas importantes sobre o mundo real (CHEN, 1976). Nesse modelo, são representados os conceitos e características observadas em um dado ambiente, concentrando-se apenas no aspecto conceitual.

No desenvolvimento de aplicações envolvendo banco de dados, o MER é o modelo mais amplamente utilizado para a representação e entendimento dos dados (DATE, 2000). Deve ser aplicado, por exemplo, para conversação, entendimento, transmissão, validação de conceitos e mapeamento do ambiente.

Os componentes de um MER na proposta original de Peter Chen são: Entidades, Relacionamentos e Atributos (COUGO, 1997; KOLP e ZIMÁNYI, 2000; HEUSER, 2001).

- Entidade: representa um conjunto de conceitos do domínio com existência independente.
- Relacionamento: é o fato, o acontecimento que liga duas ou mais entidades existentes no domínio.
- Atributos: são as características próprias, inerentes a cada um dos conceitos (entidades e relacionamentos).

Através da proposta inicial de Chen (CHEN, 1976), qualquer modelo poderia ser construído com um mínimo suficiente de semântica baseado nos conceitos de entidades, atributos e relacionamentos. Mas, com o passar do tempo, novas características semânticas demonstraram serem necessárias, para o completo entendimento dos modelos.

A partir dessas novas necessidades, alguns elementos adicionais foram agregados ao MER e constituem extensões do modelo (MACHADO e ABREU, 1996; COUGO, 1997; HEUSER, 2001). Tem-se então, as seguintes estruturas:

- Generalização/Especialização: é a estrutura que procura representar o fato de que, dado um conjunto de elementos pertencentes a subconjuntos de um conjunto maior, cada um deles deve poder ser visto como um elemento tanto pertencente aos subconjuntos distintos como ao conjunto completo.
- Agregação: é uma estrutura que denota a existência de uma junção de elementos através de um relacionamento, e que permite que essa junção seja percebida como um novo elemento a ser, por sua vez, relacionado a outro elemento.

Estes componentes reunidos de forma gráfica constituem o DER, que é uma ferramenta gráfica para projetistas e usuários de bases de dados, na descrição e comunicação de seu entendimento do mundo (DATE, 2000).

2.2.2 – Projeto Lógico

O projeto lógico, consta da transformação de um MER em um modelo lógico, cujo conteúdo é a descrição da estrutura da base que pode ser processada pelo Sistema Gerenciador de Banco de Dados (SGBD), sendo dependente do tipo particular de SGBD utilizado. O modelo lógico implementa ao nível do SGBD, os dados representados abstratamente no MER (HEUSER, 2001). Para expressar o projeto lógico, o modelo relacional é o mais comum e mais utilizado. Composto de tabelas que formam a base de dados, colunas para cada tabela, chaves de acesso e restrições de integridade, dentre outros (COUGO, 1997).

2.2.3 – Projeto Físico

O projeto físico tem como base o modelo lógico e resulta no modelo físico, que é uma descrição da implementação da base em memória secundária, descrevendo as estruturas de armazenamento e métodos de acesso utilizados para acesso efetivo aos dados, onde é detalhado o armazenamento interno das informações, que podem influenciar a performance das aplicações (MACHADO e ABREU, 1996).

O modelo físico é especificado para um SGBD específico. Cada SGBD define um modo diferente de implementação física das características e recursos necessários para o armazenamento e manipulação das estruturas de dados (estruturas de armazenamento, endereçamento, acesso e alocação física), sendo inclusive, específicos para cada sistema operacional (COUGO, 1997).

As decisões tomadas durante esta etapa, para melhorar o desempenho, podem afetar a estrutura do modelo lógico. Uma vez que os modelos lógico e físico estejam prontos, podem ser definidas e criadas as tabelas propriamente ditas.

2.3 – Usos e Benefícios da Modelagem de Dados

Um modelo de dados é um plano para a construção de bases de dados. Para ser efetivo, ele deve ser simples, preciso, conciso, objetivo e sem ambigüidades. Vários autores destacam a utilização e o benefício da modelagem de dados (McLELLAN, 1995; REYNAUD e TORT, 1996; CHEN *et al.*, 1998; CHEN, 1998; GREGERSEN e JENSEN, 1999; HEUSER, 2001).

Os principais usos e benefícios da modelagem de dados são:

- a preservação do conhecimento, facilitando o entendimento e a transmissão de conceitos, especificações e regras do negócio das organizações;
- a organização e estruturação do conhecimento, permitindo a transmissão de conceitos, especificações e regras do domínio;
- o auxílio ao melhor entendimento de um domínio específico do mundo real e melhoria na comunicação entre as pessoas;
- a descrição da estrutura do domínio de interesse, determinando as entidades que podem existir no domínio, bem como a descrição dos requisitos de dados dos especialistas da área de negócios;
- a revisão e verificação, pelos especialistas da área de negócios, dos modelos criados pelos projetistas;
- a análise, pelos especialistas do negócio, do impacto de mudanças nos requisitos do negócio;
- a segurança de que todos os dados requeridos por uma base de dados, estão representados completa e exatamente;
- o apoio ao desenvolvimento de sistemas de software com mais qualidade, através da representação dos dados, da forma como os usuários finais os vêem no mundo real;
- a interligação dos conceitos que constituem os eventos e os processos do mundo real, com a representação física destes conceitos na base de dados;
- a comunicação aos usuários finais, da estrutura dos dados requerida pela base de dados e que irá gerar a estrutura física.

Um modelo de dados efetivo representa completa e exatamente os requisitos de dados dos usuários finais. É simples o suficiente para ser entendido pelo usuário final e é detalhado o suficiente para ser utilizado pelo projetista de base de dados. O modelo elimina a redundância de dados, é independente de quaisquer restrições de hardware e software, e pode ser adaptado às mudanças de requisitos com um mínimo de esforço.

2.4 – Relação entre Ontologias de Domínio e Modelos de Dados

Existem vários trabalhos na literatura que consideram a relação entre ontologias e modelos de dados. A seguir descrevemos alguns dos trabalhos desenvolvidos observando este enfoque.

No trabalho de DOWELL *et al.* (1995), uma ontologia do conhecimento do domínio pode ser utilizada como uma entrada semântica entre diferentes esquemas de bases de dados. Especificamente, utiliza ontologias de domínio para gerar o modelo conceitual –

o modelo ER estendido, através do mapeamento de vários esquemas lógicos de diferentes bases de dados. Os rótulos das entidades, relacionamentos e atributos no MER estendido são então ligados aos itens correspondentes na ontologia do domínio. A ontologia do domínio consiste de uma hierarquia de tipos (correspondente às entidades) e uma hierarquia de ligações (correspondente aos relacionamentos e atributos).

WAND (1996) propõe um modelo baseado em princípios ontológicos que deverá servir de base teórica para a atividade de modelagem em geral, utilizando ontologia como um modelo generalizado da realidade, adaptando princípios ontológicos fundamentais que são independentes do domínio e que pode ser vista como uma meta-ontologia.

O modelo proposto chama-se *modelo de representação*, que pode ser aplicado para proporcionar um significado bem definido para construtores usados em métodos de modelagem. Como uma das aplicações do modelo proposto, faz a análise de modelos de dados semânticos com base no mapeamento definido na tabela 2.2. O modelo também prevê regras para quando usar entidades, relacionamentos e atributos, para modelar agregação e para evitar ambigüidades e inconsistências nos modelos Entidade-Relacionamento.

Tabela 2.2 – Mapeamento entre conceitos ontológicos e modelo de dados (WAND, 1996)

Conceito Ontológico	Construtor de Modelagem Semântica
Conceito	Entidade
Propriedade/atributo	Atributo
Esquema funcional	Entidade tipo
Interação/Propriedade mútua	Relacionamento
Composição	Agregação

GUARINO (1998) afirma que o uso mais óbvio de uma ontologia, no contexto de componentes de sistemas de informação, é na conexão com bases de dados e que uma ontologia pode ser comparada com o modelo conceitual de uma base de dados e pode ainda, ser utilizada na fase de modelagem conceitual.

WAND *et al.* (1999) analisa o significado de construtores comuns da modelagem conceitual provendo uma definição precisa de vários desses construtores e derivando regras para o uso de relacionamentos na modelagem conceitual Entidade-Relacionamento. Mostra ainda, como as regras resolvem ambigüidades que existem na prática atual e como elas podem aumentar a capacidade de um modelo conceitual ER para capturar o conhecimento sobre um domínio de aplicação.

Finalizando, MILTON e KAZMIERCZAK (1999), consideram o relacionamento entre ontologia e modelo de dados, argumentando que a ontologia de Roderick Chisholm é, potencialmente, uma teoria de unificação dos modelos de dados. Além disso, também estabelecem, que ontologias fundamentadas na tradição filosófica de realismo, servem ao propósito de serem uma estrutura de unificação dos modelos de dados, enriquecendo os modelos de dados e nosso entendimento sobre os mesmos.

2.5 – Considerações Finais

A modelagem conceitual é uma atividade que envolve um processo de abstração que se constitui em uma ferramenta intelectual para exploração de regularidade na organização do conhecimento (YEROOS, 2001). Tanto ontologias quanto modelos de dados são ferramentas para a execução da modelagem conceitual.

Ontologias são consideradas uma ferramenta útil para apoiar a especificação e implementação de qualquer sistema de computação complexo. Seu papel é a captura do conhecimento do domínio provendo um acordo comum através de um vocabulário comum, que define o significado dos termos no domínio e suas relações (STAAB e MAEDCHE, 2000).

Com o desenvolvimento de ontologias objetiva-se satisfazer a um ou mais dos seguintes propósitos (SILVA, 2000):

- possibilitar uma melhor compreensão de uma certa área do conhecimento;
- permitir que um consenso no entendimento sobre uma área de conhecimento seja atingido;
- estruturar e organizar o conhecimento visando possibilitar a sua reutilização;
- apoiar a implementação prática de sistemas.

A modelagem de dados tem por objetivo desenvolver um modelo detalhado o suficiente para ser utilizado pelos projetistas de bases de dados.

BENCH-CAPON (1997), argumenta que os pesquisadores de banco de dados e de sistemas especialistas devem compartilhar seus conhecimentos e experiências, devido à existência de uma grande semelhança no que a ontologia é para sistemas especialistas e o que o modelo de dados é para o projeto de bases de dados.

Já para MELLO e HEUSER (2000), do ponto de vista de banco de dados, uma ontologia pode ser considerada uma especificação parcial de um domínio que descreve basicamente conceitos, relações entre conceitos e regras de integridade. Porém, uma ontologia não é estritamente um modelo conceitual de dados, já que o mesmo descreve a estrutura de um banco de dados em alto nível de abstração. Enquanto que ontologias apenas propõem uma estrutura de consenso para grupos de usuários, sendo essa estrutura instanciada através de mecanismos de integração de dados.

Desta forma, um dos objetivos deste trabalho é investigar a complementaridade entre esses modelos, definindo um mapeamento que possa auxiliar desenvolvedores de software na modelagem de dados a partir de ontologias de domínio.

Este mapeamento será apresentado no capítulo 4, assim como um procedimento detalhado de como realizar a modelagem conceitual de dados apoiando-se em uma ontologia de domínio. Esse procedimento é definido e integrado ao processo para construção de base de dados de pesquisa apresentado no capítulo 3, que será utilizado para instanciar um ADSOD específico na Estação TABA.

Capítulo 3

Processo de Construção de Bases de Dados de Pesquisa

Este capítulo apresenta um processo padrão para construção de bases de dados de pesquisa, que pode ser utilizado em instituições de ciência e tecnologia, onde seja necessário tornar disponível rapidamente bases de dados de pesquisa que apoiem as atividades de pequenos grupos de pesquisadores.

3.1 – Processo de Software

A inexistência de um processo de software disciplinado permite que o desenvolvimento de software, ainda hoje, seja feito fortemente baseado em habilidades individuais (FALBO, 1998), fazendo que sejam freqüentes os insucessos nos projetos de desenvolvimento de software (MACHADO, 2000).

Visando aperfeiçoar o desenvolvimento de software, e na busca por um meio mais racional de produção, deve-se considerar a tarefa global de desenvolvimento como um processo que pode ser controlado, medido e melhorado, contribuindo para a diminuição dos freqüentes insucessos, além de ser a base da garantia da qualidade dos produtos gerados e para a melhoria da produtividade das equipes de desenvolvimento (FALBO, 1998; MAIDANTCHIK, 1999; MACHADO, 2000).

Um processo de software consiste de um conjunto de tarefas de engenharia de software, necessárias para a obtenção de um software a partir das transformações dos requisitos do usuário. É composto por: atividades, métodos, práticas e transformações, que guiam as pessoas na produção de software (PAULK *et al.*, 1997; FALBO, 1998; MACHADO, 2000).

Do ponto de vista da norma NBR ISO/IEC 12207 (1998) que constitui o padrão internacional para a definição de processos de ciclo de vida de software, o processo de desenvolvimento é um dos cinco processos fundamentais do ciclo de vida do software. A norma estabelece, que cada processo do ciclo de vida seja dividido em um conjunto de atividades. Cada atividade, por sua vez, é dividida em um conjunto de sub-atividades (tarefas).

A norma NBR ISO/IEC 12207 (1998) provê na verdade, uma estrutura comum que pode ser utilizada por profissionais de software, para a criação e gerência de software, abrangendo o ciclo de vida de software desde a concepção até a descontinuação. Esta estrutura é abrangente, de forma que uma organização possa, dependendo de qual seja o seu objetivo, selecionar o subconjunto de processos apropriado para satisfazer este objetivo.

Esta seleção é necessária, em função da impossibilidade de existência de um processo de software que possa ser genericamente aplicado a qualquer projeto, devido à existência de variações nas políticas e procedimentos organizacionais; métodos e estratégias de aquisição; tamanho e complexidade do projeto; requisitos e métodos de desenvolvimento do sistema, dentre outros fatores (MACHADO, 2000).

Todas essas variações contribuem para a existência de diferentes formas como um produto de software pode ser adquirido, desenvolvido, operado e mantido. Assim sendo, quando da definição de um processo de software eficaz, deve-se considerar as relações entre as atividades, as ferramentas e os procedimentos necessários, a sua adequação às tecnologias envolvidas, o tipo de software em questão, o domínio de aplicação, o grau de maturidade (ou capacitação) da equipe em engenharia de software, as características próprias da organização e as características do projeto (FALBO, 1998; ROCHA *et al.* 1999; MAIDANTCHIK, 1999; MACHADO, 2000).

Na busca pelo aperfeiçoamento e pelo estabelecimento de um processo de software disciplinado, os modelos de maturidade CMM (PAULK *et al.*, 1997) e SPICE (EMAM *et al.* 1998) estabelecem a definição de um processo padrão como uma das práticas importantes para se homogeneizar o desenvolvimento de software em projetos e organizações (MACHADO, 2000).

Um processo padrão estabelece a estrutura básica que guiará a definição de qualquer processo numa organização desenvolvedora de software, descrevendo os elementos fundamentais que cada projeto de software deve incorporar e as relações entre esses elementos (EMAM *et al.* 1998; MAIDANTCHIK, 1999; MACHADO, 2000).

A utilização de um processo padrão possibilita que o produto final seja de melhor qualidade, uma vez que a equipe de desenvolvimento realiza as atividades com o apoio de um processo formal, documentado, disciplinado e padronizado (MAIDANTCHIK, 1999). Além disso, quando se define um processo padrão, ocorre uma grande economia de tempo e esforço na definição de novos processos adequados a projetos, de forma que, ao se definir um processo padrão, deve-se buscar obter um processo facilmente adaptável e configurável (MACHADO, 2000).

A seguir apresentamos a definição de um processo padrão para construção de bases de dados de pesquisa adaptável e configurável a qualquer organização que realize este tipo de desenvolvimento.

3.2 – Processo de Desenvolvimento para Construção de Bases de Dados de Pesquisa

O processo de desenvolvimento para construção de bases de dados de pesquisa estabelece um conjunto de atividades definidas com base na estrutura comum estabelecida pela norma NBR ISO/IEC 12207 (1998) e no projeto de banco de dados utilizando a abordagem Entidade-Relacionamento (HEUSER, 2001; COUGO, 1997; MACHADO e ABREU, 1996; BATINI *et al.*, 1992). Este processo pode ser utilizado por qualquer organização que tenha como objetivo manter informações para prover apoio para pesquisa.

O ciclo de vida deste processo é constituído das seguintes fases: *Instituição do Projeto*, *Definição*, *Desenvolvimento*, *Transição* e *Manutenção*, mostradas na figura 3.1. A base de dados é construída em um único ciclo de desenvolvimento, até a sua completa operação, que ocorre na fase de Transição, caracterizando um ciclo de vida seqüencial. A fase de Manutenção pode implicar na ocorrência de um novo ciclo, que é iniciado somente se surgir um novo requisito que necessite de dados que não existam na base já construída. Se os dados necessários já existirem na base, são implementadas apenas as novas consultas ou é executada uma manutenção corretiva, não implicando na existência de novo ciclo.

Cada fase do processo é dividida em atividades, mostradas na figura 3.2, que por sua vez, são subdivididas em sub-atividades que serão descritas a seguir. Na figura 3.2 destacamos ainda o procedimento para modelagem conceitual que é utilizado na atividade Projeto Conceitual e que será apresentado no próximo capítulo.

Este processo tem como objetivo a construção de bases de dados, com vistas a assegurar aos pesquisadores que as utilizam, que os dados estruturados e armazenados refletem corretamente os resultados das pesquisas que ocorreram no mundo real. Não está sendo considerada a existência, prévia ou futura, de nenhum sistema de informação que irá utilizar os dados da base.

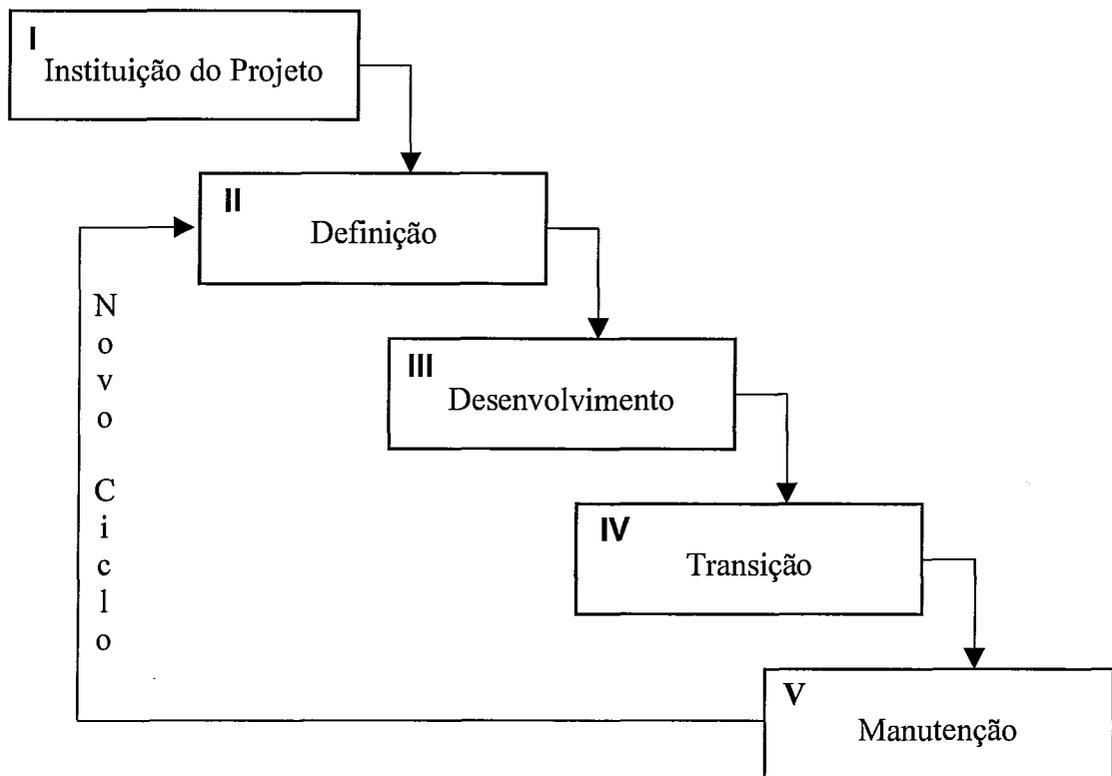


Figura 3.1 – Fases do Ciclo de Vida do processo para Construção de Bases de Dados de Pesquisa

Para a atividade de Análise e Especificação de Requisitos (na fase de Definição) sugerimos uma abordagem inspirada no GQM (*Goal-Question-Metrics*) (BASILI *et al.*, 1994). A abordagem GQM, representada na figura 3.3 visa apoiar medições de software a partir de objetivos e questões previamente estabelecidos. Assim sendo, são identificados os objetivos da medição e, a partir destes, são definidas questões. Para se ter respostas às questões, são identificadas métricas. Assim evita-se o medir por medir.

Em nossa abordagem, ilustrada pela figura 3.4, que denominamos OQD (Objetivos-Questões-Dados), visamos apoiar a construção de bases de dados de pesquisa a partir do estabelecimento dos objetivos da pesquisa que devem ser claros e explicitamente definidos. Estes podem ser objetivos a curto, médio e longo prazos. A partir destes objetivos são, então, identificadas, as questões que se deseja responder com a pesquisa e que darão origem às consultas na base de dados. Para responder às questões são necessários dados. A partir das questões da pesquisa são, portanto, identificados os dados que deverão estar na base para respondê-las².

² Esta abordagem já foi usada, com sucesso, em vários projetos de construção de bases de dados para pesquisa em Cardiologia, na Fundação Bahiana de Cardiologia/Unidade de Cardiologia e Cirurgia Cardiovascular/ Universidade Federal da Bahia.

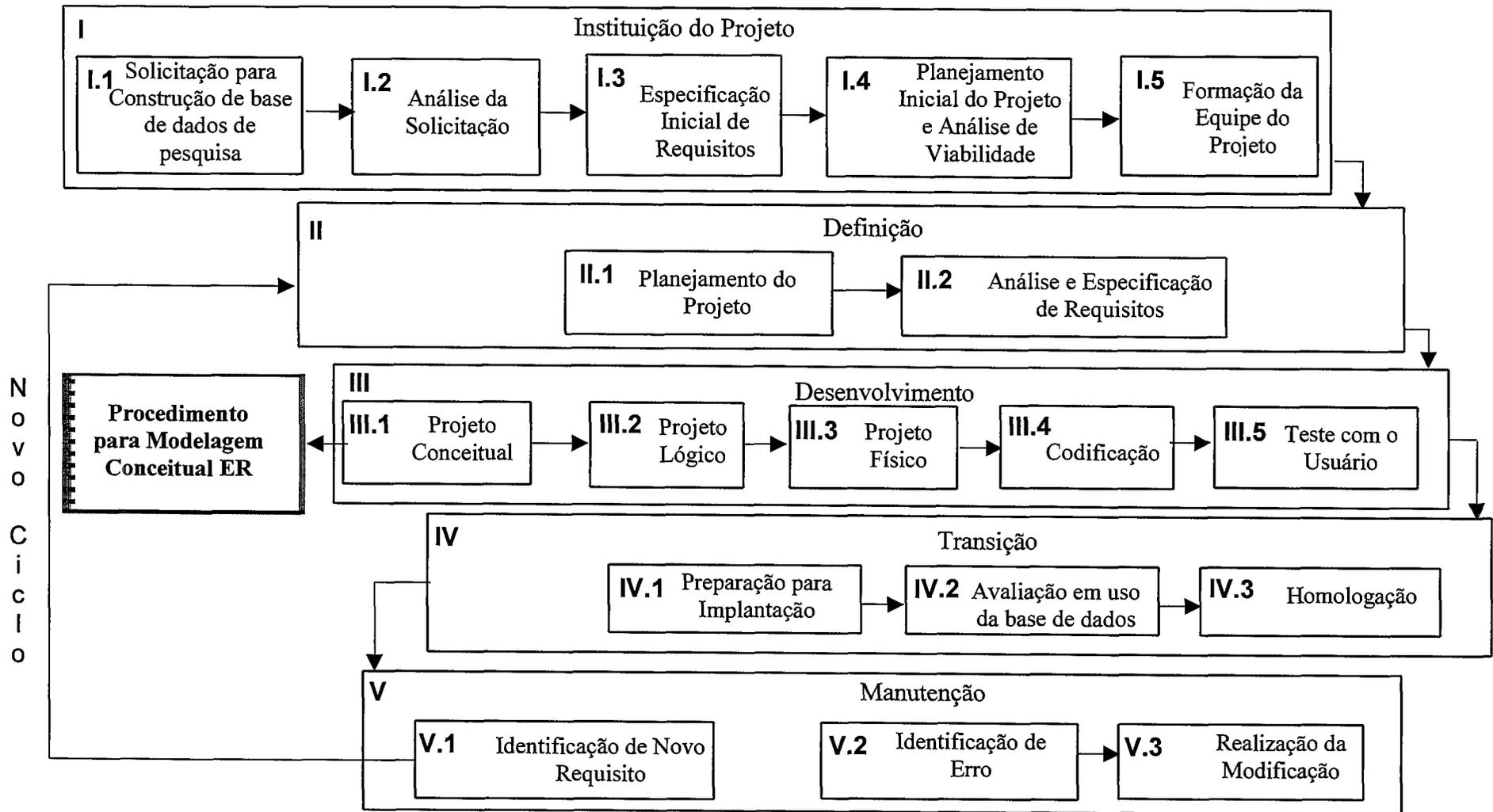


Figura 3.2 – Atividades do Processo para Construção de Bases de Dados de Pesquisa

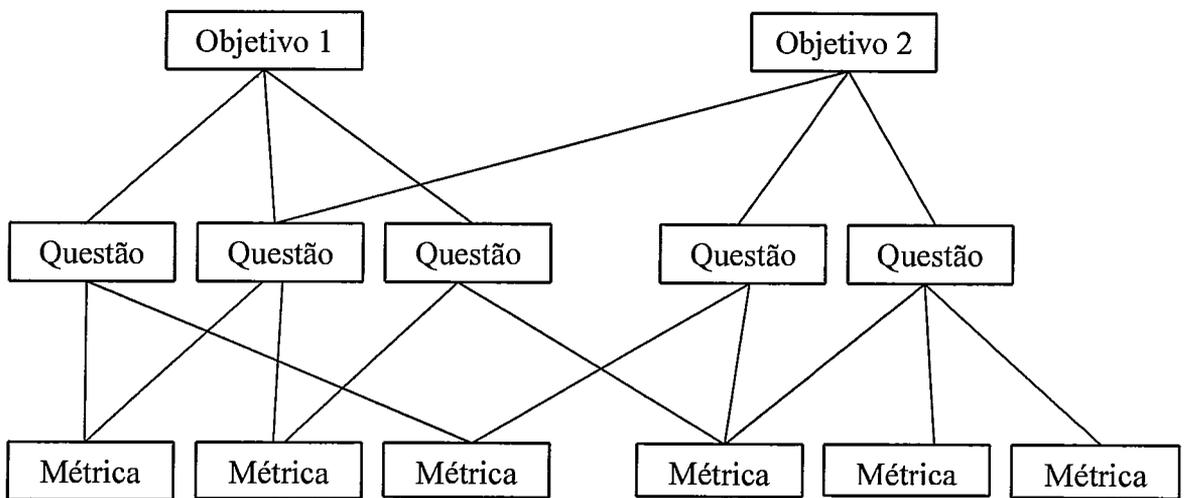


Figura 3.3 – Estrutura hierárquica da abordagem GQM (BASILI *et al.*, 1994)

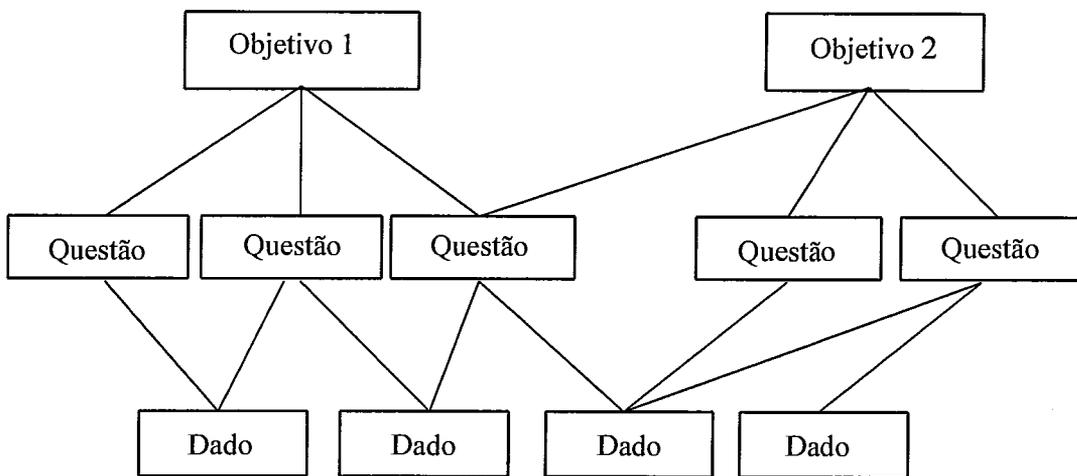


Figura 3.4 – Estrutura hierárquica da abordagem OQD

A seguir são definidas as fases do ciclo de vida, bem como as atividades relacionadas a cada fase, com a descrição de suas respectivas sub-atividades. Este processo pode ser particularizado para projetos específicos, considerado-se suas características particulares. Todavia, os elementos básicos definidos neste processo, que consideramos o processo padrão para construção de bases de dados de pesquisa, devem estar sempre presentes nos processos específicos dele derivados. Uma particularização virá do fato de se ter disponível uma Teoria do Domínio³. Neste caso certas atividades

³ A Teoria do Domínio (OLIVEIRA *et al.*, 1999b) é um modelo composto, basicamente, da organização de conceitos, propriedades e restrições do domínio através de ontologias do domínio e do mapeamento desses conceitos com as tarefas identificadas para o domínio considerado. Este assunto será abordado com mais detalhes no capítulo 5.

serão excluídas e outras incluídas. Na seção 5.3.1 mostraremos a particularização deste processo para situações onde se dispõe de uma Teoria do Domínio.

Um processo deve também, definir quem faz o quê, isto é, quem é que realiza as atividades. Neste processo consideramos seis categorias de participantes:

- coordenador do grupo de pesquisa,
- coordenador de informática,
- coordenador do projeto,
- pesquisadores,
- analistas,
- programadores

I. Instituição do Projeto⁴

Objetivo: Esta fase reúne as atividades a serem realizadas desde uma solicitação de construção de uma base de dados de pesquisa até a sua transformação ou não em um projeto. São definidos os objetivos da base de dados, e é realizado um esboço do planejamento do projeto. Nesta fase deve ser iniciado o Glossário do Projeto⁵, além de ser feita a avaliação da viabilidade de construção da base, em função de seu escopo⁶, nível de detalhamento⁷, tempo e recursos disponíveis. O Glossário deve ser ampliado ao longo do projeto de acordo com a necessidade. A fase é concluída com a aprovação do projeto e com a formação da equipe do projeto ou com a verificação de que este não é viável.

Atividades⁸

I.1. Solicitação de Construção de uma Base de Dados de Pesquisa

A área solicitante identifica a necessidade de construção de uma base de dados de pesquisa e encaminha a solicitação ao coordenador do grupo de pesquisa⁹.

⁴ A fase de Instituição do Projeto pode durar poucos dias no caso dos especialistas estarem com os objetivos da pesquisa e, conseqüentemente, da base de dados, bem definidos. Todavia, se a base for totalmente nova, com escopo e nível de detalhamento grandes, e se, os especialistas que a solicitaram, não estiverem com os objetivos da base bem definidos, esta fase pode ter uma duração mais longa e ter um custo alto.

⁵ Quando se tem disponível uma Teoria do Domínio, esta atividade deve ser excluída (ver capítulo 5)

⁶ É usado para estabelecer que dados (entidades) terão que ser modelados. É definido dentro do domínio.

⁷ Está relacionado com os dados de cada entidade modelada. Refere-se à definição da necessidade de se modelar todos os dados ou só os dados mais gerais. É definido dentro do escopo.

⁸ Ao descrever as atividades estamos considerando o desenvolvimento de uma base de dados totalmente nova (primeiro ciclo de desenvolvimento). Num ciclo de desenvolvimento posterior, estas mesmas atividades podem ser muito mais simples.

⁹ Esta solicitação algumas vezes é verbal, dada a informalidade muito comum em grupos de pesquisa, e, neste caso, não se tem nenhum documento associado.

Produtos¹⁰:

- Solicitação para Construção de Base de Dados, caso pertinente.

Responsabilidades:

Execução: área solicitante

Participação: pesquisadores

Aprovação: coordenador do grupo de pesquisa

I.2. Análise da Solicitação

Após o recebimento, pela coordenação do grupo de pesquisa de uma solicitação para construção de base de dados, este convoca uma reunião com o coordenador de informática, da qual devem participar os pesquisadores solicitantes. Nesta reunião deve ser esclarecido aos pesquisadores que o sucesso da construção da base só será alcançado se os mesmos participarem ativamente das definições iniciais e nas decisões de projeto, sendo esta a forma de garantir a qualidade do produto, a obtenção dos resultados esperados e a adequação dos custos de desenvolvimento por diminuição ou mesmo eliminação do re-trabalho.

Sub-atividades:

- Realização de uma reunião para revisão da solicitação considerando todas as suas implicações, com esclarecimento aos pesquisadores solicitantes, das suas atribuições no processo, explicação da abordagem OQD e de como elaborar o Glossário (caso este seja necessário). Solicita-se a aceitação formal do processo a ser realizado em conjunto por pesquisadores e equipe de informática.
- Definição de um cronograma para as demais atividades desta fase.

Produtos:

- Aceite Formal do Processo
- Cronograma da fase

Responsabilidades:

Execução: coordenador do grupo de pesquisa

Participação: coordenador de informática e pesquisadores

Aprovação: coordenador do grupo de pesquisa, coordenador de informática e pesquisadores.

¹⁰ Os produtos na forma de documentos deverão ter seus roteiros definidos de acordo com os padrões da Organização onde o processo esteja sendo utilizado.

I.3. Especificação Inicial de Requisitos

Nesta atividade os objetivos, questões e dados pertinentes à pesquisa são definidos pelos pesquisadores.

Sub-atividades:

- Definição dos objetivos, questões e dados, pertinentes à pesquisa e à base de dados a ser construída, utilizando a abordagem OQD.
- Avaliação dos objetivos, questões e dados definidos pelos pesquisadores numa reunião de *walkthrough*.

Produtos:

- Especificação Inicial de Requisitos
- Laudo da Reunião de *Walkthrough*

Responsabilidades:

Execução: pesquisadores

Participação: analistas

Aprovação: coordenador do grupo de pesquisa

I.4. Planejamento Inicial do Projeto e Análise de Viabilidade

Nesta atividade são realizados o planejamento inicial do projeto e o estudo de sua viabilidade.

Sub-atividades:

- Planejamento inicial do projeto, identificando os recursos necessários para a realização do projeto (custo, tempo, tecnologia, características e volume de mão de obra).
- Avaliação da viabilidade do projeto, numa reunião, considerando-se os seguintes pontos de vista: viabilidade econômica¹¹, viabilidade financeira, viabilidade de cronograma, viabilidade de mão de obra, viabilidade social¹², viabilidade ambiental¹³ e viabilidade legal¹⁴.

¹¹ Avalia a viabilidade de construção do produto sob o ponto de vista da relação custo-benefício.

¹² Avalia os impactos do produto, sobre o sistema social ao qual irá servir, caso existam.

¹³ Avalia os impactos do produto, sobre o meio ambiente, caso existam.

¹⁴ Avalia os impactos do uso do produto, sob o ponto de vista legal, caso existam implicações legais relacionadas a este uso.

Produtos:

- Plano Inicial do Projeto
- Laudo de Avaliação da Viabilidade

Responsabilidades:

Execução: coordenador de informática

Participação: pesquisadores e analistas

Aprovação: coordenador do grupo de pesquisa

I.5. Formação da Equipe do Projeto

Nesta atividade, caso o projeto tenha sido considerado viável, é constituída a equipe do projeto e é definido o seu coordenador.

Produtos:

- Equipe do projeto formada

Responsabilidades:

Execução: coordenador de informática

Participação: analistas e pesquisadores

Aprovação: coordenador do grupo de pesquisa

II. Definição

Objetivos: Esta fase tem como objetivo, definir detalhadamente os objetivos, estabelecer as consultas que a base deverá responder e os dados necessários para respondê-las. A participação efetiva dos pesquisadores nesta fase é primordial para o sucesso da obtenção dos resultados esperados, pois são definidos os requisitos completos da base de dados. A fase tem início com o planejamento do projeto.

Atividades**II.1. Planejamento do Projeto**

Nesta atividade, o coordenador do projeto elabora o Plano do Projeto. Neste momento faz-se um planejamento macroscópico, que deverá ser detalhado ao longo do desenvolvimento do projeto.

Sub-atividades:

- Elaboração do Plano do Projeto.
- Avaliação do Plano do Projeto em uma reunião de *walkthrough* com a presença do coordenador do grupo de pesquisa, do coordenador de informática e dos pesquisadores solicitantes do projeto.

Produtos:

- Plano do Projeto
- Laudo da Reunião de *Walkthrough*

Responsabilidades:

Execução: coordenador do projeto

Participação: pesquisadores

Aprovação: coordenador do grupo de pesquisa e coordenador de informática

II.2. Análise e Especificação de Requisitos

Nesta atividade, os requisitos são identificados, analisados e documentados. A atividade tem início com o detalhamento, pelos pesquisadores em conjunto com os analistas, dos objetivos da base de dados e das consultas que a base deverá responder, verificando se as mesmas atendem a todos os objetivos definidos. Devem também, ser identificadas as fontes de dados verificando se os dados respondem completamente às consultas estabelecidas. A realização desta atividade inclui ainda, a identificação dos requisitos de qualidade e de interface.

Sub-atividades:

- Detalhamento, pelos pesquisadores em conjunto com os analistas, dos objetivos da base de dados, identificando em seguida, as consultas que serão feitas à base, a curto, médio e longo prazos, verificando se estão condizentes com os objetivos estabelecidos.
- Identificação, pelos pesquisadores em conjunto com os analistas, dos dados e suas respectivas fontes, verificando se estes permitem responder, completamente, às consultas estabelecidas.
- Identificação dos requisitos de qualidade do produto (base de dados) considerando-se a norma ISO/IEC 9126/NBR 13596 (1996) e trabalhos específicos sobre qualidade de dados (PINHO, 2001).
- A coordenação do projeto refina a especificação inicial de requisitos, incluindo na mesma, todos os requisitos identificados: essenciais (escopo da base, estrutura das consultas, funções que devem estar disponíveis na interface com a base, nível de detalhamento, fontes dos dados, tempo e recursos necessários) e de interface com o usuário.

- A coordenação do projeto define a arquitetura¹⁵ para a construção da base (que também constitui a especificação de requisitos), interfaces (entre os componentes da base, e dos pesquisadores com a base construída), dados que povoarão a base e operações manuais.
- Avaliação da Especificação de Requisitos, numa reunião de *walkthrough*.

Produtos:

- Especificação de Requisitos
- Laudo da Reunião de *Walkthrough*

Responsabilidades:

Execução: pesquisadores e analistas

Participação: -

Aprovação: coordenador do grupo de pesquisa e coordenador de informática

III. Desenvolvimento

Objetivos: Esta fase tem como objetivo a construção da base de dados, tendo início com o projeto conceitual e termina com a realização dos testes da base e da sua interface.

Atividades

III.1. Projeto Conceitual

Esta atividade, cujo objetivo é a produção de um modelo conceitual de alto nível, independente do SGBD a ser utilizado, inicia-se a partir da especificação de requisitos, e resulta no modelo conceitual da base de dados.

Sub-atividades:

- Elaboração do modelo conceitual da base de dados, utilizando o procedimento para modelagem de dados baseado na abordagem Entidade Relacionamento¹⁶.
- Avaliação do Modelo Conceitual, numa reunião de *walkthrough*.

Produtos:

- Modelo Conceitual
- Laudo da Reunião de *Walkthrough*

¹⁵ Considera se vai utilizar a tecnologia cliente/servidor em LAN, se vai ter acesso via Web etc.

¹⁶ Este procedimento é definido detalhadamente na seção 4.3.1

Responsabilidades:

Execução: analistas

Participação: pesquisadores

Aprovação: coordenador do grupo de pesquisa e coordenador de informática

III.2. Projeto Lógico

O objetivo desta atividade é representar, a partir do modelo conceitual, as funções, execuções, buscas e restrições da base de dados, seus dados e controles, suas entradas e saídas, de forma independente dos dispositivos de armazenamento, resultando no modelo lógico da base e no projeto de sua interface com o usuário.

Sub-atividades:

- Evolução do modelo conceitual para o modelo lógico da base de dados, mapeando os construtores (entidade, relacionamento, atributo, generalização / especialização e agregação) do modelo conceitual em elementos do modelo lógico (tabelas com suas linhas e colunas, restrições de integridade, normalização das tabelas e chaves para recuperação das informações), considerando a topologia¹⁷ do banco de dados a ser utilizado.
- Realização de uma reunião de *walkthrough* para avaliação do Modelo Lógico
- Construção de um protótipo da interface.
- Avaliação do protótipo da interface.

Produtos:

- Modelo Lógico
- Protótipo da Interface
- Laudo da Reunião de *Walkthrough*
- Laudo da Avaliação do Protótipo

Responsabilidades:

Execução: analistas

Participação: -

Aprovação: coordenador de informática

¹⁷ Relacional, Hierárquica, Rede etc.

III.3. Projeto Físico¹⁸

Esta atividade tem como objetivo a transformação do modelo lógico no projeto físico.

Sub-atividades:

- Evolução do modelo lógico para o modelo físico da base de dados, mapeando os elementos do modelo lógico (tabelas com suas linhas e colunas, restrições de integridade, normalização das tabelas e chaves para recuperação das informações) para as estruturas físicas de implementação, considerando os recursos necessários para o armazenamento e manipulação das estruturas de dados, endereçamento, acesso e alocação física, utilizando os recursos do SGBD.
- Descrição e detalhamento das estruturas e procedimentos dos demais componentes da base, de forma que possam ser codificados e testados.
- Atualização do modelo lógico, em função do modelo físico resultante. Ou seja, verifica se as decisões para melhoria de desempenho tomadas no projeto físico, afetam a estrutura descrita no modelo lógico.
- Elaboração do projeto da interface.
- Realização de uma reunião de *walkthrough* para avaliação dos modelos lógico e físico.
- Planejamento dos testes.

Produtos:

- Projeto Físico
- Projeto Lógico atualizado
- Plano do Projeto atualizado com Plano de Testes

Responsabilidades:

Execução: analistas

Participação: -

Aprovação: coordenador de informática

III.4. Codificação

Esta atividade tem como objetivo a construção, a partir do Projeto Físico e do Projeto da Interface, da base de dados e das consultas a serem realizadas.

¹⁸ Esta atividade é totalmente dependente do SGBD a ser utilizado.

Sub-atividades:

- Os programadores codificam, testam e documentam o código.

Produtos:

- Base de dados implementada

Responsabilidades:

Execução: programadores

Participação: -

Aprovação: analistas

III.5. Teste com o Usuário

Nesta atividade, a equipe de desenvolvimento, submete a base e a interface à avaliação final, com base no Plano de Testes, através do uso de dados e consultas reais. Ao final desta etapa tem-se a base de dados testada e aceita para fins de teste em uso, objetivo da fase de Transição.

Sub-atividades:

- Realização dos testes com os usuários (pesquisadores), de acordo com o Plano de Testes.
- Realização de correções e ajustes identificados.
- Elaboração do Manual do Usuário.

Produtos:

- Base de Dados testada
- Laudo do Teste com o Usuário
- Manual do Usuário

Responsabilidades:

Execução: analistas

Participação: programadores

Aprovação: pesquisadores

IV. Transição

Objetivos: Nesta fase, o objetivo é colocar a base de dados em operação no ambiente real para que seja povoada e avaliada, de acordo com o planejamento feito. Esta fase tem início com a preparação para a Implantação e termina com a homologação do produto pelo coordenador do grupo de pesquisa.

Atividades

IV.1. Preparação para Implantação

O objetivo desta atividade é preparar a infra-estrutura necessária para dar início ao uso do produto.

Sub-atividades:

- Detalhamento do Plano do Projeto (Plano de Treinamento e Plano de Implantação e Operação).
- Definição da estrutura e responsabilidades para suporte ao usuário.

Produtos:

- Plano do Projeto atualizado
- Infra-estrutura de operação da base definida

Responsabilidades:

Execução: analistas

Participação: -

Aprovação: coordenador de informática

IV.2. Avaliação em uso da Base de Dados

Nesta atividade é iniciada a população da base e os pesquisadores iniciam a sua utilização, avaliando a facilidade de operação e o atendimento dos requisitos estabelecidos.

Sub-atividades:

- População da base de dados.
- Avaliação, em uso, da base de dados.
- Realização, pela equipe do projeto (analistas e programadores), de correções/ajustes que tenham sido identificados.
- Elaboração do Relatório Final da Fase de Transição.

Produtos:

- Base de Dados povoada e avaliada em uso
- Relatório Final da Fase de Transição

Responsabilidades:

Execução: pesquisadores

Participação: analistas e programadores

Aprovação: pesquisadores

IV.3. Homologação

O objetivo desta atividade é homologar o produto, a partir dos resultados obtidos com a utilização da base de dados durante a fase de Transição e descritos no Relatório Final da Fase de Transição.

Sub-atividades:

- Avaliação do Relatório Final da Fase de Transição e homologação do produto.

Produtos:

- Base de dados operacional
- Relatório de Homologação

Responsabilidades:

Execução: coordenador do grupo de pesquisa

Participação: pesquisadores

Aprovação: coordenador do grupo de pesquisa

V. Manutenção

Objetivos: Esta fase compreende as atividades para realização de modificações que se tornem necessárias, sejam elas por motivos de evolução, correção ou adaptação para atender ao surgimento de novos requisitos. São tratadas aqui, apenas, manutenções evolutivas de pequeno porte. Necessidades de evolução de maior porte são objeto de um novo ciclo de desenvolvimento (figura 3.2)

Atividades

V.1. Identificação de Novo Requisito

Nesta atividade, os usuários definem novos requisitos, o que implica em modificação do produto. No caso de requisitos que impliquem em mudanças substanciais no produto, esta atividade ocasiona um novo ciclo de desenvolvimento.

Sub-atividades:

- Identificação, pelos pesquisadores, de novos requisitos.
- Análise da solicitação e encaminhamento para manutenção para realização da modificação ou definição de novo ciclo de desenvolvimento.

Produtos:

- Identificação de Novo Requisito

Responsabilidades:

Execução: pesquisadores

Participação: -

Aprovação: coordenador de informática

V.2. Identificação de Erro

Nesta atividade, os usuários descrevem erros percebidos com o uso da base de dados.

Sub-atividades:

- Identificação de erros no produto
- Análise da solicitação e encaminhamento para a manutenção, para realização da modificação.

Produtos:

- Identificação de Erro

Responsabilidades:

Execução: pesquisadores

Participação: -

Aprovação: coordenador de informática

V.3. Realização da Modificação

Nesta atividade são efetuadas as modificações solicitadas no produto.

Sub-atividades:

- Definição da solução, pelos analistas.
- Realização das modificações pelos programadores.
- Realização dos Testes
- Atualização da documentação
- Aprovação das modificações efetuadas e liberação para uso no ambiente operacional.

Produtos:

- Base de dados modificada
- Documentação do produto atualizada

Responsabilidades:

Execução: analistas e programadores

Participação: -

Aprovação: coordenador do grupo de pesquisa e pesquisadores

3.3 – Considerações Finais

A diversidade de tipos de sistemas e domínios de aplicação com a qual os engenheiros de software se deparam atualmente implica na adequação de metodologias, processos e técnicas para os diferentes contextos das organizações.

Especificamente em instituições com atividades em ciência e tecnologia, a construção de bases de dados de pesquisa é uma atividade em constante desenvolvimento. Neste capítulo, foi definido um processo padrão para este tipo de necessidade, considerando atividades específicas para o levantamento dos dados, nas diversas fontes identificadas através da participação efetiva dos usuários ao longo do desenvolvimento.

No próximo capítulo verificaremos como ontologias de domínio podem apoiar a execução das etapas do processo geral de construção de bases de dados. Também é organizado um procedimento que apoia a execução da atividade Projeto Conceitual (III.1) da fase Desenvolvimento (III), do Processo de Construção de Bases de Dados de Pesquisa, definido no capítulo 3.

Capítulo 4

Definição de Bases de Dados com Apoio de Ontologias de Domínio

Este capítulo mostra como ontologias de domínio podem apoiar à construção de bases de dados, apresentando também um procedimento organizado para apoiar a execução de uma das principais atividades do processo apresentado no capítulo anterior, o projeto conceitual.

4.1 – Introdução

Ontologias de domínio podem ser utilizadas para apoiar o desenvolvedor ao longo do processo de desenvolvimento de software, uma vez que elas contêm os componentes necessários para auxiliar no entendimento do conhecimento do domínio.

O uso de ontologias de domínio para apoio à construção de bases de dados deve-se ao fato de que, as mesmas constituem um modelo conceitual que retrata diretamente o domínio em si, num nível de abstração capaz de representar o conhecimento, proporcionando a sua reutilização.

COUGO (1997) observa que a "lei do mundo" diz: o mundo está cheio de coisas que possuem características próprias e que se relacionam entre si. O modelo entidade-relacionamento (MER), objetiva modelar essa "lei". Além disso, WAND (1996) afirma que a ontologia pode ser definida como a "teoria filosófica preocupada com as características básicas do mundo". Pode-se notar então, uma forte analogia entre as "coisas do mundo", os conceitos do modelo entidade-relacionamento e os conceitos ontológicos.

Para mostrar a possibilidade do apoio por ontologias de domínio ao processo de construção de bases de dados, apresentamos nas seções 4.2 e 4.3 onde e como a ontologia pode ser usada no apoio ao desenvolvimento das etapas da construção de bases de dados. Para auxiliar nesse processo, definimos um mapeamento dos construtores da ontologia para os construtores da modelagem conceitual ER (apresentado na seção 4.4), juntamente com a equivalência entre os construtores gráficos da linguagem de modelagem de ontologias utilizada e da diagramação entidade-relacionamento.

O procedimento organizado e apresentado na seção 4.3.1 (figura 4.1) é utilizado na execução da atividade *projeto conceitual* da fase *desenvolvimento*, tal como mostra a figura 3.2.

4.2 – Uso de ontologias de domínio nas etapas do processo de construção de bases de dados

Conforme descrito no capítulo dois (seção 2.2), a modelagem de dados constitui-se de três níveis de abstrações diferentes dos dados, e que correspondem às três etapas principais da construção de bases de dados: projeto conceitual, projeto lógico e projeto físico.

Projeto Conceitual

O projeto conceitual descreve a estrutura do conteúdo da informação da base de dados, registrando no MER os dados que devem aparecer na base de dados, ao invés da estrutura de armazenamento necessária para gerenciar esta informação. Nesta etapa o uso da ontologia de domínio além de apoiar o entendimento do domínio através da semântica embutida em sua estrutura, pode ter seus construtores mapeados para os construtores do diagrama entidade-relacionamento (DER), que é um dos principais documentos do modelo conceitual, além de auxiliar na construção do dicionário de dados.

O apoio ao projeto conceitual será abordado com mais detalhes, na seção 4.3, onde mostraremos como ontologias de domínio podem apoiar na execução do procedimento apresentado na seção 4.3.1, e na seção 4.4 onde mostraremos o mapeamento dos construtores ontológicos para os construtores da modelagem conceitual.

Projeto Lógico

O modelo conceitual ER tem poder limitado de expressão, uma vez que pode representar apenas a estrutura dos objetos do domínio que irão compor a base de dados. Para fazer com que este modelo corresponda à realidade que se deseja modelar, é necessário definir restrições adicionais, uma vez que o MER não lida com restrições de integridade (DATE, 2000).

O projeto lógico estabelece essas restrições adicionais, pois define ao nível de SGBD, as estruturas de dados que implementarão os requisitos identificados no modelo conceitual. Para expressar o modelo lógico, o modelo relacional é o mais comum e mais utilizado (COUGO, 1997; HEUSER, 2001).

A ontologia de domínio pode ser utilizada no processo de transformação do MER para o modelo lógico durante a normalização, permitindo a verificação da fidelidade das estruturas das entidades e dos relacionamentos em relação ao ambiente observado e permitindo a representação das características observadas, assim como elas efetivamente são.

A ontologia de domínio, através de seus axiomas que exprimem cardinalidade mínima e máxima das relações, também permite que o projetista escolha corretamente a forma de mapeamento adotada em relação aos relacionamentos, que pode ser: tradução do relacionamento em tabela própria; tradução do relacionamento através da inserção de colunas em uma tabela correspondente a uma das entidades que participam do relacionamento; ou, tradução do relacionamento através da fusão das tabelas referentes às entidades envolvidas no relacionamento.

Para garantir a qualidade dos dados através de sua integridade, nesta etapa são definidas as restrições de integridade, cujos diferentes tipos estão resumidos na tabela 4.1 (BATINI *et al.*, 1992; COUGO, 1997; DATE, 2000). Uma restrição de integridade é uma regra de consistência de dados que é garantida pelo próprio SGBD, para que a base de dados reflita corretamente a realidade representada.

Tabela 4.1 – Tipos de Restrições de Integridade para a abordagem relacional

TIPO	DESCRIÇÃO
Integridade de Domínio (da coluna/do atributo)	Especifica os possíveis valores a serem assumidos por cada uma das colunas (observadas a partir das regras de negócio), em função de características do tipo de dado e do próprio SGBD onde essa coluna venha a ser colocada.
Integridade de Vazio (se a coluna é obrigatória ou opcional)	Especifica as características conceituais de presença ou não de um dado (colunas “permitidas” ou “não permitidas” para conter valores nulos), extraídas a partir das regras de negócio.
Integridade de Chave	Especifica que a chave primária e a alternativa não devem conter valores duplicados ou nulos.
Integridade Referencial	Especifica que não pode existir na chave estrangeira, um valor que não exista na tabela na qual ela é chave primária, para manter a consistência entre tuplas de duas relações (Se B faz referência a A, então A tem que existir).

A ontologia de domínio pode ser utilizada para definir as restrições de *integridade de domínio*, uma vez que a especificação de intervalos para domínios contínuos e domínios discretos, pode estar explicitamente representada na ontologia, através de seus axiomas, nas tabelas de descrição das propriedades de instância, ou podem ser inferidas pela própria definição dos conceitos, relações ou axiomas relacionados com a coluna que esteja sendo modelada.

OLIVEIRA (1999), ao definir uma Teoria do Domínio para cardiologia definiu o seguinte axioma expresso em linguagem natural e formalizado em lógica de primeira ordem, que estabelece o domínio para o atributo *tipo* do conceito *pulso*:

Todo pulso é arritmico, rítmico ou biferens

$$(\forall p) (\text{pulso}(p) \rightarrow (\text{tipo}(p, \text{aritmico}) \wedge \neg \text{tipo}(p, \text{rímico}) \wedge \neg \text{tipo}(p, \text{biferens})) \vee (\neg \text{tipo}(p, \text{aritmico}) \wedge \text{tipo}(p, \text{rímico}) \wedge \neg \text{tipo}(p, \text{biferens})) \vee (\neg \text{tipo}(p, \text{aritmico}) \wedge \neg \text{tipo}(p, \text{rímico}) \wedge \text{tipo}(p, \text{biferens}))$$

Ao estabelecer a integridade de domínio para uma base de dados que contenha dados sobre tipo de pulso, o projetista utilizará este axioma na sua definição.

A *integridade de vazio*, também pode ser apoiada pela ontologia de domínio, uma vez que pode estar explicitamente definido se uma propriedade de um conceito pode ou não ser vazio, se é obrigatória ou opcional. Ou, através dos axiomas que exprimem as relações e suas propriedades.

Ainda com base na Teoria do Domínio para cardiologia definida por OLIVEIRA (1999), temos novamente a definição de um axioma expresso em linguagem natural e formalizado em lógica de primeira ordem, mostrado a seguir, que servirá de base para a definição da integridade de vazio.

Um fator patogênico pode não estar associado a nenhum fator etiológico

$$(\exists p, \forall e) (\text{patogenia}(p) \wedge \text{etiologia}(e) \rightarrow \neg \text{associada}(p, e))$$

Ou seja, se existir um relacionamento, que possua atributos, entre patogenia e etiologia, o projetista define que o seu conteúdo é opcional (pode ser vazio).

As demais restrições de integridade (de chave e referencial) são garantidas, automaticamente pelos SGBD's, independente da existência ou não de uma ontologia de domínio.

As restrições que não se encaixam em nenhuma dessas categorias não são garantidas pelo SGBD, de forma automática, e são chamadas *restrições semânticas*. Este tipo de restrição é totalmente apoiado por ontologias de domínio, principalmente no que é estabelecido pelos seus axiomas. Além disso, elas requerem que o programador escreva procedimentos para garanti-las explicitamente (HEUSER, 2001), podendo utilizar os axiomas da ontologia de domínio na sua construção, conforme exemplificado abaixo, com base num axioma definido na Teoria do Domínio para cardiologia (OLIVEIRA, 1999).

Axioma	Toda patologia que é produzida por um fator etiológico não é uma má formação congênita
	$(\forall p, \exists e) (\text{patologia}(p) \wedge \text{etiopatogenia}(e) \wedge \text{produz}(e,p) \rightarrow \neg \text{ma_formacaocong}(p))$
Trecho de programa	Se PROD.PATOLOGIA="Fator Etiológico" Então PATOLOGIA <> "má formação congênita"

Projeto Físico

O projeto físico inicia do modelo lógico e resulta no modelo físico, que é uma descrição da implementação da base em memória secundária. Descreve as estruturas de armazenamento e métodos de acesso utilizados para acesso efetivo aos dados; onde é detalhado o armazenamento interno das informações, que podem influenciar o desempenho das aplicações.

O projeto físico é especificado para um SGBD específico. Cada SGBD define um modo diferente de implementação física das características e recursos necessários para o armazenamento e manipulação das estruturas de dados (estruturas de armazenamento, endereçamento, acesso e alocação física), sendo inclusive, específicos para cada sistema operacional.

As decisões tomadas durante esta etapa para melhorar o desempenho, podem afetar a estrutura do modelo lógico. Uma vez que ambos estejam prontos, os modelos lógico e físico são expressos usando a ddl (linguagem de definição de dados) do SGBD utilizado, resultando na base criada, podendo a mesma ser populada.

Como esta etapa é executada apenas por especialistas que fazem sintonia de banco de dados, procurando otimizar o desempenho, estando baseada totalmente no atendimento do que está definido na especificação de requisitos (aspectos como desempenho, flexibilidade, independência e segurança), a ontologia pode apoiar na comunicação entre os especialistas para sanar qualquer dúvida relacionada ao entendimento do domínio e também ao que se refere às restrições semânticas de integridade.

4.3 – Projeto Conceitual com base em Ontologias de Domínio

Nesta seção mostraremos como a execução da atividade *projeto conceitual* da fase *desenvolvimento*, tal como definido na seção 3.2, pode ser apoiada pela utilização do conhecimento disponível na forma de ontologias de domínio. A elaboração do MER da base de dados sendo projetada é feita com base no procedimento definido na seção 4.3.1 a seguir, organizado com base nos trabalhos de BATINI *et al.* (1992), MACHADO e ABREU (1996), COUGO (1997), DATE (2000) e HEUSER (2001).

4.3.1 – Procedimento para Modelagem Conceitual ER

A modelagem conceitual é um processo onde a percepção da realidade é refinada e enriquecida progressivamente, e o modelo conceitual é gradualmente desenvolvido. Este procedimento deve ser repetido tantas vezes quantas forem as extensões agregadas ao modelo, seja em sua fase de construção ou em suas manutenções evolutivas, adaptativas ou corretivas.

O principal objetivo da modelagem conceitual é gerar um modelo com ênfase nas *propriedades estruturais do dado* que é independente dos detalhes de implementação. Trata-se aqui de construir uma descrição da realidade que seja fácil de entender e interpretar.

O modelo Entidade-Relacionamento foi selecionado devido ser amplamente utilizado, ser baseado em poucos conceitos de modelagem, ter uma representação gráfica efetiva na qual cada elemento do modelo é mapeado para um símbolo gráfico distinto, e ser de fato um padrão de muitas ferramentas automáticas para apoio a modelagem conceitual (HEUSER, 2001). A figura 4.1 apresenta as etapas que compõem o procedimento a ser seguido para gerar o modelo conceitual de uma base de dados.

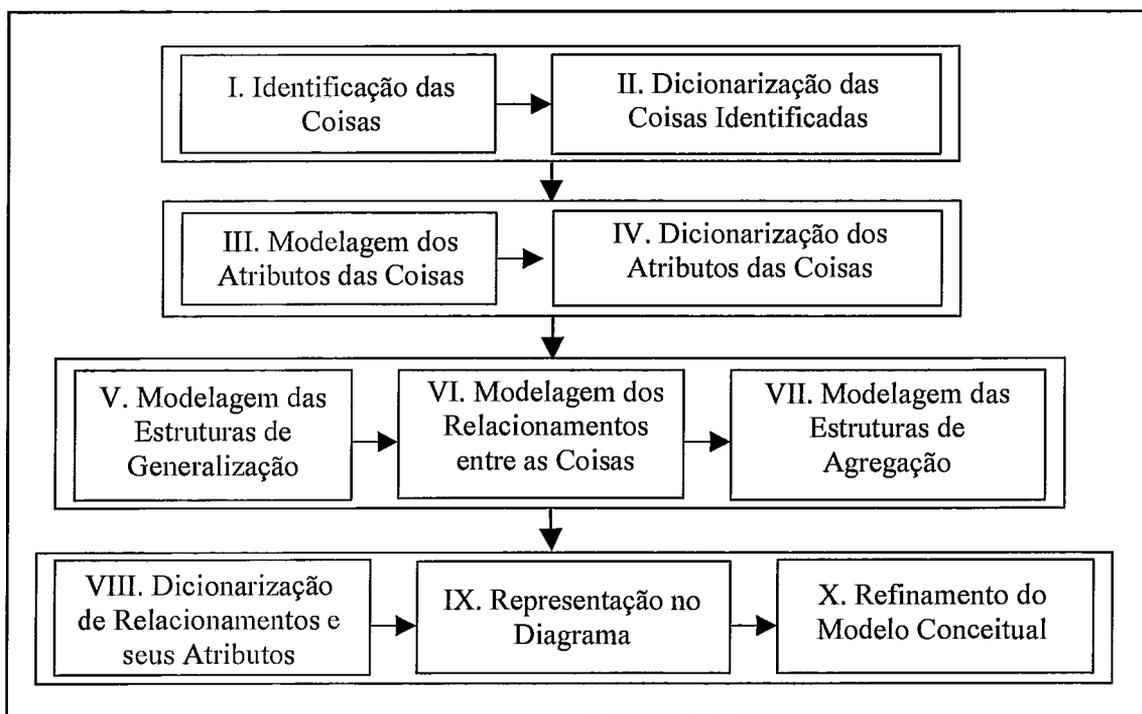


Figura 4.1 – Procedimento para Modelagem Conceitual

A seguir, são apresentadas as descrições das atividades e sub-atividades a serem executadas, constituindo o roteiro.

I. Identificação das Coisas¹⁹

Identificar as coisas existentes no domínio, que sejam elementos componentes ou de sustentação de algum conjunto de atividades do domínio e que proporcione algum interesse na manutenção de dados (informações armazenadas sobre elas).

Sub-atividades

I.1. Levantamento, investigação e análise de dados – capturar as informações que caracterizam devidamente as coisas a serem modeladas; sejam elas concretas ou imaginárias, que devem ser identificadas, entendidas e representadas.

- *Identificação das coisas* - determinar as coisas ou elementos individualizados, através da observação de cinco grandes grupos de elementos: Coisas tangíveis, Funções exercidas por elementos, Eventos ou ocorrências, Interações e Documentos.
- *Entendimento das coisas* - conceituar, entender e assimilar as coisas identificadas com todas as suas características.

¹⁹ Os conceitos com os quais um modelo conceitual se preocupa são: existem algumas *coisas* que têm certas *propriedades* e que podem estar *relacionadas* de alguma ou várias formas, com outras coisas (Bowers, 1988).

- *Classificação da abstração*²⁰ - definir os conceitos (identificar as entidades) como um agrupamento de objetos do mundo real, caracterizados por propriedades em comum (atributos, relacionamentos, generalizações e subconjuntos).

II. Dicionarização das Coisas Identificadas

Definir claramente, cada uma das coisas identificadas e representadas, associando-se seu nome, sua representação e sua definição, proporcionando o completo entendimento do conceito que estes procuram transmitir.

Sub-atividades

II.1. Referência das coisas – referenciar no dicionário, a própria coisa que ela representa, não fazendo nenhuma referência à coisa como um meio de armazenamento.

II.2. Verificação - verificar se a dicionarização satisfaz, senão todas, a grande parte das seguintes perguntas:

- O que é a coisa ?
- O que faz a coisa ?
- Para que serve a coisa ?
- O que engloba essa categoria de coisas ?
- O que está excluído dessa categoria de coisas ?
- Quando alguém ou algo passa a ser, ou deixa de ser, uma coisa desse tipo ?
- Sua permanência nessa categoria é imutável ?
- Quais estados²¹ esta coisa possui durante seu ciclo de vida ?

III. Modelagem dos Atributos das Coisas

Identificar as características próprias das coisas de um ambiente, que são inerentes a cada uma dessas coisas, e serão em princípio, comuns a todas as coisas, ou elementos individualizados, pertencentes a um mesmo conjunto.

Sub-atividades

III.1. Identificação de atributos - determinar todas as características próprias das coisas.

Um atributo é identificado quando a coisa tem uma estrutura atômica e nenhuma propriedade de interesse parece ser aplicável a mesma. Mas, observar que o atributo pode ser atômico ou composto²².

²⁰A abstração é um processo mental utilizado para selecionar algumas características e propriedades de um conjunto de objetos e excluir outras características que não sejam relevantes.

²¹ Estado, são os advérbios que são utilizados na referência do objeto durante a sua existência (ex.: aberto; fechado; cancelado; atendido).

²²Um atributo composto é um grupo de atributos que têm uma afinidade no significado ou uso. Ex.: Data = dia/mês/ano

III.2. Definir os atributos identificadores²³ - determinar os atributos que identificam as entidades.

IV. Dicionarização dos Atributos das Coisas

Definir claramente cada um dos atributos identificados e representados, associando-se seu nome, sua representação e sua definição, proporcionando a completa caracterização do universo observado.

Sub-atividades

IV.1. Entendimento dos atributos - conceituar, entender e assimilar os atributos identificados, tornando-os parte do conhecimento.

IV.2. Referência de atributos - referenciar no dicionário, toda e qualquer informação relacionada aos atributos identificados, proporcionando o completo entendimento do conceito que está sendo transmitido.

V. Modelagem das Estruturas de Generalização

Reconhecer as abstrações de generalização, definindo um subconjunto de relacionamentos entre elementos de duas ou mais entidades.

Sub-atividades

V.1. Identificação de propriedades adicionais das coisas - localizar as entidades que possuam o mesmo conjunto de atributos para descrevê-las, verificando se, uma entidade E é uma generalização de um grupo de entidades E_1, E_2, \dots, E_n observando, se cada objeto das entidades E_1, E_2, \dots, E_n é também um objeto da entidade E .

V.2. Mapeamento das generalizações - estabelecer o mapeamento da entidade genérica para o subconjunto de entidades, e utilizar o conceito de Generalização / Especialização para representá-las, verificando que cada entidade pode estar envolvida em múltiplas generalizações e que o subconjunto de entidades pode ser composto por apenas uma entidade especializada.

V.3. Observação da propriedade fundamental da generalização - verificar se todas as propriedades da entidade genérica são herdadas pelo subconjunto de entidades especializadas.

V.4. Entendimento das generalizações - conceituar, entender e assimilar as estruturas identificadas, tornando-as parte do conhecimento.

²³Um identificador de uma entidade E é uma coleção de atributos ou de entidades relacionadas a E que têm a propriedade de determinar unicamente todas as instâncias de E .

V.5. Referência das generalizações – referenciar no dicionário, trazendo ao conhecimento público toda e qualquer informação que seja de valia para o processo de compreensão e unificação de conceitos.

VI. Modelagem dos Relacionamentos entre as Coisas

Reconhecer todas as relações existentes entre as coisas observadas, considerando que cada coisa poderá se relacionar com diversas outras coisas independente de seu tipo. Identificar os relacionamentos entre pares de entidades.

Sub-atividades

VI.1. Identificação - discriminar os relacionamentos, observando além das próprias coisas, as características intrínsecas de cada uma das coisas, ou do meio onde estão situadas. Identificar e destacar os relacionamentos existentes em especializações de uma entidade com outras do modelo.

VI.2. Denominação dos relacionamentos - determinar os nomes dos relacionamentos, de forma que representem exatamente o conceito observado.

VI.3. Caracterização dos relacionamentos - identificar e mapear as características diversas dos relacionamentos. Todo relacionamento tem um nome, um significado, um grau mínimo e máximo, pode ser independente²⁴, contingente²⁵ e mutuamente exclusivo²⁶.

- *Representação das "regras de negócio"* - representar as "regras do negócio" através dos seguintes requisitos: grau, ou cardinalidade, do relacionamento; número de elementos que participam do relacionamento, condição de participação dos elementos no relacionamento e condição de estabelecimento do relacionamento.
- *Verificação da representação* - verificar se as principais informações que devem ser utilizadas para caracterizar um relacionamento de forma inequívoca, como: sua função, o que ele representa, quais são as regras de seu estabelecimento, quais são as exceções a seu estabelecimento, quando ele ocorre, quando ele pode deixar de existir; foram todas utilizadas e mapeadas. Observar que podem existir relacionamentos recursivos, que conectam uma entidade a ela mesma.

²⁴ É todo e qualquer relacionamento que possa vir a ser estabelecido sem que haja necessidade de avaliação simultânea de outro relacionamento. Tratam de conceitos diferentes, podendo ser estabelecidos tanto entre entidades diferentes como também, entre as mesmas entidades.

²⁵ Estabelece uma correlação entre dois ou mais relacionamentos, através da análise da semântica de suas associações, e também define que como regra de negócio, mais de um relacionamento deve ocorrer em um mesmo instante.

²⁶ Ocorre quando a associação de um elemento for estabelecida através de um dos relacionamentos, não pode ser estabelecida pelos demais.

- *Definição de atributos* - determinar e mapear os dados a serem manuseados, ou atributos que denotam a existência de informações que só podem ser estabelecidas ou consideradas quando da existência de uma associação entre elementos, observando que somente em relacionamentos de cardinalidade muitos-para-muitos podem existir atributos.
- *Representação de atributos* - alocar os atributos pertencentes ao relacionamento corretamente, representando-os no próprio relacionamento, cuja cardinalidade seja de muitos-para-muitos, sem a necessidade de criar uma pseudo-entidade para acomodá-los.
- *Verificação da representação* - examinar se foi identificado o local de apropriação dos atributos, se na entidade ou no relacionamento, em função de seus significados e instante de surgimento. Verificar, se em qualquer instante que o relacionamento não for mais válido e deixar de existir, os atributos também deixarão de ter significado.

VI.4. Verificação da modelagem dos relacionamentos - avaliar se foi demonstrado através do mapeamento dos relacionamentos, como um objeto se comporta em relação aos demais, qual seu grau de dependência de outros objetos, qual a associação de dados existente entre eles, entre outros fatores.

VII. Modelagem da Estrutura de Agregação

Verificar se o entendimento semântico pode ser facilitado e se os graus dos relacionamentos ternários ou de maior número podem ficar mais claros através do uso da estrutura de agregação.

Sub-atividades

- VII.1. Identificação de Agregação - estabelecer a existência de uma junção de elementos através de um relacionamento, cuja existência depende da ocorrência do outro relacionamento considerado fundamental.
- VII.2. Verificação da agregação - verificar se esta estrutura foi utilizada somente quando existiu um relacionamento de muitos-para-muitos, que representou um fato, e existiu ainda, uma terceira entidade que está relacionada a este fato.
- VII.3. Verificação de Auto agregação e Agregações Estendidas – verificar se existem duas agregações relacionadas entre si, ou se existem casos de auto-agregação, que ocorrem quando temos um auto relacionamento muitos para muitos que se relaciona com alguma outra entidade.

VIII. Dicionarização de Relacionamentos e seus Atributos

Agregar informações adicionais sobre o relacionamento estabelecido entre as entidades e das agregações identificadas.

Sub-atividades

- VIII.1. Referência de relacionamentos – referenciar de forma textual, descrevendo o relacionamento tanto quanto sua interpretação e seu sentido de leitura.
- VIII.2. Referência da agregação – referenciar de forma textual, descrevendo a agregação tanto quanto sua interpretação.
- VIII.3. Verificação – analisar, se o papel do relacionamento e da agregação, estão descritos de forma que possam ser entendidos e se o grau de ambigüidade que possa vir a ser gerado pela simples "leitura" de um modelo foi diminuído.
- VIII.4. Entendimento dos atributos - conceituar, entender e assimilar os atributos identificados, tornando-os parte do conhecimento.
- VIII.5. Referência de atributos – referenciar no dicionário, trazendo ao conhecimento público toda e qualquer informação que seja de valia para o processo de compreensão e unificação de conceitos.
- VIII.6. Condições e Regras do Relacionamento – destacar, para as agregações, quais as condições em que existe o relacionamento, observando que todo o processo de uma agregação é condicionado à pré-existência de uma ocorrência do relacionamento que estamos “agregando”.

IX. Representação no Diagrama

Representar todos os construtores identificados no diagrama entidade-relacionamento (DER), de acordo com a notação gráfica escolhida.

Sub-atividades

- IX.1. Representação das entidades - representar as entidades identificadas no diagrama ER utilizando o construtor adequado (entidade simples ou generalização / especialização).
- IX.2. Representação dos atributos - representar os atributos identificados no diagrama ER utilizando o construtor adequado, preservando a clareza do diagrama.
- IX.3. Representação conceitual dos relacionamentos - representar conceitualmente o relacionamento no diagrama ER utilizando o construtor adequado.

IX.4. Representação conceitual da agregação - representar conceitualmente a agregação no diagrama ER utilizando o construtor adequado.

X. Refinamento do Modelo de Dados

Construir o modelo final gradativamente, utilizando um esquema de transformações com as seguintes características: cada esquema de transformação tem um modelo inicial, no qual o esquema é aplicado, resultando no modelo final; cada esquema de transformação mapeia nomes de conceitos no modelo inicial em nomes de conceitos no modelo resultante; os conceitos no modelo resultante devem herdar todas as conexões lógicas definidas para o modelo inicial.

Sub-atividades

X.1. Utilização das primitivas²⁷ de refinamento - selecionar e aplicar o que for mais adequado ao domínio sendo modelado, dentro de um limitado conjunto de transformações que podem ser aplicadas ao modelo inicial.

- *Refinamento de entidades* - refinar entidades através de: um relacionamento ou em duas outras entidades; em hierarquias de generalização; dividir uma entidade em um conjunto de entidades independentes, introduzindo novas entidades, sem o estabelecimento de relacionamentos ou generalização entre elas; gerar novas entidades, quando o projetista descobrir um novo conceito com propriedades específicas que não apareceram no modelo anterior; criar uma nova entidade, definida como generalização entre as entidades definidas previamente.
- *Refinamento de relacionamentos* - refinar relacionamentos através de: dois ou mais relacionamentos entre as mesmas entidades; verificar se um relacionamento entre dois conceitos deveria ser expresso via um terceiro conceito, que não estava visível no modelo anterior; gerar um novo relacionamento entre duas entidades previamente definidas; verificar a cardinalidade de todos os relacionamentos.
- *Refinamento de atributos* - refinar atributos das seguintes formas: introduzir novos atributos em entidades e relacionamentos; transformar um atributo atômico em atributo composto ou em um grupo de atributos.

X.2. Verificação do modelo resultante - verificar se o modelo resultante preserva as informações do modelo inicial e se a nova organização melhorou a clareza do modelo.

²⁷ Uma aplicação rigorosa das primitivas auxilia os projetistas, especialmente os iniciantes, devido a sua facilidade de aplicação e devido elas dividirem o processo de modelagem em passos simples e corretos (ver BATINI *et al.*, 1992).

- *Verificação da completude* – analisar se o modelo representa todas as características relevantes do domínio, através da verificação de que todos os requisitos estão representados no modelo.
- *Verificação da correção* - verificar se o modelo utiliza corretamente os conceitos do modelo ER, sintática e semanticamente, e se a representação gerada, associada aos objetos já existentes previamente no modelo, continua a formar um conjunto, em sua totalidade, coerente e fiel aos conceitos encontrados.
- *Verificação da expressividade* - analisar se o modelo representa todos os requisitos de forma natural e se pode ser facilmente entendido.
- *Verificação da estética*²⁸ – colocar o fato principal²⁹ objeto da modelagem no centro do diagrama. Derivar para o lado direito as entidades significativas de negócio e os relacionamentos entre elas. Ao lado esquerdo colocar as entidades que identificam cadastros básicos e relacioná-las com os devidos fatos e/ou entidades.
- *Verificação da facilidade de leitura* - verificar se o DER permite o entendimento do modelo, inclusive por leigos.
- *Verificação da extensibilidade* - analisar se o modelo é facilmente adaptável a mudanças de requisitos, verificando se pode ser decomposto em módulos ou visões, onde as modificações possam ser introduzidas.

Como a finalidade de um roteiro é a geração de um ou mais documentos (FALBO, 1998), ao final da execução deste, obteremos o Diagrama ER e um descritivo no Dicionário de Dados de cada entidade com seus atributos, e dos relacionamentos existentes entre estas entidades com seus atributos e expressões de relacionamento, constituindo o modelo final (MER).

A tabela 4.2 mostra o resultado da execução das atividades que compõem o procedimento, com base no conteúdo de ontologias de domínio. Ao utilizar o conhecimento disponível em uma ontologia de domínio, estamos reutilizando o conhecimento modelado e disponível num ADSOD (tal como definido no capítulo 5), mostrando também como ocorre a utilização do conhecimento do domínio numa atividade específica do processo padrão definido no capítulo 3, o projeto conceitual.

²⁸ A estética é fundamental para que um diagrama possa ser entendido.

²⁹ Sempre existe um relacionamento principal em todo o domínio que funciona como centro do negócio.

Tabela 4.2 – Realização do procedimento para modelagem conceitual com base no conteúdo de uma ontologia de domínio

Atividade	Utilização do conhecimento de Ontologia
I. Identificação das Coisas	Utilizar o conteúdo da ontologia (<i>árvores conceituais</i> e <i>dicionário de conceitos</i>) para identificar, visualizar e entender as coisas que serão traduzidas para as entidades que compõem o MER.
II. Dicionarização das Coisas Identificadas	Definir claramente cada uma das coisas modeladas, referenciando-as no dicionário de dados, através da utilização das definições conceituais disponíveis na ontologia, acrescentando as informações estruturais importantes no contexto do MER.
III. Modelagem dos Atributos das Coisas	Identificar as características próprias dos objetos, traduzindo-os para os atributos das coisas, utilizando o conhecimento disponível na ontologia na forma de <i>dicionário de conceitos</i> e <i>dicionário das propriedades dos conceitos</i> para obter o entendimento dos atributos identificados.
IV. Dicionarização dos Atributos das Coisas	Caracterizar completamente os atributos das coisas identificadas, com base nas informações disponíveis na ontologia na forma de <i>dicionário das propriedades dos conceitos</i> e <i>axiomas correspondentes a domínios de valores</i> , acrescentando as informações estruturais importantes na composição do MER.
V. Modelagem das Estruturas de Generalização	Identificar as propriedades das coisas que caracterizam a existência de estruturas de Generalização, utilizando as <i>árvores conceituais</i> , o <i>dicionário de conceitos</i> e os <i>axiomas epistemológicos</i> disponíveis na ontologia.
VI. Modelagem dos Relacionamentos entre as Coisas	Identificar, visualizar e entender os relacionamentos existentes entre as coisas, utilizando as relações disponíveis nas <i>árvores conceituais</i> ou expressos através de <i>axiomas</i> disponíveis na ontologia.
VII. Modelagem das Estruturas de Agregação	Identificar a existência de relacionamento dependente da existência prévia de outro relacionamento, com base nas <i>árvores conceituais</i> , no <i>dicionário de conceitos</i> e nos <i>axiomas</i> disponíveis na ontologia.
VIII. Dicionarização de Relacionamentos e seus Atributos	Caracterizar completamente os relacionamentos, utilizando as informações disponíveis na ontologia na forma de <i>dicionário de conceitos</i> e <i>dicionário das propriedades dos conceitos</i> para obter o entendimento dos relacionamentos identificados, acrescentando as informações estruturais importantes para a composição do MER.
IX. Representação no Diagrama	Construir o modelo inicial a partir da identificação dos construtores (atividades I a VIII) construindo o DER. Esta atividade é apoiada pelo mapeamento dos construtores ontológicos e da modelagem conceitual mostrada na seção 3.4.3.
X. Refinamento do Modelo Conceitual	Refinar os construtores (entidades, relacionamentos, atributos), com base no conteúdo da ontologia.

Para a execução das atividades (I a X) o projetista deve utilizar o conhecimento disponível na forma de ontologia de domínio. Este conhecimento permite agilizar a execução das atividades, possibilitando que a análise das estruturas dos dados necessários à construção de uma base de dados seja feita de forma mais fácil devido à disponibilidade do conhecimento para uso imediato.

Mas, ressaltamos que a modelagem conceitual é uma atividade totalmente dependente do projetista, que tem a inteira responsabilidade pelo processo de entendimento e transformação dos requisitos no projeto conceitual usando a ontologia disponível. Assim, o mesmo deve estar atento aos elementos da ontologia que estão ligados unicamente a funções de comunicação e entendimento sobre o domínio.

4.4 - Mapeamento dos construtores da ontologia para os construtores da modelagem conceitual

Conforme já visto no capítulo 2, ontologias são teorias de conteúdo sobre *tipos de objetos*³⁰, *propriedades de objetos* e *relações* possíveis entre objetos no conhecimento de um domínio específico. Elas provêm termos potenciais para descrição do conhecimento sobre o domínio. Ou seja, ontologia é um vocabulário de representação dos conceitos que os termos do vocabulário pretendem capturar, freqüentemente especializado para um domínio (CHANDRASEKARAN *et al.*, 1999).

No contexto do Projeto TABA (ROCHA *et al.*, 1990), ontologias são construídas e representadas por uma linguagem gráfica chamada LINGO (FALBO, 1998), que é basicamente, uma linguagem para se falar de ontologias e, portanto, incorpora uma meta-ontologia. As primitivas básicas capazes de representar a conceituação de um domínio são: conceitos e relações. No anexo 3, encontra-se um breve resumo sobre a linguagem LINGO e seus principais componentes.

O modelo ER, tal como abordado no capítulo 2, também pretende capturar o conhecimento sobre um domínio, descrevendo os conceitos do mundo real com *entidades* e *relacionamentos* entre as entidades. É expresso graficamente através do diagrama entidade-relacionamento (DER).

Nesta seção, estabelecemos a relação de equivalência entre os construtores básicos de ontologias com os construtores da modelagem conceitual baseada na abordagem entidade-relacionamento. Este relacionamento foi estabelecido com base nos trabalhos de WAND (1996), PARSONS e WAND (1997), WAND *et al.* (1999) e KOLP e ZIMÁNYI (2000). Mostramos ainda, a equivalência através de exemplos, dos construtores gráficos da ontologia, disponíveis em LINGO (representados nas figuras a seguir pelos construtores no lado esquerdo), com os construtores gráficos da modelagem conceitual disponíveis no DER (representados nas figuras a seguir pelos construtores no lado direito).

³⁰ Objeto aqui é utilizado significando um conceito sobre o domínio. É a representação de um objeto pelo pensamento, nas suas características gerais.

Construtores

O construtor fundamental necessário para modelar um domínio é a representação de um *objeto*. Identificar os objetos (conceitos) em um domínio é um pré-requisito para a elicitação da semântica do domínio. O construtor genérico da modelagem conceitual usada para representar um objeto é uma instância.

Na abordagem de modelagem ER, uma instância é representada como uma *entidade*, que descreve conceitos do mundo real. Uma entidade é um objeto do domínio com existência independente que pode ser identificada distintamente e deve ser utilizada somente para representar objetos.

Além disso, todo objeto do mundo real possui propriedades. Cada propriedade pode ser intrínseca a um objeto e é representada por atributos, que constituem a percepção das propriedades pelas pessoas.

No modelo ontológico, objetos podem ser ligados somente por *propriedades mútuas ou relacionais*³¹ (usualmente refletindo interações) e não por outros objetos. Como propriedades mútuas não são objetos, elas não devem ser representadas por construtores que representam objetos. *Relacionamentos*, na modelagem ER, representam as propriedades mútuas dos objetos. O mesmo construtor deve ser usado para representar um relacionamento binário e um relacionamento de maior ordem. Ambos os relacionamentos refletem atributos representando propriedades mútuas.

A figura 4.2 abaixo mostra um exemplo de uma *relação binária* entre os conceitos de Pessoa e Empresa, indicando que Empresas contratam Pessoas, e que Pessoas são contratadas por Empresas, no contexto de uma organização. A relação contrato é a representação de uma propriedade mútua de Pessoa e Empresa.

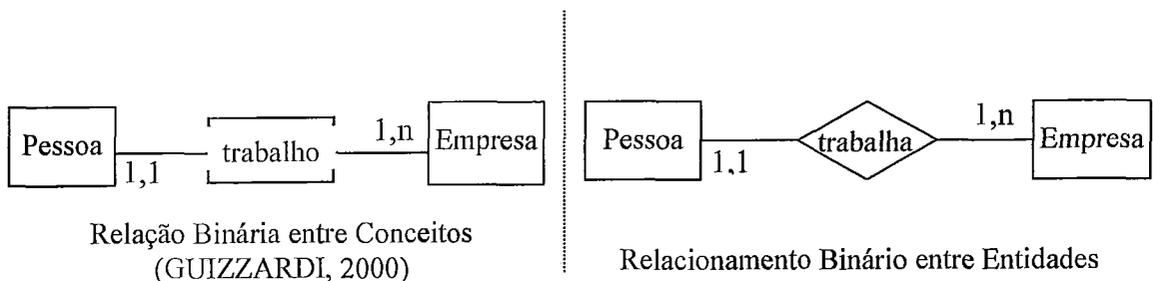


Figura 4.2 – Equivalência entre relação binária ontológica e relacionamento binário do DER

³¹ São propriedades que dependem de dois ou mais objetos.

Similarmente, no DER os *conceitos* são equivalentes às *entidades* que participam do *relacionamento*, que por sua vez é equivalente à *relação* com as mesmas *cardinalidades*, que são usadas para mostrar quantas instâncias de um conceito podem participar da relação e significam, que uma Empresa pode contratar uma ou várias Pessoas, mas que uma Pessoa só pode ser contratada por uma Empresa.

Relações também podem ser de ordem superior, tais como relações ternárias, como mostrado na figura 4.3.

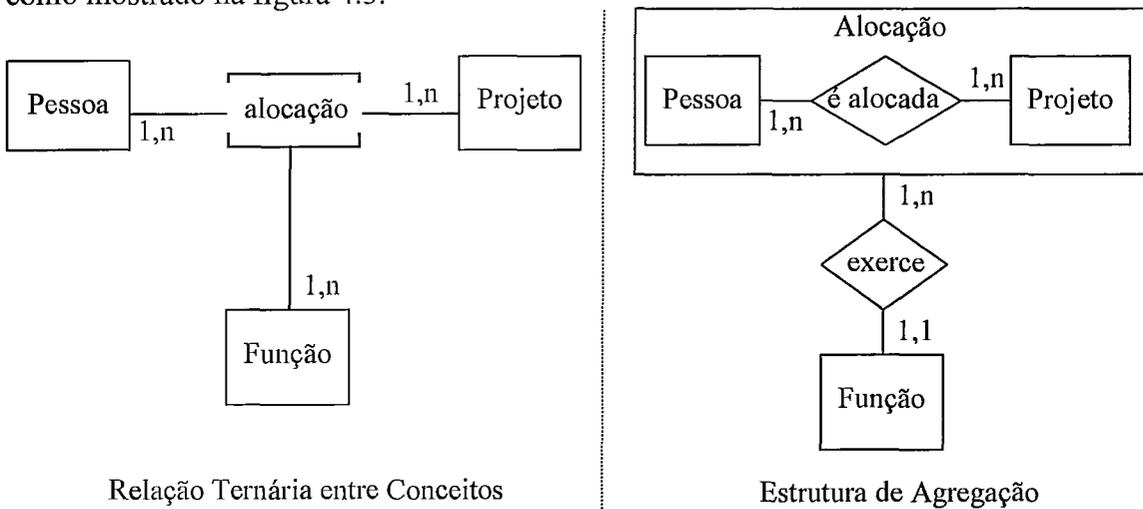


Figura 4.3 – Equivalência entre relação ternária ontológica e estrutura de agregação do DER

Neste caso, o mapeamento da *relação ternária* foi feito para a estrutura de *agregação* no DER, uma vez que não foi prevista a possibilidade de associar uma entidade com um relacionamento na modelagem ER. Então, o projetista terá que analisar com base no que está definido na ontologia do domínio e nas regras de negócio, com qual das duas entidades a terceira entidade está relacionada, observando a cardinalidade do relacionamento.

Se Função for associada à Pessoa, não teremos a informação sobre em qual Projeto ela exerce a Função; por outro lado, se Função estivesse associada a Projeto, não saberíamos que Pessoa exerce a Função. Neste caso, uma Pessoa pode exercer diversas Funções e uma Função pode ser exercida por várias Pessoas. A Pessoa pode estar alocada a vários Projetos, mas em cada um dos Projetos cada Pessoa só pode exercer uma Função. O que não impede, que uma Pessoa acabe por estar alocada a diversas Funções em diferentes Projetos.

As *relações entre instâncias* de um mesmo conceito também são válidas (figura 4.4). Em algumas relações (como essa), é importante o uso de *papéis*, cujo principal objetivo é deixar claro as responsabilidades de cada uma das partes envolvidas na relação.

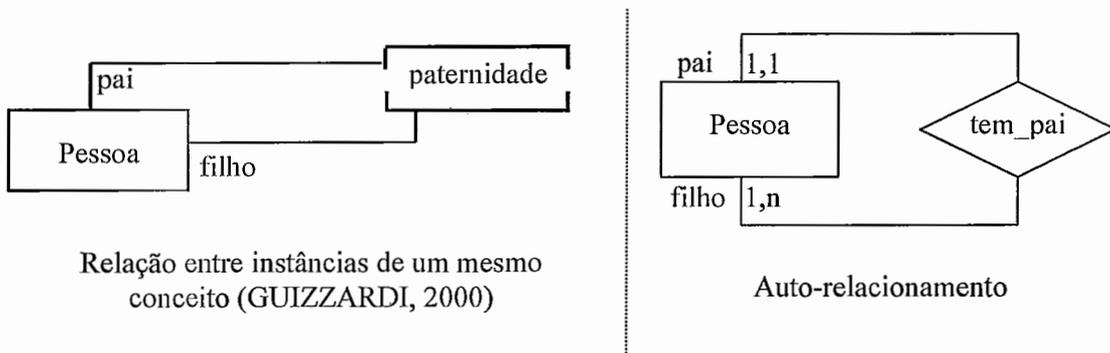


Figura 4.4 – Equivalência da relação entre instâncias ontológicas e o auto-relacionamento do DER

Tanto em LINGO como no DER, temos que a própria interpretação da *relação / relacionamento*, mostra as instâncias do *conceito* e *ocorrências da entidade* assumindo papéis diferentes, conforme seu posicionamento no relacionamento.

Composição é um conceito ontológico fundamental que estabelece a noção de que um objeto existe a partir da composição de outros objetos; isto é, *tem partes*. Cada um desses objetos, por sua vez, é um componente ou parte do *objeto composto*. Além disso, objetos compostos têm pelo menos uma *propriedade emergente*³² de interesse.

Ontologicamente, a composição e o relacionamento todo-parte, são uma composição e seus componentes são objetos. Composição é um *objeto complexo* e relacionamento todo-parte é *objeto simples* ou *composto*. A composição e seus componentes são relacionados via uma propriedade mútua obrigatória (WAND *et al.*, 1999).

Em DER, um construtor de relacionamento é freqüentemente usado para mostrar que uma composição e seus componentes são relacionados. Em ontologia, o relacionamento todo-parte é um tipo de atributo mútuo obrigatório, mas é diferente de um relacionamento de conexão, que também é um tipo de atributo mútuo obrigatório.

Tanto objetos simples quanto compostos, devem ser representados usando o mesmo construtor entidade. Um objeto composto deve possuir propriedades emergentes e deve, ser mostrado com atributos ou relacionamentos não possuídos por nenhum de seus componentes.

³² É uma propriedade não possuída por nenhuma das partes componentes.

LINGO, conforme mostrado no anexo 3, também dispõe de uma estrutura que representa a *relação todo-parte* (figura 4.5), que importa e aplica uma teoria abstrata de composição de elementos ao conteúdo da ontologia em desenvolvimento. As relações individuais entre o todo e cada uma de suas partes também podem ser vistas como *relações binárias* com os respectivos papéis de *todo* e *parte*. Nesta relação de composição, as partes são exclusivas de um único todo.

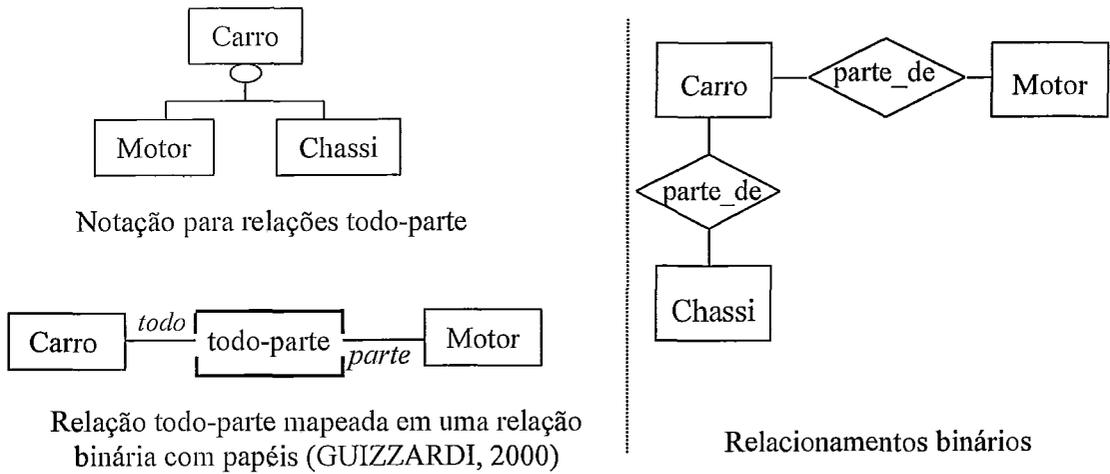


Figura 4.5 – Equivalência da estrutura todo-parte ontológica e relacionamentos binários do DER

Não existe no DER uma estrutura que corresponda diretamente à estrutura todo-parte de LINGO. Desta forma, o mapeamento é feito de forma que cada uma das partes esteja relacionada com o todo através de *relacionamentos binários*.

Especialização é o processo de formação de um objeto a partir de outro objeto, pela adição de propriedades ao conjunto de propriedades possuídas pelo objeto final. *Generalização* é o processo de formação de um objeto a partir de outros objetos, pela identificação de propriedades que são comuns a estes últimos objetos (subtração de propriedades que não são comuns a eles).

No DER, generalização é um mecanismo de abstração, envolvendo duas ou mais entidades, chamadas respectivamente, *entidade genérica* e *entidades especializadas*. Uma entidade genérica pode ter várias entidades especializadas e vice-versa. Além disso, cada ocorrência da entidade especializada possui, além de suas próprias propriedades (atributos, relacionamentos e generalizações/especializações), também as propriedades da ocorrência da entidade genérica correspondente.

Em LINGO, existe uma estrutura que corresponde a uma relação *subtipo-de/supertipo-de*, conforme mostra a figura 4.6, onde a associação ocorre entre conceitos e não entre suas instâncias.

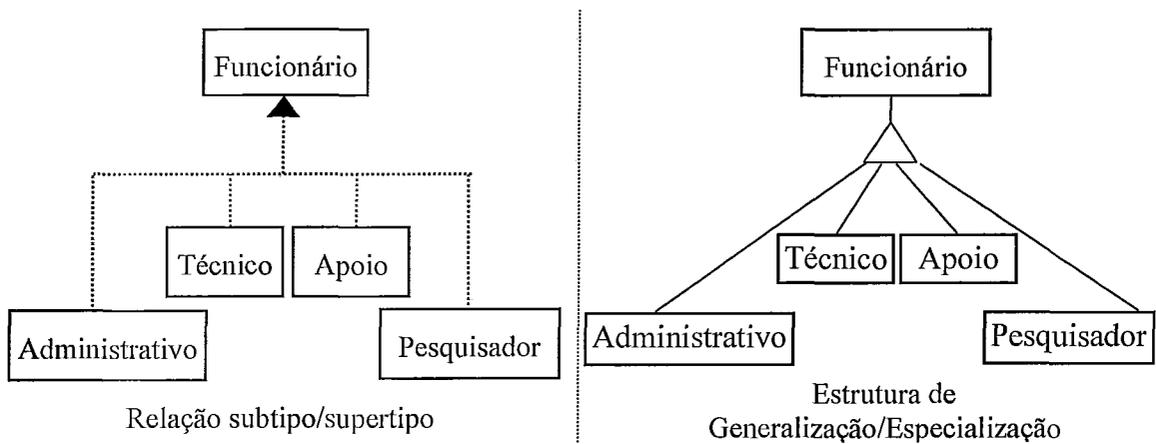
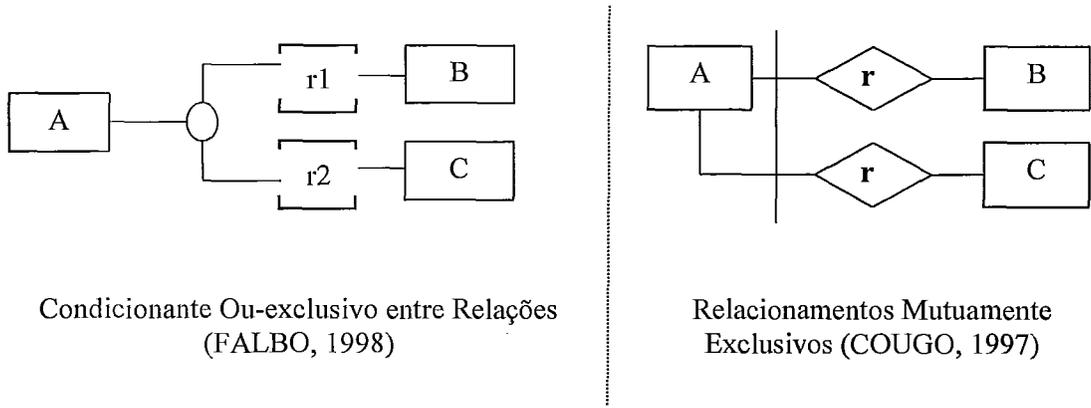


Figura 4.6 – Equivalência da estrutura subtipo-de/supertipo-de ontológica e a estrutura de generalização/especialização do DER

Através do uso da estrutura de generalização/especialização no DER, pode ser observado que a entidade *Funcionário* subdivide-se em outras quatro. Visualizando isso, não como um conjunto de cinco entidades distintas, mas como o triângulo sendo um prisma no qual, ao incidir a entidade *Funcionário*, pode ser visto refletir do outro lado seus quatro espectros distintos: *Administrativo*, *Técnico*, *Apoio*, *Pesquisador*. De um dos lados do triângulo está o agrupamento total dos elementos com todas as características que lhes são comuns, e do outro, os subconjuntos possíveis obtidos a partir da especialização do todo.

Além destes construtores básicos, existe ainda o conceito de atributo no modelo ontológico, que é a representação de uma propriedade de um objeto no mundo real e que corresponde ao construtor atributo no DER. Mas, para fins de clareza do diagrama, os atributos, com suas respectivas cardinalidades, não são representados diagramaticamente.

LINGO prevê ainda as estruturas para expressar condicionantes entre relações, conforme as figuras 4.7 e 4.8. No DER, a representação gráfica de relacionamentos condicionais demonstra quando um ou outro elemento permite a ausência ou obrigatoriedade de associações.

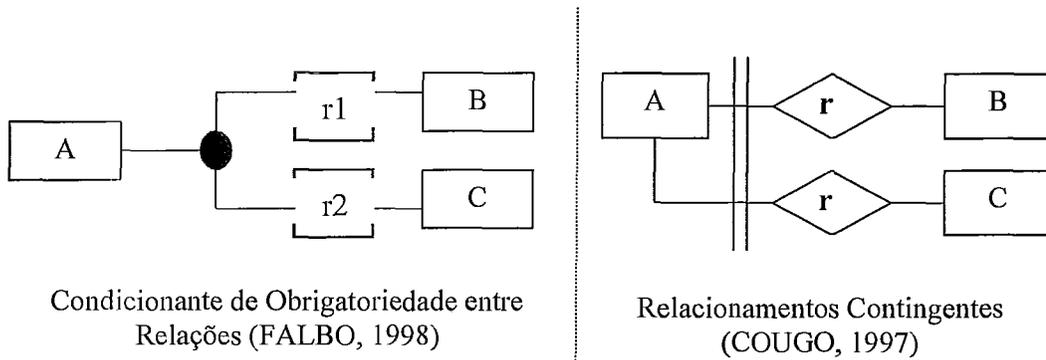


Condicionante Ou-exclusivo entre Relações (FALBO, 1998)

Relacionamentos Mutuamente Exclusivos (COUGO, 1997)

Figura 4.7 – Equivalência da estrutura condicionante Ou-exclusivo ontológica e a estrutura de relacionamentos Mutuamente exclusivos do DER

Neste caso, dois relacionamentos estabelecidos entre os objetos de nosso modelo podem, em função de características próprias do domínio, caracterizar-se por serem mutuamente exclusivos, implicando que, se a associação de um elemento for estabelecida através de um dos relacionamentos, não poderá ser estabelecida pelo outro.



Condicionante de Obrigatoriedade entre Relações (FALBO, 1998)

Relacionamentos Contingentes (COUGO, 1997)

Figura 4.8 – Equivalência da estrutura condicionante de Obrigatoriedade ontológica e a estrutura de relacionamentos Contingentes do DER

Aqui existe a dependência entre os relacionamentos, impondo o estabelecimento simultâneo de associações entre os vários elementos envolvidos. Isto define que, como regra de negócio, mais de um relacionamento deve ocorrer em um mesmo instante.

A tabela 4.3 resume o mapeamento que estabelece as diretivas de correspondência entre os conceitos dos construtores ontológicos, que representam o conhecimento do domínio, e os construtores do modelo ER, que descrevem as estruturas dos dados das bases de dados. Esse mapeamento foi feito baseado nos trabalhos de WAND (WAND, 1996; PARSON e WAND, 1997; WAND *et al.*, 1999).

Tanto ontologias de domínio quanto modelos ER, estão em busca da representação da semântica do domínio modelado, de forma a representar o mais fiel possível, o ambiente observado na forma de um modelo conceitual. Porém, os níveis da representação do conhecimento são diferentes (FALBO, 1998). A ontologia representa o conhecimento no nível da significação (ontológico) e o MER, no nível de estruturação (epistemológico).

Tabela 4.3 – Equivalência entre Construtores Ontológicos e Construtores do MER

Construtores ontológicos	Construtores modelo ER
Objeto (Conceito) simples ou Objeto composto	Entidade
Propriedade intrínseca	Atributo de uma entidade
Propriedade mútua ou relacional	Relacionamento binário ou n-ário
Composição/Relação todo-parte	Relacionamento parte-de
Relação Entre Instâncias	Auto-relacionamento
Generalização/Especialização (relação entre conceitos)	Generalização/Especialização
Cardinalidade Mínima e Máxima	Cardinalidade Mínima e Máxima

Além dos construtores básicos, uma das características principais de ontologias formais, é o conjunto de suposições explícitas que dizem respeito ao significado planejado dos termos do vocabulário específico. Ou seja, axiomas apropriados são adicionados para expressar outros relacionamentos entre os objetos ou para restringir sua interpretação (FALBO, 1998).

Desta forma, todas as notações de LINGO foram projetadas para serem capazes de capturar certos axiomas de forma implícita (ver Anexo 3), possuindo uma semântica equivalente a um conjunto de axiomas, de modo que ao utilizar estas notações, descrevemos de fato o conjunto de axiomas que elas representam. Estes axiomas estabelecem significados formais para alguns termos do vocabulário do domínio, permitindo restringir a interpretação da teoria definida pela ontologia. A tabela 4.4 mostra uma possível utilização dos axiomas disponíveis na ontologia, na construção do MER.

Tabela 4.4 – Uso dos Axiomas Formais da Ontologia na Modelagem Conceitual

Axiomas	Uso na modelagem conceitual
Epistemológico ³³	Utilizado para verificar a consistência do MER, através das estruturas usadas na sua construção.
Ontológico ³⁴	Utilizado para verificar a consistência do MER, podendo gerar novas entidades e novos relacionamentos entre as mesmas, ou apenas novos relacionamentos entre as entidades já existentes.
De Consolidação ³⁵	Utilizado para verificar a consistência do MER e complementar o dicionário de dados.

Por exemplo, os axiomas referentes ao uso da estrutura representada na figura 4.2 são:

Cardinalidade mínima 1	$(\forall e) (\text{Empresa}(e) \rightarrow (\exists p) (\text{contrato}(e,p)))$ $(\forall p) (\text{Pessoa}(p) \rightarrow (\exists e) (\text{contrato}(p,e)))$
Cardinalidade máxima 1	$(\forall e,p1,p2) (\text{contrato}(e,p1) \wedge \text{contrato}(e,p2) \rightarrow p1=p2)$
Cardinalidade máxima N	$(\forall e1,e2,p1,p2) (\text{contrato}(e1,p1,p2) \wedge \text{contrato}(e2,p1,p2) \rightarrow e1=e2 \wedge p1 \neq p2)$

Estes axiomas são utilizados para o entendimento e a verificação do uso correto da cardinalidade, se está de acordo com a realidade (regras do domínio) que se está modelando para a construção da base de dados.

Não existe um mapeamento direto destes axiomas para o DER. Mas eles permitem que o diagrama seja melhor elaborado, permitindo a verificação da consistência de determinados usos, visando o mapeamento fiel da realidade observada. Além disso, esses axiomas podem ser utilizados diretamente no projeto do dicionário de dados (atividades de dicionarização - II, IV e VIII - do procedimento definido na seção 4.3.1).

4.5 – Considerações Finais

O MER é uma representação do conhecimento no nível epistemológico, que mostra a existência de conceitos no domínio observado, e os relacionamentos entre instâncias desses conceitos, modelando a estrutura dos conceitos e seus relacionamentos, sendo insuficiente para representar o conhecimento no nível da significação (ontológico). Para a representação do conhecimento nesse nível, necessitamos da existência de ontologias formais, cujas primitivas satisfaçam postulados formais de significação, restringindo a interpretação de uma teoria (FALBO, 1998).

³³ Axiomas derivados simplesmente da estrutura dos conceitos e não de seus significados particulares e que estão embutidos na representação gráfica de LINGO (FALBO, 1998).

³⁴ Axiomas que representam restrições entre os conceitos que compõem a ontologia (FALBO, 1998).

³⁵ Axiomas que têm por objetivo verificar a coerência das informações existentes e não derivam novas informações (FALBO, 1998).

Por este motivo, neste capítulo, verificamos como o projeto conceitual é realizado com o apoio de ontologias de domínio, através de seu uso nas etapas do processo geral de construção de bases de dados e do mapeamento dos construtores disponíveis nas linguagens gráficas utilizadas (LINGO e DER).

Também foi organizado um procedimento com base na abordagem entidade-relacionamento (CHEN, 1976), voltado para a atividade *projeto conceitual* da fase *desenvolvimento do processo padrão* definido no capítulo 3, que traz como resultado a definição das necessidades de armazenamento de dados independente de implementação, na forma de um modelo conceitual. Mostramos também, como utilizar ontologias de domínio no apoio ao desenvolvimento das atividades organizadas neste procedimento.

No próximo capítulo mostraremos a construção do ADSOD *Insecta*, para o domínio de Entomologia, onde é utilizado o processo definido no capítulo 3. Esta classe de ADS, definido por OLIVEIRA (1999), tem como principal característica o uso do conhecimento do domínio, modelado através de ontologias de domínio, ao longo do processo de desenvolvimento.

Capítulo 5

INSECTA: Um Ambiente de Desenvolvimento de Software Orientado ao Domínio de Entomologia

*Este capítulo apresenta inicialmente, a Estação TABA e Ambientes de Desenvolvimento de Software Orientados a Domínio (ADSOD), enfatizando sua origem, definição e principais características. Mostra ainda, a definição da Teoria do Domínio para Entomologia, tendo Ontologias como base, focalizando cada uma das atividades desenvolvidas, além da sua utilização como componente de um Ambiente de Desenvolvimento de Software Orientado a Domínio chamado *Insecta*, definido e instanciado a partir da infra-estrutura da Estação TABA. Neste capítulo, também é apresentada a modelagem conceitual de uma pequena base de dados de pesquisa para o domínio de Entomologia, feita utilizando a modelagem conceitual de dados com base em ontologias de domínio, definida no capítulo anterior.*

5.1 – A Estação TABA e Ambientes de Desenvolvimento de Software Orientados a Domínio

Iniciamos esta seção apresentando a definição e características básicas de ambientes de desenvolvimento de software. Em seguida descrevemos a Estação TABA no que se refere às suas funcionalidades e ferramentas disponíveis que apóiam a definição e geração desses ambientes.

5.1.1 – Ambiente de Desenvolvimento de Software

Ambiente de Desenvolvimento de Software (ADS) é um sistema computacional que provê suporte ao desenvolvimento, reparo e melhorias em software e para o gerenciamento e controle destas atividades (MOURA e ROCHA, 1992), objetivando o apoio automatizado ao desenvolvimento de produtos de software, de forma integrada e controlada (OLIVEIRA *et al.* 2000a).

Ambientes de Desenvolvimento de Software possuem estruturas de organização que propõem diferentes componentes para constituir os ambientes e um conjunto de funcionalidades para prover o apoio desejado aos desenvolvedores de software de forma integrada. Além disso, a arquitetura de ADS deve permitir que as ferramentas possam ser conectadas ao ambiente e que o ambiente forneça suporte metodológico ao desenvolvimento (TRAVASSOS, 1994; TRAVASSOS e ROCHA, 1994a e 1994b; OLIVEIRA, 1999).

Acreditando que o uso do conhecimento do domínio pode contribuir para a redução de dificuldades existentes no processo de desenvolvimento decorrentes da falta de conhecimento de conceitos do domínio pelos desenvolvedores, OLIVEIRA (1999) definiu Ambientes de Desenvolvimento de Software Orientados a Domínio (ADSOD).

Esses ambientes tornam disponível o conhecimento sobre o domínio da aplicação numa representação simbólica utilizando ontologias do domínio e a identificação de tarefas possíveis de serem realizadas no domínio em questão. ADSOD pode ser visto, então, como uma evolução dos ambientes de desenvolvimento de software tradicionais, cujo aspecto central é a introdução e uso do conhecimento do domínio no ADS, provendo apoio aos desenvolvedores de software, em domínios específicos, ao longo de todo o processo de desenvolvimento (OLIVEIRA *et al.* 2000a; SILVA, 2000).

O domínio é uma área de aplicação na qual vários produtos de software serão desenvolvidos. O uso do conhecimento do domínio durante o desenvolvimento de software tende a tornar o processo de entendimento do problema mais fácil e agradável tanto para os desenvolvedores, quanto para os especialistas do domínio que muitas vezes não são profissionais da área de informática (FISCHER, 1994).

Dessa forma, a principal característica de um ADSOD, é a existência do conhecimento do domínio explicitamente definido no ambiente, apoiando o desenvolvimento de software que utiliza esse conhecimento. Para atender este aspecto, OLIVEIRA (1999) definiu a Teoria do Domínio e introduziu uma atividade nova no processo de desenvolvimento de forma a estabelecer o apoio ao desenvolvimento através do uso do conhecimento.

Teoria do Domínio

A Teoria do Domínio (OLIVEIRA *et al.*, 1999b) é um modelo composto basicamente da organização de conceitos, propriedades e restrições do domínio através de ontologias do domínio e do mapeamento desses conceitos com as tarefas identificadas para o domínio considerado. Pode ser dividida em sub-teorias, uma vez que, identificar e definir os conceitos de

um domínio pode ser um processo muito longo devido um domínio ser geralmente muito amplo e rico em detalhes.

Cada sub-teoria é uma ontologia sobre uma parte específica do domínio que se refere a um mesmo contexto dentro do domínio considerado e que deve estar relacionada com as demais sub-teorias identificadas, através da definição e descrição de restrições axiomáticas.

Além disso, cada sub-teoria é ainda mapeada com potenciais tarefas pertinentes ao domínio. Uma tarefa é geralmente independente do domínio, sendo genérica e aplicável a diferentes domínios. Cada uma dessas tarefas possui características próprias que serão adaptadas a qualquer domínio escolhido.

A definição da Teoria do Domínio é executada seguindo o processo de construção proposto por OLIVEIRA (1999) e utilizado por SILVA (2000), conforme ilustra a figura 5.1.

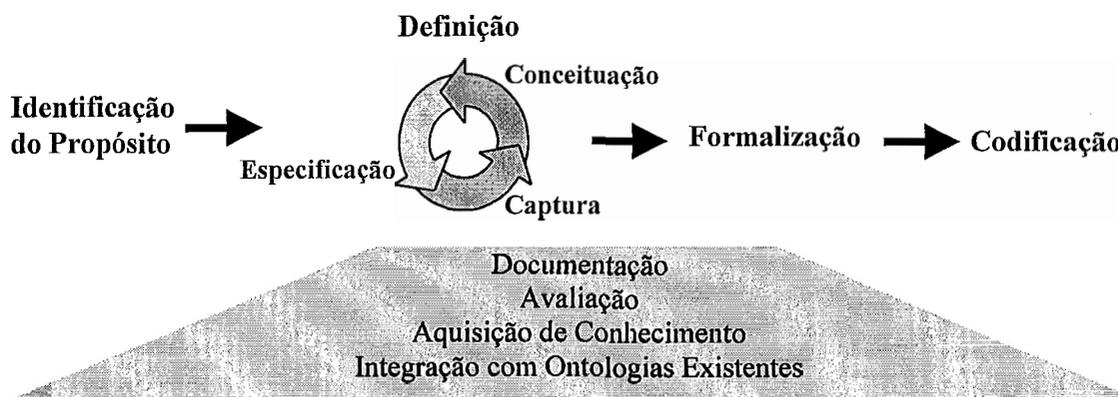


Figura 5.1 – Processo de Desenvolvimento da Teoria do Domínio (OLIVEIRA, 1999).

5.1.2 – A Estação TABA

O projeto TABA, concebido e em desenvolvimento na COPPE/UFRJ desde 1990 (ROCHA *et al.*, 1990), tem como objetivo a criação de um meta-ambiente capaz de instanciar ADS configuráveis a partir dos requisitos de diferentes projetos, domínios de aplicação e tecnologias. A inspiração para a concepção do projeto deveu-se à constatação do fato de que, os domínios de aplicação e projetos específicos possuem características próprias e estas características devem estar presentes de forma customizada nos ADS utilizados na gerência e controle do desenvolvimento de aplicações (MACHADO, 2000; OLIVEIRA *et al.*, 2000a; SILVA, 2000).

A Estação TABA é, então, um meta-ambiente de desenvolvimento de software que permite a especificação e instanciação de ambientes de desenvolvimento de software de acordo com as particularidades de processos de software específicos, domínios de aplicação e tecnologias (FALBO *et al.*, 1999a). ADSOD, tal como estabelecido por OLIVEIRA (1999) são instanciados, na Estação TABA, a partir da definição de um processo e de uma Teoria do Domínio.

Para apoiar as atividades de definição de processos e Teoria do Domínio, de forma a permitir a instanciação de ADS e ADSOD particulares a projetos específicos, a Estação TABA provê um conjunto de ferramentas, dentre as quais podemos destacar as ferramentas para *Definição do Processo de Desenvolvimento*, a ferramenta para *Definição da Teoria do Domínio e de Tarefas*. Além disso, um ADSOD instanciado pela Estação TABA provê um conjunto de ferramentas que apóiam o entendimento do domínio ao longo do processo de desenvolvimento.

Ferramentas para Definição do Processo de Desenvolvimento foram desenvolvidas com o objetivo de permitir a especialização e instanciação de processos a partir de um processo padrão:

- a) **EDIT-PRO** (ZLOT e SANTOS, 1999) é uma ferramenta para a introdução de um processo previamente definido pelo engenheiro de software.
- b) **ASSIST-PRO** (FALBO *et al.*, 1999b) é uma ferramenta que fornece assistência inteligente na escolha do modelo de ciclo de vida e na descrição de um processo específico para o ambiente a ser instanciado.
- c) **DEF-PRO** (MACHADO, 2000) é uma ferramenta mais completa que apoia desde a definição do processo padrão até a definição de um processo instanciado, considerando não apenas a norma NBR ISO/IEC 12207 (1998) e modelos de maturidade CMM (PAULK *et al.*, 1997) e SPICE (EMAN *et al.*, 1998), mas também características da organização, tipo de software, tecnologia de desenvolvimento e o tipo de ambiente que se deseja instanciar (orientados a domínio ou não).

A *Ferramenta para Definição da Teoria do Domínio (EDITED)* foi desenvolvida com o objetivo de apoiar a definição da Teoria de Domínio e permitir sua evolução. Essa ferramenta é utilizada para permitir a introdução do conhecimento do domínio na Estação TABA.

A Estação TABA também provê **EDITAR** (OLIVEIRA, 1999), uma ferramenta para edição de tarefas, feita através de simples documentos que descrevem as principais características da tarefa, fornecendo referências bibliográficas para a mesma.

Ferramentas para apoio ao Entendimento do Domínio, são ferramentas desenvolvidas com o objetivo de apoiar o acesso ao conhecimento e facilitar o entendimento do domínio (SILVA, 2000). Essas ferramentas englobam:

- a) **REGCON**, uma ferramenta para apoiar o registro de uso de conceitos da Teoria do Domínio, permitindo ao desenvolvedor identificar e relacionar, quais conceitos da Teoria do Domínio, foram utilizados em determinado projeto.
- b) **GENESIS**, uma ferramenta para apoiar o registro de instâncias de conceitos da Teoria do Domínio, permitindo ao desenvolvedor o registro de exemplos dos conceitos da Teoria do Domínio que foram utilizados em determinado projeto.
- c) **NAVEGUE**, um assistente de aprendizado do domínio, que permite ao desenvolvedor a pesquisa de conceitos do domínio, obtendo informações sobre o significado de conceitos, suas características específicas, situações de uso em sistemas já desenvolvidos e uma relação de exemplos desses conceitos.

Após esta breve apresentação, podemos verificar que a Estação TABA possui facilidades para a definição de processos, métodos e ferramentas *CASE* que serão utilizadas no processo de desenvolvimento, além de permitir a inclusão de conhecimento específico do domínio no qual o desenvolvedor irá executar suas atividades. Todos estes elementos estão organizados em um modelo de integração do ambiente, permitindo que diferentes ambientes sejam definidos e instanciados (OLIVEIRA *et al.*, 2000a), apoiando o gerenciamento e controle das atividades que serão executadas.

Na seção seguinte, apresentamos a modelagem do conhecimento do domínio, através de uma Teoria do Domínio para o domínio de Entomologia, como uma das etapas da construção do ADSOD *Insecta*.

5.2 – Teoria do Domínio para Entomologia³⁵

Conforme abordagem da seção anterior, ADSOD tem como objetivo auxiliar os desenvolvedores de software no entendimento do domínio e no desenvolvimento de software, orientados por descrições específicas do conhecimento do domínio e organizadas em um modelo denominado Teoria do Domínio.

A Teoria do Domínio descrita nesta seção apresenta o conhecimento básico de Entomologia, necessário ao desenvolvimento das atividades do entomólogo³⁶. Apesar de abrangente, não é uma Teoria completa, por tratar de uma área de conhecimento muito ampla. A partir de sua utilização a mesma deverá evoluir, incluindo novos conceitos e relações que retratem o novo conhecimento identificado.

A construção da Teoria do Domínio para Entomologia teve como motivação, uma situação real existente na Embrapa Amazônia Oriental³⁷, onde existe a necessidade de construção de bases de dados temáticas (diferentes domínios) como um dos produtos principais de qualquer instituição de ciência e tecnologia (C&T), e deve-se considerar na sua construção, o conhecimento do domínio. A Entomologia foi o domínio inicialmente selecionado.

A primeira atividade realizada foi, a identificação do propósito da Teoria. A partir deste propósito, houve a interação com três especialistas da Embrapa Amazônia Oriental e um especialista do Museu Paraense Emílio Goeldi³⁸. A partir do conhecimento dos mesmos foram enumeradas questões de competência geral e especificado um modelo amplo contendo este conhecimento inicial. Além destes especialistas, também foi consultada literatura específica (GALLO *et al.*, 1988; BORROR e DeLONG, 1984; STORER, 1984; LARA, 1992; PAPAVERO, 1983; SILVEIRA NETO *et al.*, 1976).

As questões de competência geral serviram de base para a próxima etapa, a captura do domínio. Nesta etapa, foi realizada a elaboração de modelos usando LINGO (FALBO, 1998), conforme notação descrita no Anexo 3, e a descrição de conceitos e axiomas utilizando linguagem natural.

³⁵ Entomologia é um ramo da zoologia que estuda os insetos sob todos os seus aspectos

³⁶ São todos os profissionais especialistas no estudo dos insetos sob todos os seus aspectos.

³⁷ Embrapa Amazônia Oriental é uma unidade da Empresa Brasileira de Pesquisa Agropecuária, que atua na Amazônia Oriental há sessenta e dois anos, com sede em Belém-PA.

³⁸ O Museu Paraense Emílio Goeldi é uma instituição de pesquisa que produz e difunde conhecimentos e acervos científicos sobre sistemas naturais e sócio-econômicos relacionados à Amazônia, com sede em Belém-PA.

Numa próxima iteração da etapa de definição da Teoria, um outro especialista do Museu Nacional³⁹ avaliou estes modelos, sugeriu algumas melhorias, e nova literatura específica foi consultada (SANTOS, 1982; SANTOS, 1985).

Nas subseções a seguir serão apresentados os resultados da execução de cada uma das etapas realizadas para a definição desta Teoria, segundo o processo utilizado (figura 5.1).

5.2.1. Identificação do Propósito

A execução desta atividade tem como resultado a descrição do objetivo da construção da Teoria do Domínio.

No contexto da Entomologia na Embrapa Amazônia Oriental, a definição de uma Teoria do Domínio visa apoiar a construção de bases de dados de pesquisa através da utilização explícita do conhecimento do domínio nas etapas da construção destas bases.

Esta definição deve estabelecer um vocabulário comum, permitindo um melhor entendimento do domínio e dos requisitos da base de dados a ser construída. É provável que o tempo despendido no entendimento do domínio por parte dos desenvolvedores, engenheiros de conhecimento, especialistas e usuários em geral seja menor, devendo facilitar ainda, a execução da atividade projeto conceitual, tal como definido na seção 3.2.

O termo entomologia tem sua origem nas palavras gregas *entomon* (inseto) e *logia* (estudo). O objetivo maior da entomologia é estudar as relações dos insetos com o homem, com as plantas e com outros animais. A entomologia usa seus conhecimentos para resolver problemas na agricultura, como os programas de controle integrado de pragas, assim como a criação e manejo de insetos benéficos, dentre outros. Para isso, a entomologia está ligada a diferentes áreas do conhecimento como, por exemplo, a genética, a fisiologia, a ecologia e a taxonomia.

Devido à ligação da entomologia com essas diferentes áreas, a Teoria do Domínio de entomologia aqui descrita, procura definir, hierarquizar e organizar não só os conceitos que se referem às características morfológicas básicas, biológicas e comportamentais dos insetos que são importantes para a sua identificação, manejo e controle, mas, também, os conceitos que caracterizam as principais atividades desenvolvidas pelos entomólogos.

³⁹ O Museu Nacional é uma instituição de ensino e pesquisa vinculado à UFRJ, com sede no Rio de Janeiro-RJ.

5.2.2. Definição

É a segunda e a mais longa das etapas do processo de construção de Teorias do Domínio (figura 5.1) e refere-se à modelagem do conhecimento do domínio propriamente dita. Consiste de três atividades distintas: especificação de requisitos, captura do domínio e conceituação. A especificação de requisitos implica na definição das questões de competência do domínio. A captura do domínio refere-se à identificação dos conceitos, relacionamentos, propriedades e restrições do domínio. A conceituação envolve a definição precisa e não ambígua para o domínio capturado.

Estas atividades são feitas iterativamente, de forma que os conceitos mais significativos são capturados inicialmente, procurando-se entendê-los e defini-los, depois, especifica-se os requisitos e retorna-se para uma captura e conceituação mais detalhadas. Esta iteratividade é necessária e definida em função do escopo do domínio a ser modelado e da experiência da equipe que está realizando a tarefa.

Após todas as definições dos conceitos, é realizada a definição das restrições axiomáticas para os conceitos e sub-teorias. Axiomas são definidos, em linguagem natural para validação pelos especialistas do domínio, visando estabelecer restrições nas relações entre conceitos ou entre propriedades (atributos) de um conceito.

Foram definidas, inicialmente, questões de competência geral com base no objetivo definido para a Teoria do Domínio. Dessa forma, a ênfase foi dada no que seria importante e útil para solucionar os diversos problemas, que têm na entomologia a base para a sua solução, e que deveriam estar disponíveis em bases de dados de pesquisa.

Um conjunto de conceitos foi então levantado a partir do conhecimento e do processo de trabalho dos entomólogos que colaboraram nesta definição. Esses conceitos contêm informações de diferentes níveis de abstração. Os conceitos semanticamente mais relacionados e em um mesmo nível de abstração deram origem à identificação das questões de competência geral descritas a seguir, que por sua vez originaram as diferentes sub-teorias que compõem a Teoria do Domínio (OLIVEIRA, 1999).

- i. Como se caracterizam os insetos ?
- ii. O que deve ser considerado na classificação de um inseto ?
- iii. Quais são as principais categorias taxonômicas dos insetos ?
- iv. Como se caracterizam as principais atividades dos entomólogos ?

Estas questões de competência definem as cinco sub-teorias que irão respondê-las:

- a) Anatomia de insetos: responde às questões i e ii. Contém os conceitos referentes à estrutura morfológica dos insetos.
- b) Biologia de insetos: responde às questões i e ii. Apresenta os conceitos sobre como é o ciclo de vida dos insetos.
- c) Ecologia de insetos: responde às questões i e ii. Aborda os conceitos básicos relacionados ao comportamento dos insetos.
- d) Classificação de insetos: responde à questão iii. Mostra os conceitos envolvidos na identificação dos insetos.
- e) Processo de trabalho do entomólogo: responde à questão iv. Apresenta conceitos envolvidos nas atividades desenvolvidas pelos entomólogos.

A figura 5.2 ilustra a relação entre estas sub-teorias. A anatomia de insetos constitui-se na base para a classificação de insetos. Já a biologia dos insetos, com seu ciclo de vida, exerce influência sobre a anatomia. A ecologia dos insetos nos apresenta como os insetos se relacionam entre si e com todo o meio ambiente, e juntamente com a anatomia e a biologia de insetos constituem fontes de informação que auxiliam a execução do processo de trabalho do entomólogo, além de ter uma grande importância na classificação de insetos, um dos principais resultados do processo de trabalho do entomólogo.

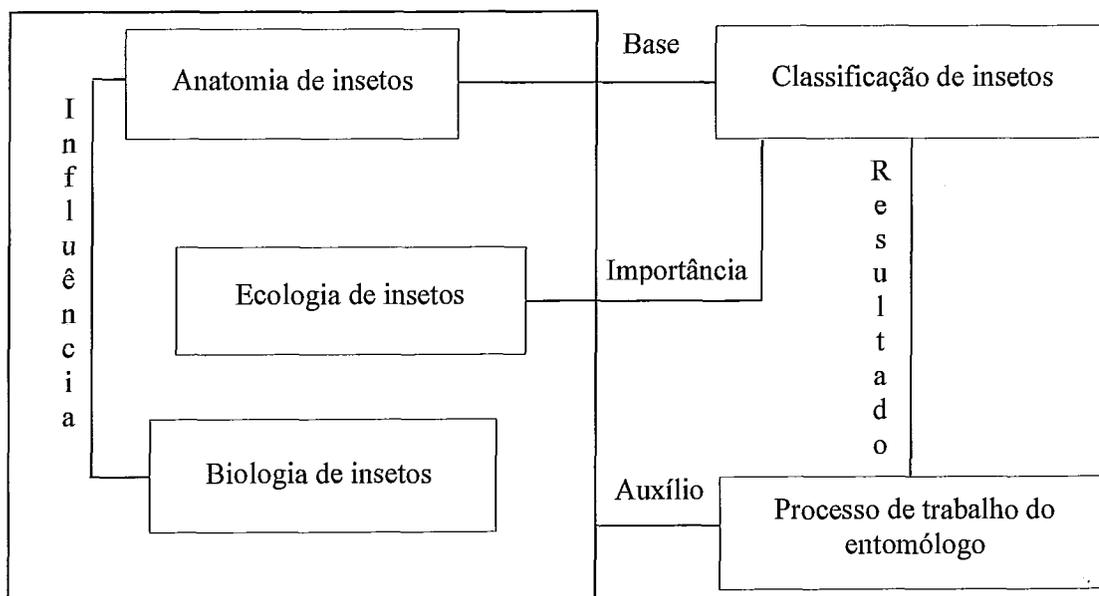


Figura 5.2 – Sub-teorias da Teoria de Domínio de Entomologia

Nas próximas sub-seções, cada sub-teoria terá sua especificação, captura e conceituação detalhadas. É importante ressaltar que, durante a fase de captura dos conceitos, surgiram novas questões de competência, que são apresentadas juntamente com a captura e conceituação de cada sub-teoria, para facilitar o seu entendimento e a visualização da abrangência de cada sub-teoria.

Os modelos, contendo os conceitos e relações de cada sub-teoria, estão representados graficamente em LINGO (FALBO, 1998). No Anexo 1, são apresentadas as tabelas com a descrição de todos os conceitos identificados e das propriedades das instâncias dos conceitos que constituem a Teoria do Domínio.

5.2.2.1. Sub-teoria Anatomia de Insetos

Esta sub-teoria apresenta os conceitos referentes à estrutura morfológica dos insetos, utilizada pelo entomólogo no desenvolvimento de suas atividades, principalmente a *atividade de identificação*⁴⁰.

▪ Especificação

O *inseto* constitui-se no objeto de estudo da entomologia. Todo o trabalho do entomólogo está baseado na *classificação*⁴¹ do inseto e sua fonte de trabalho é a anatomia que, por sua vez, está preocupada com a *morfologia*, forma e estrutura apresentada pelo inseto. A morfologia é o que caracteriza a *espécie* que identifica cientificamente o inseto.

A sub-teoria de anatomia de inseto apresenta apenas os conceitos relacionados às características mais gerais da estrutura e forma dos insetos, que fornecem uma base para a comparação com outros *artrópodos*⁴², possibilitando a sua identificação.

A sub-teoria de anatomia de insetos deve ser capaz de responder às seguintes questões de competência:

1. Quais as características morfológicas básicas dos insetos ?
2. Quais são as partes do corpo dos insetos ?
3. Como se caracterizam as partes do corpo dos insetos ?

▪ Conceituação

Com base nas questões de competência anteriormente relacionadas, foi identificado um conjunto de conceitos relacionados com a anatomia dos insetos, que são capazes de respondê-las.

⁴⁰ Atividade e Identificação são conceitos da sub-teoria Processo de trabalho do entomólogo (seção 5.2.2.5).

⁴¹ Classificação e Espécie são conceitos da sub-teoria Classificação de insetos (seção 5.2.2.4).

⁴² Instância da categoria taxonômica filo (definida na seção 5.2.2.4), à qual pertencem os insetos.

Para responder à primeira questão de competência, a *morfologia* dos insetos está dividida em *morfologia externa*, que se constitui do *corpo* dos insetos e do seu revestimento externo por uma estrutura, chamada *exoesqueleto*, que por sua vez é formado pela sobreposição de três *camadas*.

A *morfologia interna*, que constitui a outra subdivisão da morfologia, observa a caracterização dos *sistemas* e dos *órgãos* que os compõem, e o *endoesqueleto* que se constitui na estrutura de sustentação interna do *corpo* dos insetos, sendo considerada a extensão interna do exoesqueleto.

A morfologia é o principal instrumento para a identificação dos insetos, uma vez que todo inseto apresenta uma morfologia, que por sua vez caracteriza a espécie que identifica o inseto. No entanto, a morfologia interna só será utilizada se a partir da morfologia externa, o inseto não puder ser identificado.

A segunda questão de competência é respondida através da observação da composição do corpo dos insetos por três *regiões*: a *cabeça*, o *tórax* e o *abdome*, que estão interligadas entre si, e o exoesqueleto que é coberto por uma *superfície* composta de *suturas* e *escleritos*. Cada região apresenta externamente, *apêndices* e *segmentos*.

A terceira questão de competência obtém sua resposta através das taxonomias de apêndice e segmento que caracterizam os diversos tipos que compõem as três regiões do corpo dos insetos, e através da taxonomia do abdome, que é estabelecida de acordo com o tipo de interligação mantida com o tórax.

Os apêndices que são fixados nos segmentos se caracterizam de acordo com a sua localização nas regiões do corpo. Os *apêndices da cabeça*, que podem ser *apêndices móveis* ou *apêndices fixos*; os *apêndices torácicos* que são os *apêndices locomotores* e os *apêndices abdominais*.

Os segmentos também são classificados de acordo com sua localização nas regiões do corpo, e estão presentes também nos apêndices. Os *segmentos da cabeça* são as áreas delimitadas pelas suturas, chamadas cada uma delas de *área intersutural*; o *segmento torácico* composto pelo *protórax*, *mesotórax* e *metatórax*, onde ficam fixados os apêndices locomotores; e os *segmentos abdominais*, bem mais visíveis que os segmentos da cabeça e que os segmentos torácicos, são compostos pelos subconjuntos de segmentos chamados *segmentos pré-genitais*, *segmento genital* e *segmento pós-genital*. Esta denominação é dada devido à genitália dos insetos ser uma característica morfológica muito importante e bastante utilizada para a identificação dos insetos.

O abdome se caracteriza de acordo com as formas de conexão com o tórax e pode ser dos seguintes tipos: *abdome sésil*, *abdome livre* ou *abdome pedunculado*.

O conhecimento da anatomia dos insetos é essencial para o entendimento de como os insetos vivem e de como podem ser distinguidos uns dos outros e dos demais animais. Todas estas partes que compõem os insetos possuem outras subdivisões, com uma extensa nomenclatura anatômica, que não será capturada no escopo deste trabalho, mas que será incorporada à Teoria do Domínio, à medida que sejam necessárias quando da modelagem de bases de dados que necessitem dessas informações.

A figura 5.3 ilustra a representação em LINGO da organização dos conceitos e suas relações, identificados para esta sub-teoria. Em cinza, observa-se o conceito *espécie*, proveniente da sub-teoria *Classificação de Insetos*, definida na seção 5.2.2.4.

Para especificar as definições de conceitos na Teoria do Domínio e restrições sobre sua interpretação, definindo a semântica dos conceitos, são definidos axiomas. Uma vez que LINGO traz embutido em suas estruturas os axiomas epistemológicos, expressos em lógica de primeira ordem, tal como mostra o anexo 3, os mesmos não serão apresentados neste trabalho, pois sempre que as notações de subtipo e parte-de forem empregadas, assumimos que os axiomas deste tipo existem.

Os axiomas identificados para esta sub-teoria (três de consolidação), são descritos a seguir, inicialmente em linguagem natural para que seja possível sua validação pelos especialistas do domínio, cujo objetivo é consolidar as respostas às questões de competência para esta sub-teoria.

- a) Todo apêndice e Segmento da cabeça estão localizados na cabeça.
- b) Todo apêndice e segmento torácicos estão localizados no tórax.
- c) Todo apêndice e segmento abdominais estão localizados no abdome.

Os conceitos e propriedades de instâncias referentes a esta sub-teoria, estão descritos nas tabelas A1.1 e A1.2 do Anexo 1.

5.2.2.2. Sub-teoria Biologia de Insetos

Esta sub-teoria define os conceitos que caracterizam o ciclo de vida dos insetos visto que isto influencia a observação das características morfológicas dos mesmos, auxiliando no desenvolvimento do processo de trabalho do entomólogo.

▪ Especificação

A biologia do inseto está preocupada com o seu *ciclo de vida*, desde o momento da sua fecundação até a sua morte. Comumente, a *reprodução* nos insetos é dependente do encontro dos dois sexos e da fecundação do óvulo pelo espermatozóide. Assim, a vida de um inseto começa no primeiro *estágio* do seu *desenvolvimento*, o *ovo*. Seu desenvolvimento é acompanhado por uma série de mudanças de forma, até atingir o estágio de *adulto* e iniciar a sua reprodução, dando início a um novo ciclo. Desta forma, a sub-teoria sobre a biologia de insetos deverá ser capaz de responder às seguintes questões de competência:

1. Como se caracteriza o ciclo de vida dos insetos ?
2. Como se caracteriza o desenvolvimento dos insetos ?
3. Quais os tipos de reprodução dos insetos ?

▪ Conceituação

Considerando as questões de competência anteriormente relacionadas, foi identificado um conjunto de conceitos relacionados com a biologia do inseto, que são capazes de respondê-las.

A primeira questão de competência é respondida pela observação de que todo inseto possui um ciclo de vida, que é composto pelas fases de *reprodução* e desenvolvimento. O ciclo de vida pode ser de dois tipos⁴³: heterodinâmico e homodinâmico, de acordo com a existência ou não de um período de dormência (inatividade) durante o seu desenvolvimento.

Para responder à segunda questão de competência, considera-se que durante o seu desenvolvimento, os insetos mudam de forma e passam por vários *estágios*, que são diferentes entre si. Alguns insetos passam por mudanças de forma muito pequena e os jovens e os *adultos* são muito semelhantes, exceto no tamanho; em outros casos, os jovens e os adultos são muito diferentes em hábitos e na forma.

O primeiro estágio da vida de um inseto é o ovo, que dá nascimento à *larva*, que constitui o segundo estágio da vida do inseto. As larvas, de acordo com a sua forma, podem ser dos tipos: eruciforme, escarabeiforme, campodeiforme, elateiforme e vermiforme.

Aos poucos as larvas, semelhantes a vermes, vão desenvolvendo asas e olhos rudimentares e, depois da última mudança larval, entram no estágio de *pupa*, durante o qual costumam ficar encerradas em um casulo⁴⁴. As pupas, de acordo com os apêndices existentes neste estágio, podem ser dos tipos: obtecta, exarada e coarctata. Depois de imobilizada, por certo tempo, a pupa rompe sua pequena prisão para atingir o último estágio, *adulto*.

Na *paurometabolía*⁴⁵, a *ninfa*, pode parecer uma versão em miniatura de um adulto. À medida que cresce, precisa se desfazer de sua carapaça rígida, o *exoesqueleto*⁴⁶.

Desta forma o ciclo de vida do inseto se completa e dá início à sua reprodução. Para responder à terceira questão de competência, a reprodução, que gera o ovo, de acordo com a forma e com o estágio no qual a mesma ocorre, pode ser dos tipos⁴³: fecundação, partenogênese, pedogênese e poliembrionia. A figura 5.4 mostra a representação em LINGO da organização dos conceitos identificados para esta sub-teoria, onde podemos observar em cinza, o conceito *inseto* proveniente da sub-teoria *Anatomia de Insetos* (seção 5.2.2.1).

Todo esse conhecimento sobre a biologia do inseto é necessário para que o entomólogo possa realizar suas atividades, principalmente as atividades de manejo e controle, além de servir de embasamento para a análise morfológica do inseto.

Os conceitos e propriedades de instâncias referentes a esta sub-teoria, estão descritos nas tabelas A1.3 e A1.4 do Anexo 1.

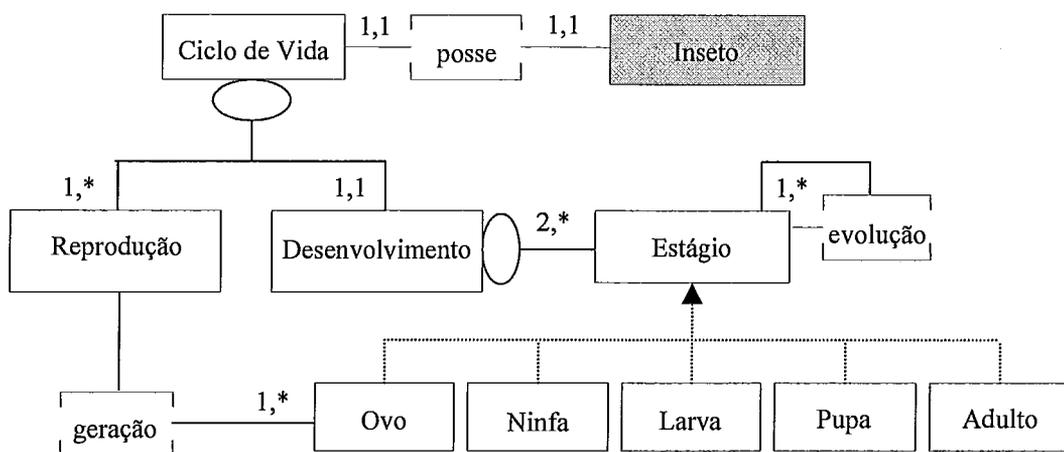


Figura 5.4 – Conceitos da Sub-teoria Biologia de Insetos.

⁴³ Tipo é uma propriedade dos conceitos: ciclo de vida, larva, pupa e reprodução.

⁴⁴ É o invólucro que envolve o inseto durante seu estágio de pupa.

⁴⁵ Paurometabolía e holometabolía são tipos de metamorfose que é uma propriedade do conceito desenvolvimento.

⁴⁶ Exoesqueleto é um conceito da sub-teoria Anatomia de inseto.

Para esta sub-teoria, três axiomas ontológicos são identificados a seguir, descritos em linguagem natural que respondem à segunda questão de competência:

- a) Todo desenvolvimento com metamorfose *paurometabolia* não tem como componentes os estágios larva e pupa.
- b) Todo desenvolvimento com metamorfose *holometabolia* não tem como componente o estágio ninfa.
- c) Todo ciclo de vida do tipo *heterodinâmico* apresenta o estágio de desenvolvimento do tipo *pupa*.

5.2.2.3. Sub-teoria Ecologia de Insetos

A ecologia é uma ciência básica para a exploração do ambiente, mas os problemas que aparecem são extremamente difíceis, envolvendo todos os campos científicos. Esta sub-teoria aborda os conceitos ecológicos básicos relacionados com o comportamento dos insetos, a partir da identificação dos mesmos. É importante saber sobre o comportamento de insetos, porque é através dele que ocorre a interação com o meio ambiente.

▪ Especificação

O termo ecologia é derivado do grego *oikos*, significando casa, lar, habitação ou ambiente. Ecologia significa então, o estudo da relação dos organismos com o meio ambiente. Esta sub-teoria expressa, através de seus conceitos, o objeto de estudo dos ecologistas, que examinam a dinâmica da *população*, a interação entre insetos (como a competição e alimentação), o *meio ambiente* (que inclui a disponibilidade de nutrientes) e o fluxo de energia através do *ecossistema*, e os padrões de distribuição e sucessão.

O estudo ecológico conduz o homem a uma compreensão do funcionamento dos sistemas naturais e o habilita para a prática da conservação da natureza. Como os insetos detêm o maior número de espécies no reino⁴⁷ animal, havendo uma estimativa de que haja mais de 900.000 espécies, o estudo de sua ecologia também é muito importante para a determinação de sua *classificação*, bem como para o conhecimento da sua *distribuição* no globo terrestre, permitindo a prescrição do manejo ou controle do inseto.

⁴⁷ Reino é um conceito da sub-teoria classificação de insetos (seção 5.2.2.4).

Além disso, o *ciclo de vida*⁴⁸ de cada *espécie* ajusta-se intimamente às condições climáticas do meio ambiente, podendo ser observada a influência que os fatores do ambiente exercem sobre os insetos, através do seu *desenvolvimento, comportamento* e distribuição.

Então, a sub-teoria sobre a ecologia de insetos deverá ser capaz de responder às seguintes questões de competência:

1. Quais os principais componentes do ecossistema que influenciam o comportamento dos insetos ?
2. Quais os principais fatores ecológicos que influenciam o comportamento dos insetos ?
3. O que caracteriza uma comunidade de insetos ?
4. O que caracteriza uma população de insetos?
5. Qual a função dos insetos dentro da comunidade ?
6. O que caracteriza a distribuição das populações dos insetos ?
7. O que caracteriza o comportamento dos insetos ?

▪ **Conceituação**

Em função das questões de competência identificadas, um conjunto de conceitos que representam a sub-teoria foi identificado, de forma que fosse capaz de respondê-las. As questões de competência são respondidas pela sub-teoria representada em LINGO, na figura 5.5, onde podemos observar em cinza, o conceito *inseto* proveniente da sub-teoria *Anatomia de Insetos*.

Respondendo à primeira questão de competência, observa-se que, em geral, a unidade básica da ecologia é o *ecossistema*. Um ecossistema consiste numa rede complexa de relações de mútua influência entre a flora⁴⁹, a fauna⁵⁰ e os microrganismos⁵¹ de uma determinada área ou região (*localidade*) e todos os elementos físicos naturais (geológicos, climáticos, tempo, etc.).

⁴⁸ Ciclo de vida e Desenvolvimento são conceitos da sub-teoria Biologia de insetos (seção 5.2.2.2).

⁴⁹ É o conjunto das plantas que crescem numa região.

⁵⁰ Conjunto das espécies animais, que vivem em um espaço geográfico ou em determinado habitat.

⁵¹ Organismos microscópios que promovem fermentação e causam doenças infecciosas.

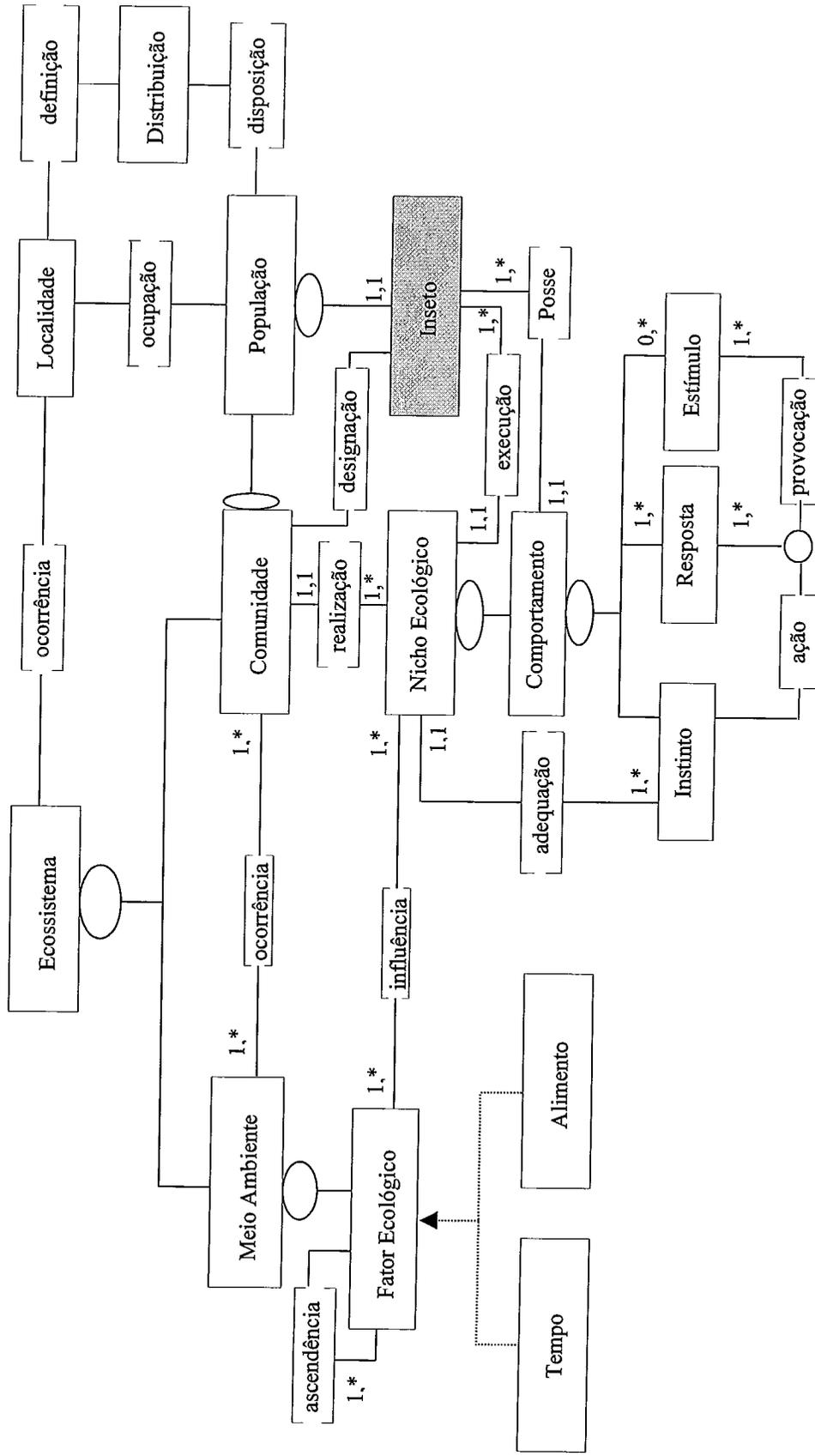


Figura 5.5 – Conceitos da Sub-teoria Ecologia de Insetos.

Nesta sub-teoria, os componentes do ecossistema que serão observados são: o *meio ambiente* que cada inseto ocupa e que influencia o seu modo de vida particular, e a *comunidade* que ocorre no meio ambiente e se constitui do conjunto de populações que ocupam uma localidade comum e se inter relacionam.

Em resposta à segunda questão de competência, a ação de fatores físicos e biológicos (*fatores ecológicos*) produz uma grande variedade de meios ambientes nas diversas partes da Terra. Pode-se medir cada um dos fatores e observar seus efeitos sobre os insetos, mas todos são relacionados entre si e nenhum age independentemente. Os principais fatores ecológicos que influenciam o comportamento dos insetos são o *alimento* e o *tempo* (estado da atmosfera), composto de vários fatores físicos variáveis num determinado momento e localidade como radiação, temperatura, umidade, luz, vento e gravidade.

Para responder à terceira questão de competência, as comunidades são designadas tendo como referência um inseto característico dominante (por exemplo, comunidade de besouros) ou uma característica abiótica⁵² dominante (por exemplo, comunidade de águas correntes).

Ao grupo de insetos que ocupam uma localidade, capazes de se reproduzir entre si e que compartilham determinado recurso genético, compondo uma comunidade, é denominado de *população*. As populações são consideradas como unidades evolutivas, visto que ocorre um processo evolutivo sempre que a frequência dos genes de uma população se altera nitidamente em gerações sucessivas, o que responde à quarta questão de competência.

Respondendo à quinta questão de competência, o *nicho ecológico* é a execução de uma função por um inseto dentro da comunidade, englobando também a localidade ocupada. Implica ainda, em todos os componentes químicos, físicos e biológicos, como o *alimento*, o período do dia em que os seres se alimentam, a temperatura, a umidade, as partes da localidade que usa (por exemplo, as árvores ou o descampado), a forma como se reproduz e a forma como se comporta.

A sexta questão de competência é respondida pela relação de ocupação de uma localidade pela população e a relação de definição da *distribuição* da população de insetos pela localidade onde os mesmos ocorrem e habitam.

⁵² Cada um dos fatores físicos e químicos que influenciam os seres vivos.

A sétima questão de competência é respondida pelos conceitos: estímulo, resposta e instinto. Os insetos respondem a muitos *estímulos*, o que corresponde ao comportamento como parte do seu nicho ecológico dentro da comunidade, no seu ecossistema. Dentre os diversos estímulos, podemos destacar os fatores ecológicos como a luz, a força da gravidade e a temperatura, além da água, dos produtos químicos, do tato ou contato, das correntes de ar ou de água.

A *resposta* mais simples de um inseto a um estímulo é um reflexo. Além disso, os insetos também possuem os comportamentos instintivos (*instinto*), que exercem uma ação sobre a resposta dada. Os instintos são herdados e evoluíram para serem adaptativos, adequando o inseto a seu papel específico (nicho ecológico).

A ecologia de insetos, além da observação do comportamento individual das espécies, isto é, do seu nicho ecológico, observa ainda, a comunidade como um todo, enfocando: seu desenvolvimento, sua população e as relações entre os insetos componentes, além de sua distribuição pela Terra. Tudo isso compreende os estudos ecológicos das populações, comunidades e ecossistemas.

Os conceitos e propriedades de instâncias referentes a esta sub-teoria, estão descritos nas tabelas A1.5 e A1.6 do Anexo 1. Para esta sub-teoria, não foi identificado nenhum axioma, devido tratar-se de uma parte da entomologia pouco estudada pelos entomólogos em geral, existindo poucos especialistas e dados disponíveis.

5.2.2.4. Sub-teoria Classificação de Insetos

Esta sub-teoria mostra os conceitos envolvidos na identificação dos insetos, resultado de sua classificação. Os animais podem ser classificados de várias formas, mas a classificação seguida pelos zoólogos é baseada primariamente em caracteres estruturais: os animais com certas estruturas morfológicas em comum são classificados em um grupo e os com outras estruturas em outros grupos.

▪ Especificação

A *classificação* de organismos é o agrupamento de *espécies* próximas, que possuem características que as aproximam e que as distinguem. Os insetos como qualquer animal também seguem uma forma de agrupamento.

A taxonomia ou sistemática, é a área que estuda a classificação dos insetos. Fundamenta-se, principalmente, na análise comparativa de seus caracteres morfológicos (*morfologia*⁵³), auxiliada pela observação do seu *comportamento*⁵⁴ (emissão de som e de luz, padrões de corte, dentre outros), e atualmente através das tecnologias de biologia molecular. Mas ainda hoje, com raríssimas exceções, toda classificação animal se fundamenta no estudo comparativo dos caracteres morfológicos. Esta sub-teoria deverá então, responder às seguintes questões de competência:

1. Em quais categorias taxonômicas os insetos podem ser classificados ?
2. Quais são as subdivisões das categorias taxonômicas dos insetos mais utilizadas ?

▪ Conceituação

Os insetos são identificados através da sua classificação, que é constituída de várias *categorias taxonômicas*, originadas com base na sua morfologia. Para responder à primeira questão de competência, o *reino* animal é dividido em cerca de uma dúzia de grupos maiores chamados *filos*. Os insetos pertencem ao filo *arthropoda* (*athro*, articulação ou segmento; *poda*, pé ou apêndice), que corresponde a, aproximadamente, 78% do reino animal. Mas a classificação dos animais não termina com filós. Cada filo é dividido, com base em caracteres estruturais, em grupos chamados *classe*. Os insetos constituem uma delas que é a classe *Insecta*.

As classes dos animais são subdivididas em *ordens*, as ordens em *famílias*; as famílias em *gêneros* e os gêneros em *espécies*. Logo, as principais categorias taxonômicas da classificação animal são: filo, classe, ordem, família, gênero e espécie. Os insetos existentes e identificados podem ser colocados em 26 ordens, compreendendo perto de 1000 famílias e muitos milhares de gêneros. A categoria básica, ou unidade zoológica de classificação é a espécie. Cada espécie recebe um nome próprio regional (nome trivial ou popular).

A segunda questão de competência tem sua resposta na observação da existência de categorias intermediárias, que podem ser estabelecidas para uma melhor disposição taxonômica de determinado grupo, como: *subfilo*, *subclasse*, *subordem*, *superfamília*, *subfamília*, *tribo*, *subgênero*, *subespécie*, dentre outras⁵⁵. A figura 5.6 mostra a organização dos conceitos identificados para esta sub-teoria, onde os conceitos em cinza *inseto* e *morfologia* são provenientes da sub-teoria *Anatomia de Insetos*.

⁵³ Morfologia é um conceito da sub-teoria Anatomia de inseto.

⁵⁴ Comportamento é um conceito da sub-teoria Ecologia de inseto.

⁵⁵ Os sistematas podem criar novos agrupamentos (categorias) visando facilitar a atividade de identificação.

Os conceitos e propriedades de instâncias referentes a esta sub-teoria, estão descritos nas tabelas A1.7 e A1.8 do Anexo 1. Para esta sub-teoria, foi identificado apenas um axioma epistemológico descrito em linguagem natural: *Tribo é a subdivisão de uma subfamília.*

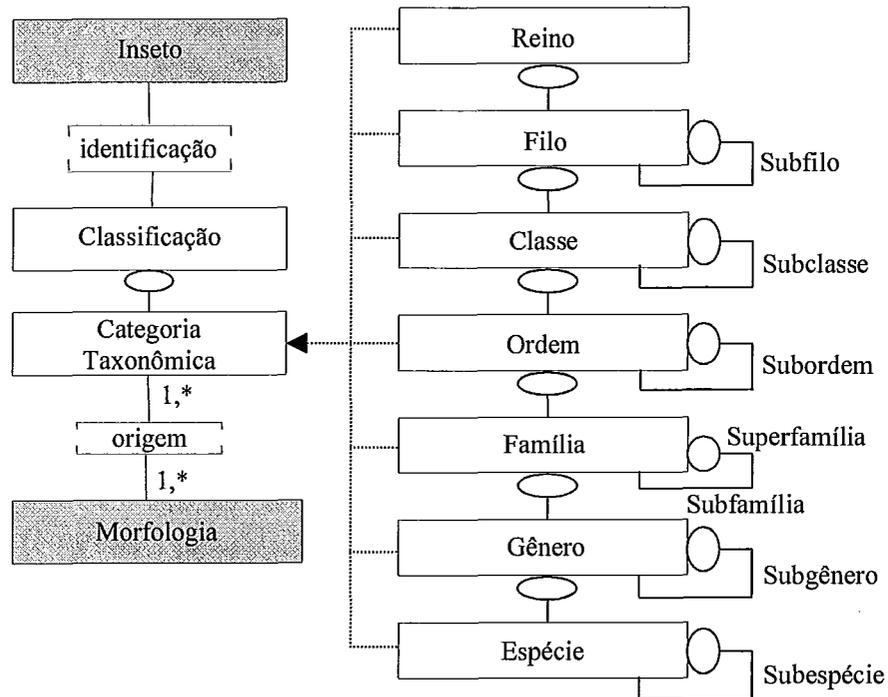


Figura 5.6 – Conceitos da Sub-teoria Classificação de Insetos.

5.2.2.5. Sub-teoria Processo de Trabalho do Entomólogo

Para o entendimento do domínio de entomologia é importante o conhecimento sobre o que envolve o processo de trabalho do entomólogo, já que as bases de dados modeladas também visam apoiar este processo. Então, nesta sub-teoria são apresentados os conceitos envolvidos no processo de trabalho do entomólogo.

▪ Especificação

O estudo dos insetos é a atividade principal de um *entomólogo*. Algumas das investigações feitas pelos entomólogos são executadas no intuito de reduzir os problemas causados pelos insetos que transmitem doenças ou que atacam plantações. Assim, é importante também o conhecimento dos conceitos envolvidos no processo de trabalho do entomólogo que utiliza os conceitos descritos nas sub-teorias anteriores. A figura 5.7 mostra a organização dos conceitos identificados para esta sub-teoria, onde

em cinza, observa-se os conceitos *classificação* e *inseto*, provenientes das sub-teorias *Classificação* e *Anatomia de Insetos*, respectivamente.

Esta sub-teoria deve responder às seguintes questões de competência:

1. Que fatores estão envolvidos no processo de trabalho de um entomólogo ?
2. Quais são os tipos de atividades desenvolvidas por um entomólogo ?
3. Quais os recursos utilizados pelas atividades ?
4. Quais produtos são consumidos ou produzidos pelas atividades ?
5. Quem realiza as atividades ?
6. Quais os tipos de controle que podem ser executados ?

▪ **Conceituação**

Para responder à primeira questão de competência, observa-se que o processo de trabalho do entomólogo é constituído de diversas *atividades*. Cada atividade consome e produz diversos *produtos*, além de utilizar na sua execução vários *recursos*.

A segunda questão de competência é respondida através da classificação das atividades em: atividades de *pré-identificação* (*coleta*, *tratamento* e *armazenamento*), atividade de *identificação*, atividade de *manejo* e atividade de *controle*. Para a execução das atividades, existe uma ordem específica, caracterizando a precedência entre atividades.

A terceira questão de competência tem sua resposta através da observação de que, na execução das atividades são utilizados *recursos*, que podem ser classificados em: *procedimentos*, *código de nomenclatura* e *recursos humanos*. O *código de nomenclatura* zoológica procura assegurar nome único e distinto para cada espécie e rejeitar os nomes que causam confusão ou ambigüidades e tem como objetivo promover a estabilidade e a universalidade dos nomes científicos⁵⁶.

Em resposta à quarta questão de competência, os diversos *produtos* resultantes ou consumidos pela execução das atividades, podem ser classificados em: *chave de classificação*, *coleção* entomológica, o *espécime* que representa o inseto na coleção, com sua *classificação*, *trabalhos científicos* e *receituários de controle*.

O nome científico é a chave para a documentação passada sobre cada um dos espécimes de insetos e o elo de ligação entre todos os passos do conhecimento e do *trabalho científico*. É a pedra angular do entendimento.

⁵⁶ Nome científico é uma propriedade da categoria taxonômica.

Respondendo à quinta questão de competência, alguns dos recursos humanos que executam as atividades têm que ser especialistas, como o próprio *entomólogo* que precisa ser especialista em sistemática (*sistemata*) para executar a identificação de insetos; outros podem ser apenas experientes como no caso do *coletor* e do *curador*.

A sexta questão de competência tem sua resposta pela observação de que existem vários tipos de controle que podem ser executados, e são recomendados pelos entomólogos através do *receituário de controle*, cuja classificação é: *controle mecânico*; *controle químico*, que pode ser classificado em *inseticida* e *formulação*; *controle integrado*; *controle cultural*; *controle biológico*, classificado em *inimigo natural* e *patógeno*; *controle comportamental*; e *controle genético*.

Detalhando um pouco mais, a atividade de identificação pode ser feita por: análise do espécime por especialista (*sistemata*), comparação direta com espécimes devidamente identificados em coleções entomológicas, e através de bibliografia usando-se as descrições de espécies ou *chaves de identificação*. Toda coleção é administrada por uma curadoria (*curador*), que é responsável por toda a política prática e científica da *coleção*.

Os conceitos e propriedades de instâncias referentes a esta sub-teoria, estão descritos nas tabelas A1.9 e A1.10 do Anexo 1. Para esta sub-teoria, foram identificados três axiomas de consolidação descritos em linguagem natural a seguir. O primeiro axioma responde à quinta questão de competência e os outros dois axiomas respondem à terceira questão de competência.

- a) Toda atividade de identificação utiliza o recurso humano *sistemata*.
- b) Toda atividade de identificação sempre utiliza o recurso código de nomenclatura.
- c) Toda atividade de pré-identificação utiliza um procedimento na sua realização.

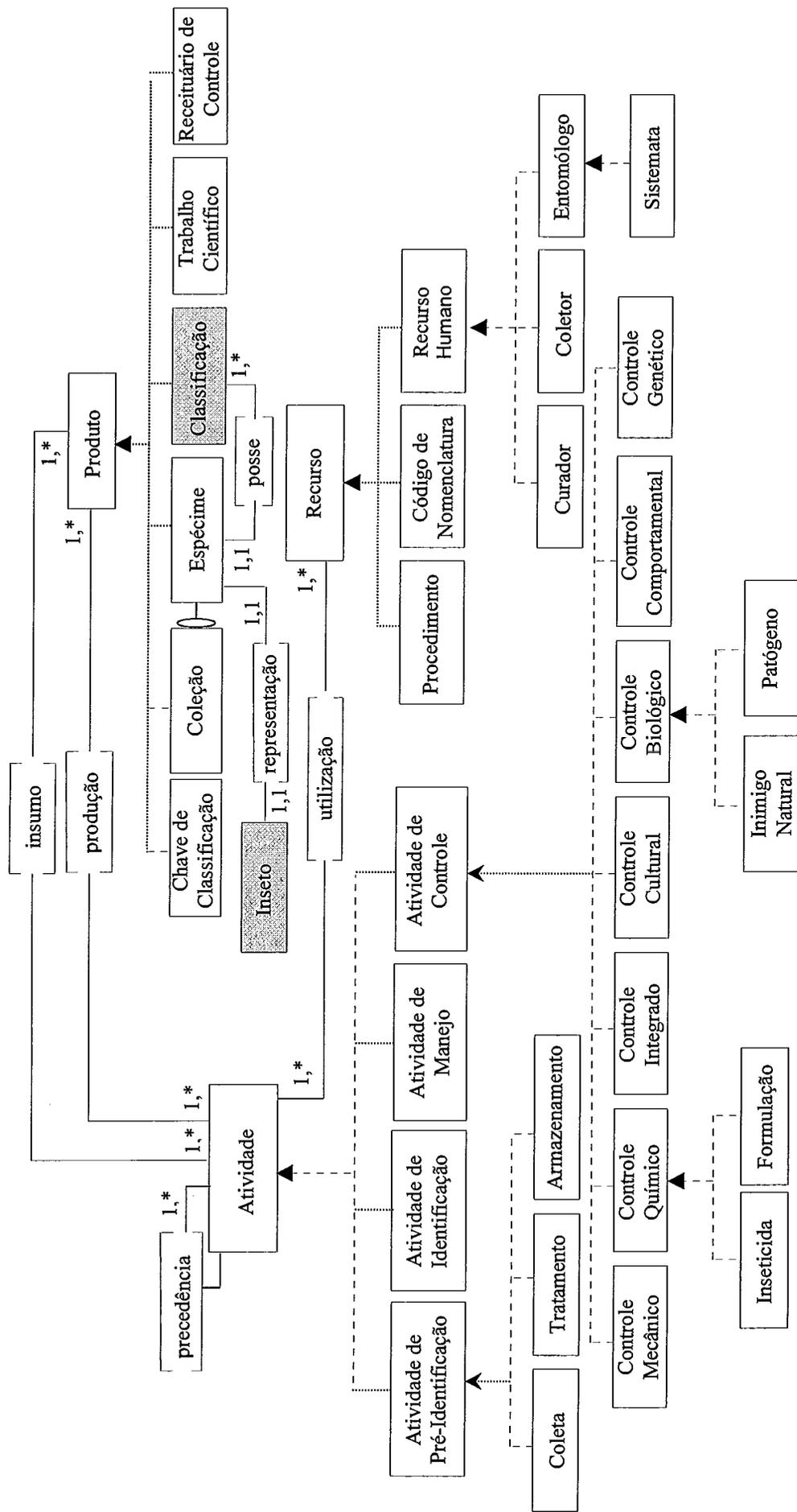


Figura 5.7 – Conceitos da Sub-teoria Processo de Trabalho do Entomólogo.

Na última iteração da atividade de definição, faz-se um mapeamento entre as sub-teorias identificadas e as tarefas mais típicas realizadas no domínio de Entomologia. O objetivo deste mapeamento é simplesmente a indicação do que deve ser considerado quando do desenvolvimento de uma aplicação (OLIVEIRA, 1999).

É importante considerarmos as tarefas realizadas no domínio, pois além da consciência de que existem objetos no mundo, que possuem propriedades que por sua vez podem ter valores, e que esses objetos podem existir em várias relações com outros objetos, considera-se também que as propriedades e relações podem mudar com o tempo e que existem eventos que ocorrem em diferentes instantes do tempo. Além disso, existem processos nos quais os objetos participam ao longo do tempo.

As tarefas mais típicas realizadas no domínio de entomologia são: *classificação* do inseto sob vários aspectos além da sua identificação, *diagnóstico* e *prescrição* do manejo e controle, *planejamento* e *controle* das atividades do entomólogo, *avaliação* de coleções e das classificações existentes, originando as chaves de classificação. A tabela 5.1 relaciona as sub-teorias que devem ser consideradas para as tarefas mencionadas.

Tabela 5.1 – Correspondência entre tarefas e sub-teorias

<i>TAREFA</i>	<i>SUB-TEORIAS</i>
Classificação	Anatomia de Insetos, Biologia de Insetos, Ecologia de Insetos e Classificação de Insetos.
Diagnóstico	Ecologia de Insetos e Classificação de Insetos.
Planejamento	Processo de Trabalho do Entomólogo.
Controle	Processo de Trabalho do Entomólogo
Prescrição	Biologia de Insetos, Ecologia de Insetos e Classificação de Insetos.
Avaliação	Anatomia de Insetos, Classificação de Insetos e Processo de Trabalho do Entomólogo.

5.2.3 - Formalização e Codificação

A etapa de formalização consiste em escrever a ontologia na linguagem de representação selecionada, ou seja, nesta etapa deve-se definir explicitamente os conceitos e relações através de uma linguagem formal.

Os axiomas epistemológicos de subtipo e os referentes à composição todo-parte, são descritos implicitamente através do uso de LINGO na definição das sub-teorias. Os axiomas de consolidação, epistemológicos e ontológicos identificados ao longo da etapa de definição deverão ser formalizados em lógica de primeira ordem, devido ser esta a representação formal utilizada atualmente na Estação TABA (FALBO, 1998).

No contexto da Estação TABA, as ontologias de domínio são incorporadas através da arquitetura de Servidores de Conhecimento (FALBO, 1998), utilizando a ferramenta EDITED, desenvolvida por OLIVEIRA (1999). Os conceitos, relações e axiomas são definidos como módulos de conhecimento em Prolog, que por sua vez são incorporados ao Servidor de Conhecimento. No contexto deste trabalho, não foi identificada a necessidade de formalização dos axiomas, devido nosso enfoque estar na modelagem conceitual, onde os axiomas em linguagem natural permitem verificar a consistência do modelo ER definido.

A etapa de codificação corresponde à implementação dos axiomas escritos em linguagem formal. Na Estação TABA isto é feito em Prolog. A partir desta codificação, a Teoria do Domínio passa a estar disponível no ambiente, na forma de módulos de conhecimento, podendo então, ser instanciada e acessada no ambiente gerado, formando uma base de conhecimentos sobre o domínio de entomologia.

5.2.4 - Atividades de Apoio

As atividades de documentação, avaliação, aquisição do conhecimento e integração com ontologias existentes, devem estar presentes ao longo de todo o processo de desenvolvimento da Teoria do Domínio (figura 5.1), variando seu grau de influência conforme a etapa que esteja sendo executada.

▪ Documentação

A documentação gerada é específica para cada uma das etapas, sendo mais detalhada na fase de definição, que no caso desta Teoria do Domínio constitui-se basicamente de:

- dicionário de conceitos com seus significados, sinônimos, exemplos de instâncias e suas propriedades (ver tabelas A1.1, A1.3, A1.5, A1.7 e A1.9 no Anexo 1);
- organização dos conceitos das sub-teorias em árvores de classificação (quando pertinentes), definindo taxonomias de especialização e composição (vide figuras 5.3, 5.4, 5.5, 5.6 e 5.7 da seção 5.2.2);

- dicionário de atributos (propriedades) de conceitos definindo significado, tipo e valores de domínio (ver tabelas A1.2, A1.4, A1.6, A1.8 e A1.10 no Anexo 1).

▪ **Avaliação**

Esta atividade deve ser realizada com mais de um especialista e deve buscar avaliar segundo alguns dos critérios de qualidade definidos na literatura (tabela 2.1).

No contexto deste trabalho, além das avaliações normais com os especialistas do domínio envolvidos na definição da Teoria do Domínio, feitas ao longo da etapa de definição, foi utilizado um formulário (Anexo 2) para avaliar o conteúdo da Teoria do Domínio quanto aos critérios de: clareza, consistência, generalidade, completude, concisão e robustez (OLIVEIRA, 1999). A figura 5.8 mostra os gráficos das avaliações sob o ponto de vista dos avaliadores em função de seu perfil em relação ao conhecimento tanto de ontologia quanto do domínio de Entomologia.

Foi solicitado a nove pessoas, com diferentes perfis, que fizessem a avaliação. Seis delas responderam, sendo dois especialistas do domínio que participaram da definição da Teoria do Domínio, dois avaliadores com razoável conhecimento de ontologia, um analista de sistemas e um bibliotecário que atua na área de organização de informação.

No formulário (Anexo 2), foi solicitado que as características fossem avaliadas segundo seu grau de existência nos itens avaliados, com valores variando de zero a quatro (nenhum, insuficiente, razoável, bom ou excelente, respectivamente). Os gráficos da figura 5.8 nos permitem verificar que nenhuma das características teve sua média abaixo de dois (razoável), sendo que a Teoria do Domínio, como um todo, foi considerada boa, de acordo com o nível dos critérios em geral.

Também foi solicitado aos avaliadores que identificassem possíveis utilizações e formas de utilizar a Teoria do Domínio. Dois avaliadores sugeriram duas formas de utilização da Teoria do Domínio, mostrada na figura 5.9. O primeiro exemplo é a visão de utilização da Teoria do Domínio de um analista de sistemas, não especialista em ontologia e com nenhum conhecimento em entomologia, e o segundo exemplo é a visão de um entomólogo.

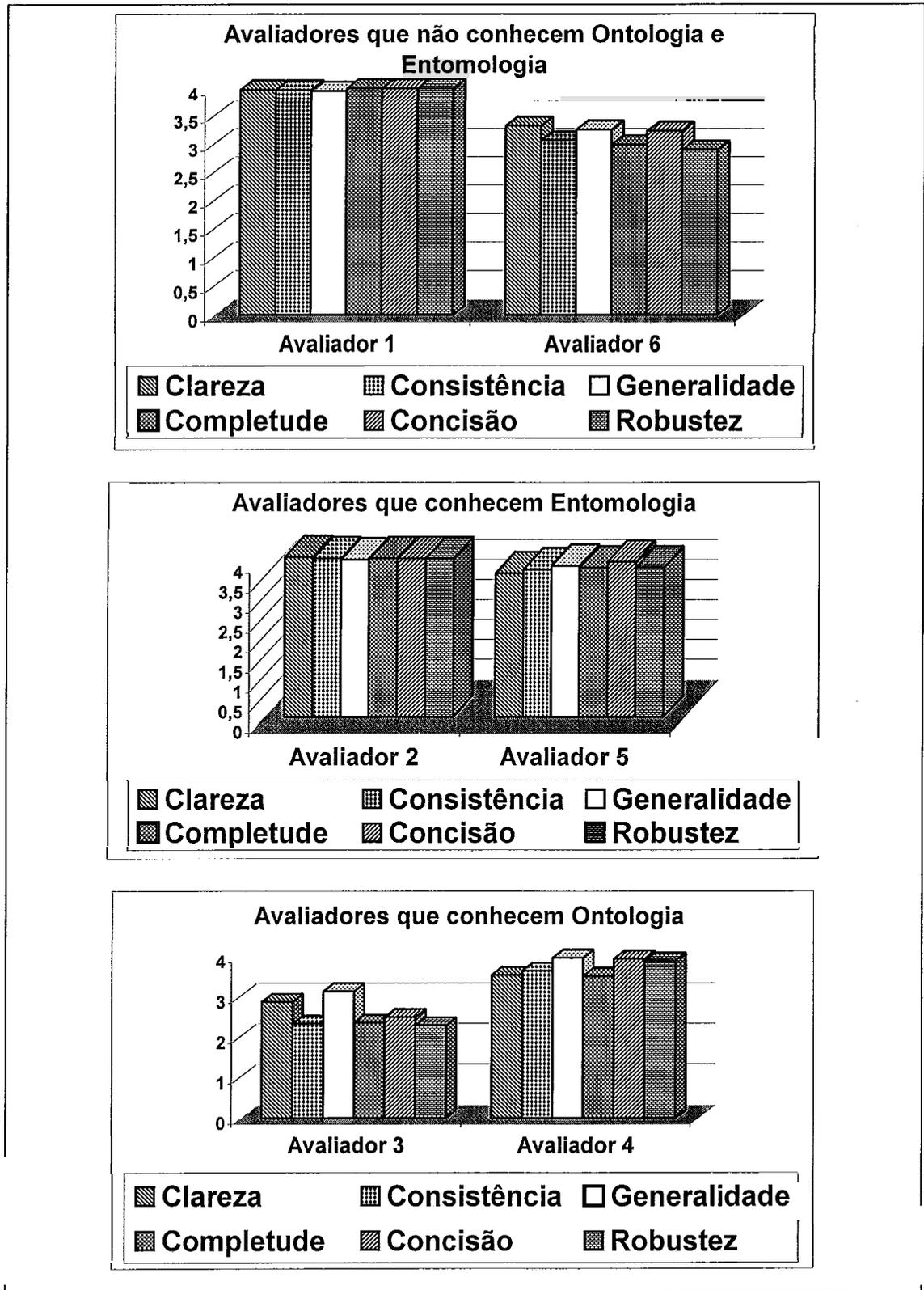


Figura 5.8 – Avaliação da Teoria do Domínio de Entomologia

12	<p>Em qual (is) situação (ões) você consegue visualizar a utilização da Teoria do Domínio de Entomologia ?</p> <p>PARA A DESCRIÇÃO DO AMBIENTE ATUAL E AUTOMATIZADO, NA CONSTRUÇÃO DE UMA BASE DE DADOS ENTOMOLÓGICA. SEM ESSA TEORIA DO DOMÍNIO, ESSA ESPECIFICAÇÃO SERIA PRATICAMENTE IMPOSSÍVEL DE REALIZAR.</p>
13	<p>Se você conseguiu visualizar uma ou mais utilizações para a Teoria do Domínio de Entomologia, de que forma ela poderia ser utilizada ?</p> <p>NO CASO ESPECÍFICO DE UMA BASE DE DADOS, SERIA UTILIZADA NA NORMALIZAÇÃO E NA MODELAGEM DOS DADOS.</p>
12	<p>Em qual (is) situação (ões) você consegue visualizar a utilização da Teoria do Domínio de Entomologia ?</p> <ul style="list-style-type: none"> • Em treinamentos a estudantes, como uma introdução à entomologia (não só estudantes, como também q pessoas interessadas no tema em questão e/ou profissionais, técnicos, etc.) • Cursos de Entomologia Básica (conceitos)
13	<p>Se você conseguiu visualizar uma ou mais utilizações para a Teoria do Domínio de Entomologia, de que forma ela poderia ser utilizada ?</p> <ul style="list-style-type: none"> • Na forma de uma apostila, com aulas teóricas e práticas, complementando um curso sobre Entomologia Básica

Figura 5.9 – Sugestões de utilização da Teoria do Domínio de Entomologia

▪ **Aquisição de conhecimento e integração com ontologias existentes**

Técnicas de aquisição de conhecimento são intensamente utilizadas na etapa de definição enquanto que, a integração com ontologias existentes pode ser uma atividade bastante comum quando existem várias sub-teorias a serem definidas. No escopo deste trabalho, à medida que as sub-teorias foram sendo definidas, as mesmas foram sendo integradas através das relações entre os conceitos das diferentes sub-teorias. Na figura 5.10 podemos visualizar a integração das diferentes sub-teorias que compõem a Teoria do Domínio de entomologia.

Podemos observar que a integração é feita em torno do conceito central da Teoria do Domínio de entomologia, o inseto, pois se trata do objeto de estudo do entomólogo. Por motivo de clareza e melhor visualização, não foram incluídos todos os conceitos e relações *supertipo/subtipo* e *parte de*, de todas as sub-teorias, mas apenas aqueles que estavam diretamente relacionados com o conceito central.

Apresentamos na seção seguinte as etapas seguidas para a instanciação do ADS *Insecta* a partir da infraestrutura da Estação TABA.

5.3 – Definição e Instanciação do Ambiente *Insecta*

A construção de um ADSOD, no contexto da Estação TABA, é feita através da utilização de ferramentas de definição de ambientes, que têm como objetivo permitir a edição do processo e da Teoria do Domínio, que serão utilizados no ambiente construído. Na construção do ADSOD *Insecta* serão utilizados uma especialização do processo definido no capítulo 3 (seção 3.2) e a Teoria do Domínio definida na seção 5.2.

A Estação TABA utiliza um esquema de definição de processo de desenvolvimento de software conforme ilustrado pela figura 5.11 (a) onde, inicialmente é definido um processo padrão para uma organização em particular, a partir do qual, diferentes processos de desenvolvimento de software podem ser especializados de acordo com as características de cada tipo de software (OLIVEIRA *et al.*, 2000).

Um tipo de software refere-se a um conjunto de sistemas de software que utiliza uma determinada tecnologia de desenvolvimento (como, por exemplo, sistemas baseados em conhecimento, sistemas de informação, etc), seguindo um paradigma específico (como, por exemplo, orientado a objetos, estruturado, etc.).

MACHADO (2000) acrescenta que, além do tipo de software, deve-se observar características do desenvolvimento (como, por exemplo, o tipo de ADS). Nesse momento, novas atividades podem ser definidas e incluídas no processo especializado. No entanto, todos os elementos básicos definidos no Processo Padrão deverão sempre estar presentes nos processos especializados.

Para ser utilizado em um projeto específico, o processo especializado mais adequado para um determinado tipo de software deve ser instanciado para atender às características do projeto.

No contexto deste trabalho, porém, o esquema de definição do processo de desenvolvimento de software foi modificado, uma vez que a padronização do processo foi feita considerando o tipo de software, *bases de dados de pesquisa*, a partir do qual, diferentes processos de desenvolvimento de software podem ser especializados de acordo com as características da organização ou mesmo para a orientação ao domínio, conforme ilustrado pela figura 5.11 (b).

Segundo OLIVEIRA (1999), as atividades mais comuns e que em geral estão presentes em um processo de desenvolvimento de software são, direta ou indiretamente, dependentes do conhecimento do domínio. Além disso, observou que em algumas atividades ocorre a existência de uma maior necessidade do conhecimento do domínio do que em outras. Para definir a orientação a domínio, nessas atividades, considerou o fato de que no processo de desenvolvimento de software cada uma das atividades é composta de várias sub-atividades. A partir desta consideração, definiu uma sub-atividade específica, chamada *Investigação do Domínio*, cujo objetivo é a realização da pesquisa e o uso do conhecimento do domínio necessário na execução da atividade correspondente. A *Investigação do Domínio* deve então ser incluída nas várias atividades do processo, sempre que se julgue necessário o entendimento do domínio para a realização da mesma.

Segundo o esquema definido e apresentado pela figura 5.11 (b), pode-se verificar que para a construção de ADSOD, a sub-atividade *Investigação do Domínio*, deve ser introduzida como atividade na especialização do processo padrão, visando garantir que todos os processos instanciados considerarão a orientação ao domínio na execução dos projetos específicos.

Esta utilização da atividade *Investigação do Domínio* no nível do processo especializado, ao invés de ser considerada no processo padrão, é devido à existência de Teorias do Domínio não serem comum nas organizações, e o seu processo de construção não estar incluído no processo padrão definido (seção 3.2). Desta forma, a orientação ao domínio só será considerada se existir previamente uma Teoria do Domínio disponível. Caso a organização queira incluir a construção da Teoria do Domínio no seu processo de construção de bases de dados, poderá acrescentar as atividades do processo definido (figura 5.1) ao processo padrão (figura 3.2), e a *investigação do domínio* será incluída na especialização do processo.

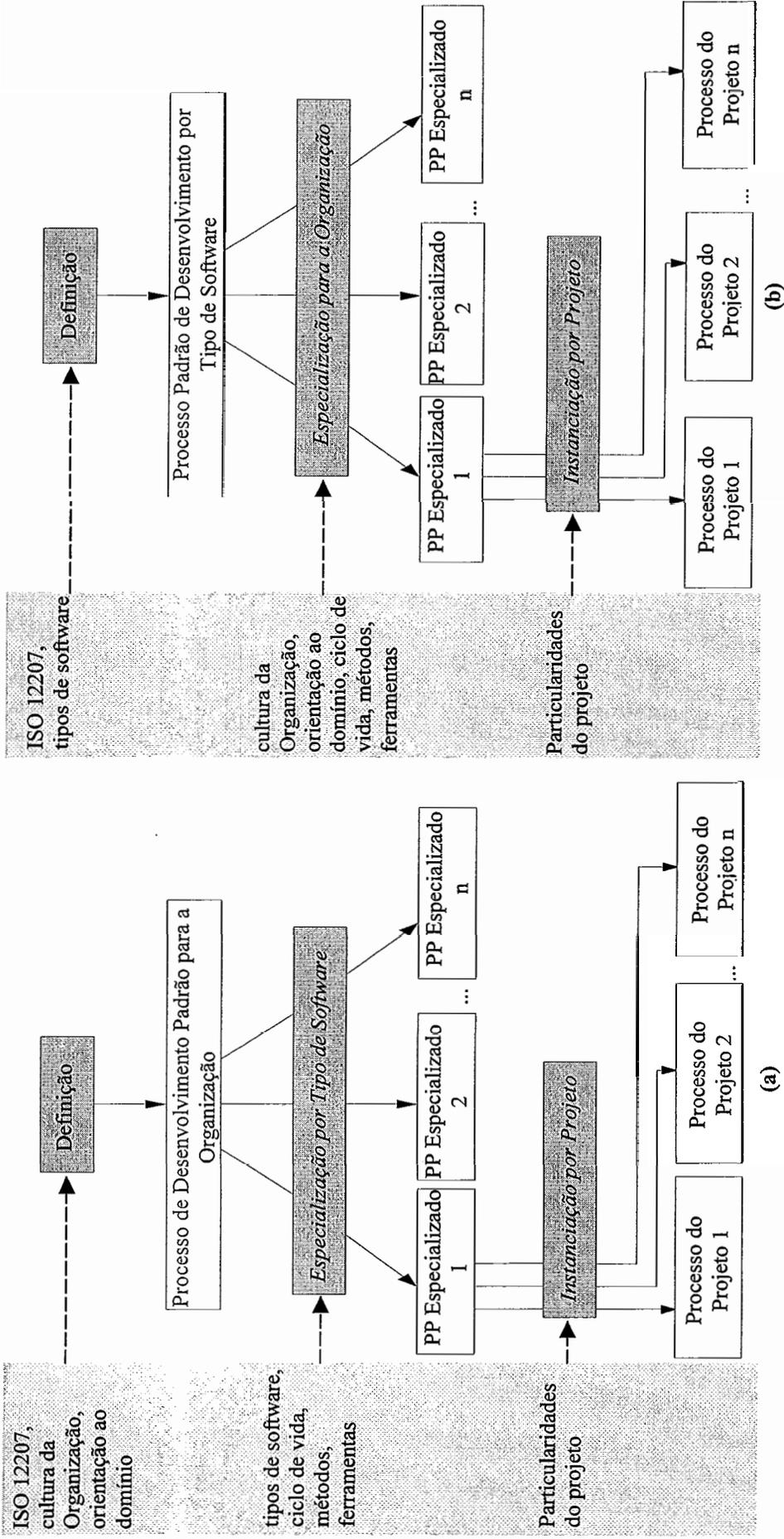


Figura 5.11 – Possíveis formas de definição, especialização e instanciação de processo de software para um ADSOD na Estação TABA (OLIVEIRA, 1999, SILVA, 2000)

5.3.1 – Especialização do Processo para o Ambiente *Insecta*

A primeira etapa para a instanciação do ambiente *Insecta* foi a especialização do processo padrão definido na seção 3.2, através da inclusão da sub-atividade Investigação do Domínio nas atividades onde se identificou como necessária, a utilização do conhecimento disponível na Teoria do Domínio para apoiar a sua execução. As atividades identificadas estão representadas na figura 5.12. Ressaltamos que a sub-atividade investigação do domínio é realizada pela equipe de informática envolvida no projeto.

Na fase de instituição do projeto (I), identificamos como necessária a investigação do domínio na atividade planejamento inicial do projeto e análise de viabilidade (I.4), devido ser necessária a existência de um conhecimento mínimo do domínio para que a análise de viabilidade social, de mão-de-obra, ambiental e legal possam ser executadas. Da mesma forma, para a formação da equipe do projeto (I.5), deve ser considerado o perfil dos profissionais em relação ao conhecimento do domínio, bem como a facilidade de seu aprendizado.

Na fase de definição (II), o conhecimento do domínio é necessário para a execução de ambas atividades (II.1 e II.2) que constituem a fase, pois, ao elaborar o Plano do Projeto, essa investigação é essencial para o entendimento do objetivo e da complexidade do domínio. No que se refere à especificação de requisitos, consideramos que o bom entendimento do domínio é crucial para a qualidade da base de dados construída e para a adequação aos requisitos do usuário. Na fase de desenvolvimento (III), a investigação do domínio será necessária à execução das atividades III.1, III.2 e III.3. Durante a fase de transição (IV) não foi identificada nenhuma atividade que necessite da inclusão da investigação do domínio para sua execução. Já na fase de manutenção (V), a sub-atividade investigação do domínio deverá fazer parte da execução das atividades V.1 e V.3. As demais atividades devem ser executadas de acordo com o que foi definido no processo padrão.

Após a definição das atividades onde a investigação do domínio pode ser incluída, a sub-atividade investigação do domínio é inserida na Estação TABA, constituindo as atividades definidas no processo padrão, que serão utilizadas na instanciação do ambiente *Insecta* (OLIVEIRA, 1999).

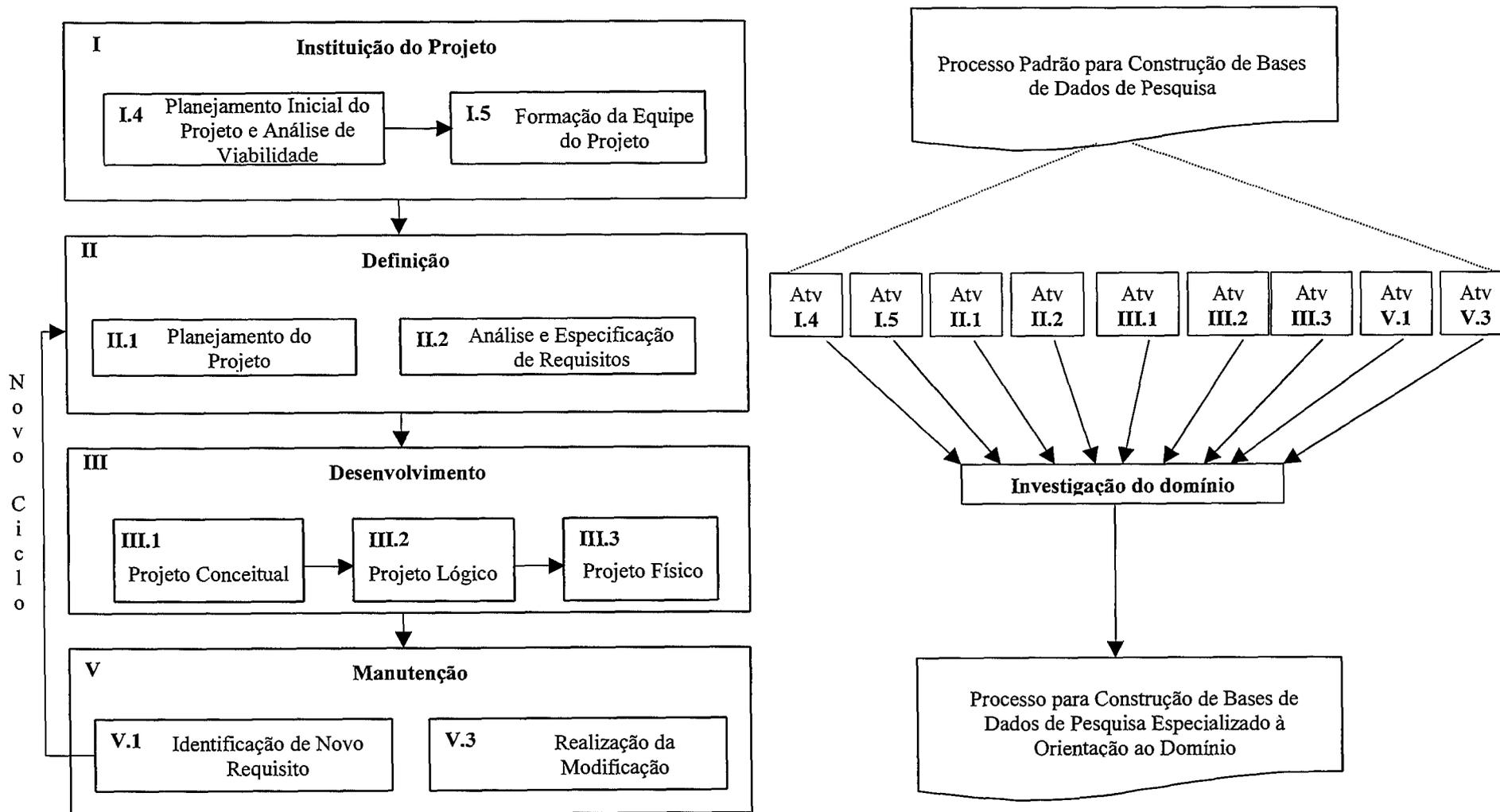


Figura 5.12 – Identificação das atividades do Processo Padrão para Construção de Bases de Dados de Pesquisa, onde se incluiu a sub-atividade Investigação do Domínio

A figura 5.13 apresenta a tela da ferramenta EDIT-PRO com a edição das características da atividade *Projeto Conceitual*. Nela podemos observar o nível de detalhamento das características que a ferramenta permite sejam inseridas para as diversas atividades que compõem o processo.

A imagem mostra a interface de usuário da ferramenta EDIT-PRO. No topo, há uma barra de menu com as opções: 'Artefatos', 'Métodos', 'Técnicas' e 'Diretrizes'. Abaixo disso, há uma barra de sub-atividades com 'Sub-atividades' selecionado e 'Mapeamento Norma ISO 12207'. A barra de recursos contém 'Geral', 'Recursos Humanos', 'Recursos Software' e 'Recursos Hardware'. O formulário principal contém os seguintes campos:

- Nome:** Projeto Conceitual
- Descrição:** Elaborar e avaliar o modelo conceitual de alto nível, independente do SGBD a ser utilizado, da base de dados a ser construída.
- Tipo:** Construção
- Granularidade:** Atividade de nível intermediário
- Origem:** (Gênérica)
- Organização:**
- Modelo:**
- Maturidade:**
- Nível:**
- Maturidade:**
- Processo de Software:** Documentação

Na base do formulário, há quatro botões: 'OK', 'Cancelar', 'Aplicar' e 'Ajuda'.

Figura 5.13 – Edição da atividade Projeto Conceitual através do uso da ferramenta EDIT-PRO

5.3.2 – Edição da Teoria do Domínio para Entomologia

A edição da Teoria do Domínio na Estação TABA é realizada através da ferramenta EDITED. Esta ferramenta permite a inserção de todas as informações relativas aos conceitos e sub-teorias; além da inserção de relações entre conceitos e de relações entre propriedades destes conceitos, a partir da entrada de restrições escritas de maneira informal e de axiomas descritos em Prolog. Ou seja, a ferramenta considera as sub-teorias, conceitos, características desses conceitos e suas relações com as tarefas da Teoria do Domínio (OLIVEIRA, 1999; OLIVEIRA *et al.*, 1999a; OLIVEIRA *et al.*, 2000b).

A segunda etapa para a instanciação do ambiente *Insecta* foi a edição da Teoria do Domínio definida na seção 5.2 através da ferramenta EDITED. As figuras 5.14 (a), (b), (c) e (d) mostram as telas ilustrando algumas das opções da ferramenta EDITED utilizadas na edição do conteúdo da Teoria do Domínio de Entomologia.

Na figura 5.14 (a) podemos observar as informações que podem ser inseridas na Estação TABA através da ferramenta EDITED, relacionadas com a Teoria do Domínio: Subteoria, Mapeamento, Conceito e Relação entre Conceitos e entre Propriedades.



Figura 5.14 (a) – Opções do menu Inserir da ferramenta EDITED

Na figura 5.14 (b) podemos observar que para o conhecimento sobre Subteorias, podemos inserir seu nome e descrição, e fazer o seu mapeamento com as tarefas realizadas no domínio.

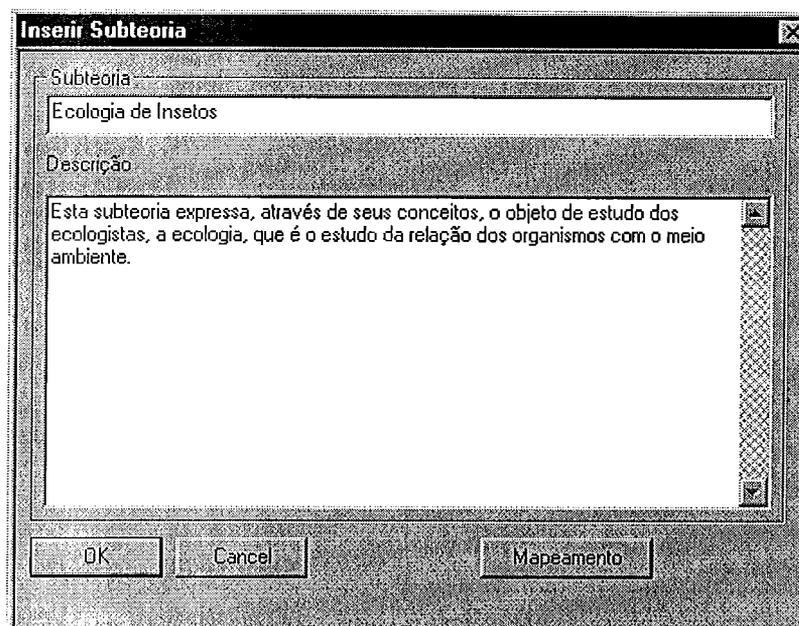


Figura 5.14 (b) – Inserção de subteorias na ferramenta EDITED

Na figura 5.14 (c) verificamos que a ferramenta permite incluir Conceitos da Teoria do Domínio com sua descrição, sinônimo e termo em inglês, além de relacioná-los com a subteoria a qual pertencem, com suas propriedades e com as bibliografias que os citam.

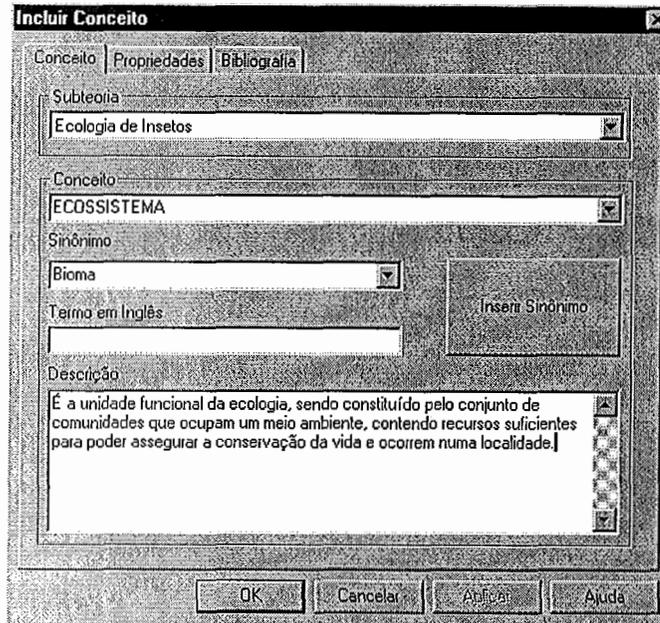


Figura 5.14 (c) – Inserção de Conceitos na ferramenta EDITED

Na figura 5.14 (d) podemos verificar que a ferramenta permite incluir as Relações entre os Conceitos da Teoria do Domínio com sua descrição e cardinalidades.

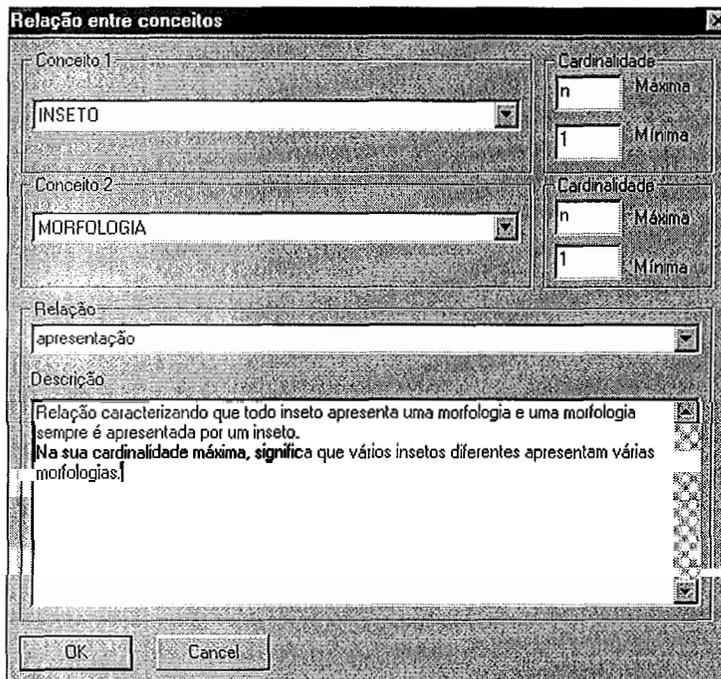


Figura 5.14 (d) – Inserção de Relações na ferramenta EDITED

5.3.3 – Instanciação do ADSOD *Insecta*

Para instanciar o ambiente foi utilizada a ferramenta para instanciação de ADS, disponível na Estação TABA (OLIVEIRA, 1999, ZLOT e SANTOS, 1999), onde são fornecidas as características do ambiente. As figuras 5.15 (a), (b) e (c) apresentam as telas de acesso à opção Ambiente e as telas de definição e instanciação do ambiente *Insecta*.

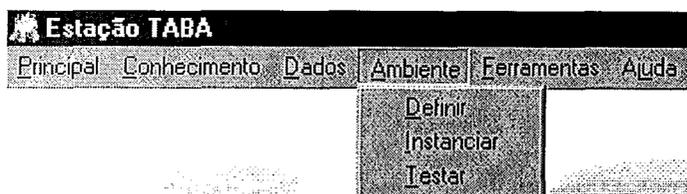


Figura 5.15 (a) – Opções da Ferramenta para Instanciação de ADS

Na figura 5.15 (a) verificamos que a ferramenta permite a definição, instanciação e teste de ADS.

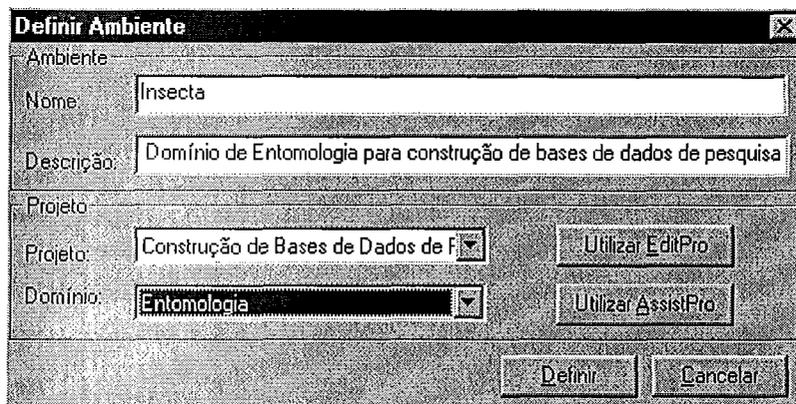


Figura 5.15 (b) – Definição do ambiente *Insecta*

Na figura 5.15 (b) verificamos que ao definir o ambiente *Insecta*, além de informarmos seu nome, descrição e projeto para o qual está sendo definido, selecionamos o domínio no qual o projeto está inserido. Ao selecionar o domínio, estaremos tornando disponível a Teoria do Domínio construída previamente, no ambiente definido. Devemos ainda utilizar uma das ferramentas para definição de processo e definir ou selecionar as atividades previamente definidas referentes ao projeto que será executado. Ao definir o ambiente, o mesmo será incorporado à Estação TABA, estando pronto para ser instanciado.

Para se instanciar um ambiente, conforme ilustrado pela figura 5.15 (c), é necessário escolher um ambiente previamente definido e ainda não instanciado e indicar um diretório onde o executável deverá ser salvo. Opcionalmente, pode-se escolher um ícone para a aplicação e solicitar que seja feita uma cópia do código fonte do ambiente instanciado e deve-se informar ao sistema a localização das classes do *framework* e do arquivo “*vcvars32.bat*”. O executável gerado é perfeitamente funcional, podendo ser executado a partir da Estação TABA ou externamente.

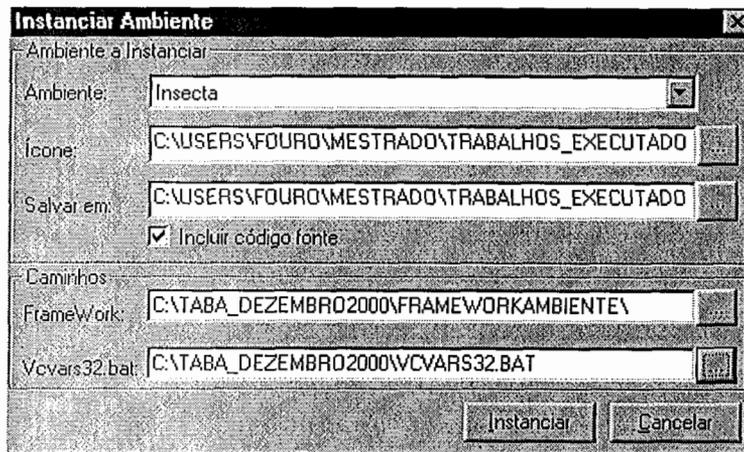


Figura 5.15 (c) – Instanciação do ambiente *Insecta*

A figura 5.16 mostra o ADSOD *Insecta* com as atividades que serão executadas para o projeto construção de bases de dados de pesquisa. No detalhe podemos observar as sub-atividades do Projeto Conceitual e as ferramentas que poderão ser utilizadas na execução da investigação do domínio.

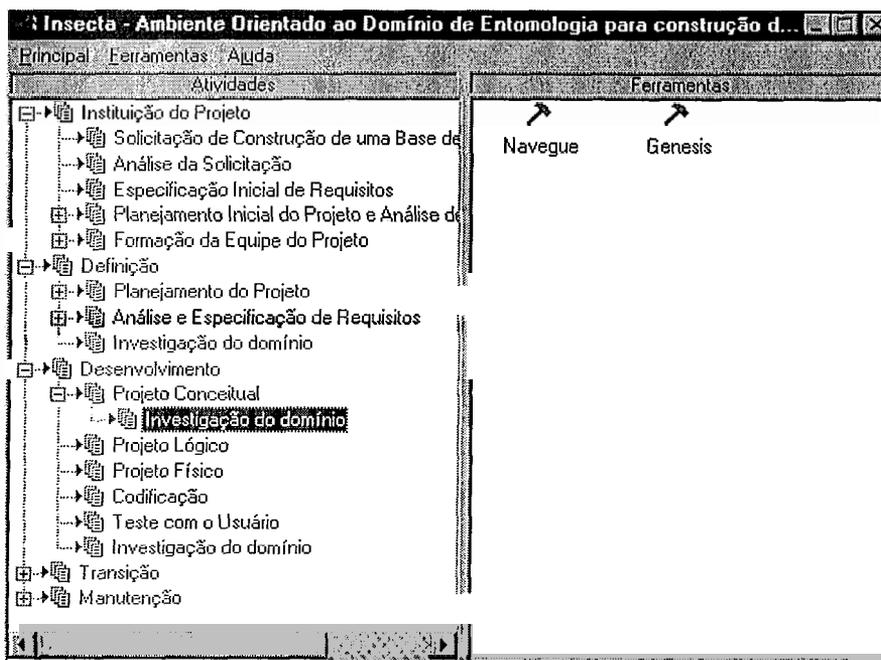


Figura 5.16 – Ambiente *Insecta*

5.4 – Exemplo da Execução da Atividade Projeto Conceitual para uma Base de Dados no Domínio de Entomologia

Projetos de pesquisa necessitam de recursos para sua realização. Atualmente, a decisão da aplicação dos recursos públicos ou privados que os financiam é feita de forma racional baseada em estratégia e informação, e considera, dentre outras coisas, quem será beneficiado com os resultados da pesquisa (CASTRO *et al.*, 1999). Por este motivo, o entomólogo, tal como qualquer pesquisador, necessita planejar as suas atividades de pesquisa.

O objetivo da modelagem da base de dados a seguir é permitir o planejamento das atividades de pesquisa do entomólogo, que deverá ser feito com base nas pesquisas executadas na base de dados projetada. O planejamento se refere especificamente às atividades de controle, identificação e manejo de insetos em determinada localidade de interesse.

Nesta seção, apresentamos a modelagem de uma base de dados de pesquisa para o domínio de entomologia, através da execução da atividade Projeto Conceitual, tal como definida na seção 3.2, utilizando o procedimento definido na seção 4.3.1, com base no mapeamento apresentado na seção 4.4. Esta base de dados está sendo construída para uso no Laboratório de Entomologia da Embrapa.

5.4.1 – Especificação de Requisitos de Dados

A atividade Projeto Conceitual é feita com base na especificação de requisitos de dados. Utilizamos a abordagem OQD (seção 3.2) para definição dos objetivos, questões e dados da base de dados a ser modelada. Foram definidos três objetivos principais das pesquisas feitas à base. Para cada objetivo definimos um conjunto de questões que permitam alcançá-los. Essas questões darão origem às consultas feitas à base. A tabela 5.2 a seguir mostra os objetivos e as questões definidas.

Para responder às questões, são necessários dados, que são identificados pelos pesquisadores a partir das questões das pesquisas, e que deverão estar na base de dados para respondê-las.

Tabela 5.2 – Definição dos Objetivos e Questões

Objetivo	Questões Derivadas
I. Identificar os tipos de insetos que precisam ter formas de controle pesquisadas	a) Quais os tipos de insetos com ciclo de vida heterodinâmico que ocorrem na localidade <i>X</i> cuja densidade populacional esteja em nível de dano econômico ? b) Qual a diversidade de espécies da ordem <i>Collembola</i> que apresentam apêndice torácico saltatório ? c) Qual a diversidade de espécies que desempenham seu nicho ecológico na localidade <i>X</i> podem ser controladas através do controle biológico ?
II. Identificar os tipos de insetos que podem ser utilizados na polinização de espécies florestais comerciais que deverão ter formas de manejo pesquisadas	d) Quais os tipos de insetos que ocorrem na localidade <i>X</i> e que possuem como uma das características de seu nicho ecológico a prática da polinização ? e) Qual a diversidade de espécies na comunidade de abelhas que constitui o ecossistema floresta na Amazônia ? f) Qual o percentual de espécies da ordem Phasnatodea que possuem apêndice móvel do tipo filiforme ?
III. Identificar os tipos de insetos que necessitam sejam feitos trabalhos de pesquisa para sua identificação	g) Qual a incidência de trabalhos científicos sobre a espécie <i>Cycloneda sanguinea</i> (L.) ? h) Qual o percentual de ocorrência de espécimes do inseto <i>X</i> nas coleções ? i) Dentro da ordem <i>Dípteros</i> qual a incidência de espécies benéficas em relação a espécies pragas ? j) Qual o percentual de gêneros que apresentam a reprodução do tipo pedogênese e ocorrem na localidade <i>X</i> ?

A figura 5.17 a seguir mostra a estrutura OQD definida para a base de dados no domínio de entomologia, onde podemos observar a identificação de cinco categorias de dados que irão responder às questões definidas na tabela 5.2. Os dados morfológicos são todos os dados relacionados com a anatomia dos insetos utilizados para sua identificação (seção 5.2.2.1). Os dados biológicos estão relacionados com o ciclo de vida dos insetos (seção 5.2.2.2). Dados ecológicos são aqueles que descrevem o comportamento dos insetos no ecossistema em que vivem (seção 5.2.2.3). Dados taxonômicos são os dados sobre a identificação taxonômica dos insetos (seção 5.2.2.4) e os dados complementares estão relacionados com as atividades de identificação, manejo e controle, executadas pelos entomólogos (seção 5.2.2.5) e que precisam ser planejadas. A partir destas definições, apresentaremos na seção seguinte a execução da atividade projeto conceitual.

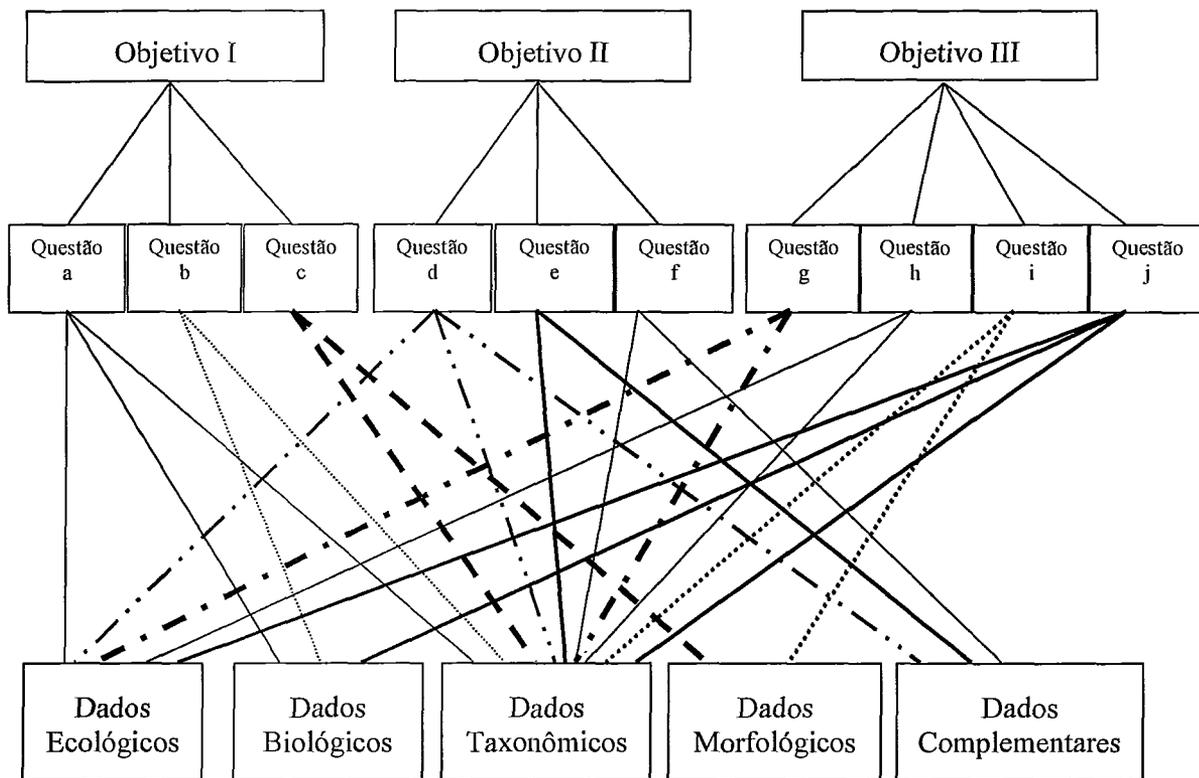


Figura 5.17 – Definição dos Dados que devem responder às Questões

5.4.2 – Projeto Conceitual

Inicialmente, o projetista deve analisar a especificação de requisitos de dados (seção 5.4.1) e executar a *investigação do domínio* com vistas a identificar quais os construtores disponíveis na Teoria do Domínio (seção 5.2) que podem ser utilizados para apoiar a execução do projeto conceitual. Essa investigação pode ser feita analisando-se as questões de competência da Teoria do Domínio e identificando quais as sub-teorias disponíveis que as respondem.

A partir da análise da especificação de requisitos de dados (abordagem OQD) e do conhecimento disponível na forma de Teoria do Domínio, o projetista é capaz de identificar que partes (sub-teorias) estão mais diretamente relacionadas com a base de dados a ser modelada. Neste caso, serão utilizados dados de todas as sub-teorias que compõem a Teoria do Domínio de entomologia na forma de árvores conceituais (figuras 5.3, 5.4, 5.5, 5.6, 5.7 e 5.10), vocabulário do domínio (Anexo 1), axiomas em linguagem natural e atributos de instâncias (Anexo 1).

Após a identificação do que, dentro do conhecimento disponível, será utilizado, podemos dar início à execução do projeto conceitual com base em ontologias de domínio, tal como abordado na seção 4.3, iniciando pelo procedimento definido na seção 4.3.1 e considerando o mapeamento disponível na seção 4.4.

Ao executar a atividade I do procedimento - *Modelagem de Objetos* - teremos identificado as entidades que irão compor o diagrama. Com base no mapeamento estabelecido na seção 4.4 temos que os conceitos disponíveis na Teoria do Domínio correspondem às entidades no DER. A figura 5.18 mostra a identificação das entidades que devem responder cada uma das questões estabelecidas através da abordagem OQD definida para a base de dados. Nela podemos observar a identificação de dezoito entidades diferentes.

- ! Questão a: *Inseto, Ciclo de Vida, Localidade e População*
- ! Questão b: *Inseto, Localidade e Nicho Ecológico*
- ! Questão c: *Inseto, Espécie, Ordem e Apêndice Torácico*
- ! Questão d: *Inseto, Espécie, Nicho Ecológico, Localidade e Controle Biológico*
- ! Questão e: *Inseto, Trabalho Científico e Espécie*
- ! Questão f: *Inseto, Espécime, Classificação e Coleção*
- ! Questão g: *Inseto, Ordem e Espécie*
- ! Questão h: *Inseto, Espécie, Comunidade e Ecossistema*
- ! Questão i: *Inseto, Ordem, Espécie e Apêndice Móvel*
- ! Questão j: *Inseto, Gênero, Reprodução e Localidade*

Figura 5.18 – Identificação das Entidades que devem responder às Questões

A tabela 5.3 a seguir relaciona as entidades identificadas com os dados que respondem as questões identificadas na abordagem OQD. A entidade inseto corresponde a todos os dados em conjunto.

Tabela 5.3 – Relação entre os Dados da Abordagem OQD e as entidades identificadas

Dados	Entidades Derivadas
Ecológicos	Localidade, população, nicho ecológico, comunidade e ecossistema
Biológicos	Ciclo de vida e reprodução
Taxonômicos	Classificação, espécie, ordem, gênero
Morfológicos	Apêndice torácico e apêndice móvel
Complementares	Controle biológico, trabalho científico, espécime e coleção

O entendimento dos objetos (entidades) identificados bem como de todos os demais componentes do modelo, dentro do ADSOD *Insecta*, é obtido através do uso da ferramenta *Navegue* (SILVA, 2000) na execução da atividade *Investigação do Domínio* inserida na especialização do processo (seção 5.3.1). Isto permite que o projetista consulte o especialista apenas para sanar dúvidas que persistam após consulta ao conhecimento disponível na Teoria do Domínio.

A execução da atividade II do procedimento - *Dicionarização dos Objetos Modelados* - utiliza o vocabulário disponível na Teoria do Domínio. A partir deste conhecimento, o projetista acrescenta as informações que são importantes para compor o dicionário de dados e que não estejam disponíveis na Teoria do Domínio. A figura 5.19 mostra a dicionarização da entidade Inseto, desenvolvida pelo projetista com base no entendimento obtido com a investigação do domínio. Nela podemos observar que a descrição da entidade foi totalmente refeita, em relação à descrição do conceito inseto na Teoria do Domínio.

Dicionário de Dados

Entidade
Nome: Inseto
Descrição: É todo animal pertencente ao filo Arthropoda, englobando apenas a classe <i>Insecta</i> , em qualquer estágio do seu ciclo de vida, que de acordo com seu comportamento em relação ao foco de pesquisa do entomólogo pode ser considerado benéfico ou praga.
<small>OBS.: A descrição deve responder às perguntas: O que é ? / O que faz ? / Para que serve ? / O que engloba essa categoria de objetos ? / O que está excluído dessa categoria de objetos ? / Quando alguém ou algo passa a ser, ou deixa de ser, um objeto dessa categoria ? / Sua permanência nessa categoria é imutável ? / Quais estados esta categoria de objetos possui durante seu ciclo de vida ?</small>

Figura 5.19 – Dicionarização da Entidade Inseto que comporá o MER

A partir das propriedades de instâncias disponíveis no vocabulário do domínio e do que foi definido através da abordagem OQD da base de dados, o projetista executa a atividade III do procedimento - *Modelagem dos Atributos de Objetos* - definindo os atributos das entidades identificadas.

A figura 5.20 exemplifica a execução da atividade através da identificação dos atributos da entidade Inseto. O atributo nome é proveniente diretamente da Teoria do Domínio, mas o atributo tipo somente foi identificado como necessário na execução desta atividade.

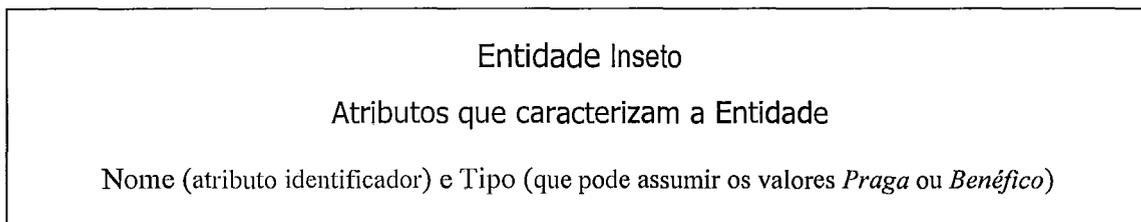


Figura 5.20 – Identificação dos Atributos da Entidade Inseto

A atividade IV do procedimento - *Dicionarização dos Atributos de Objetos* - tem seu resultado obtido a partir da análise, pelo projetista, do que foi estabelecido na Teoria do Domínio na forma de definição das propriedades de instâncias, acrescentando as informações que julgar sejam relevantes para compor o MER da base de dados sendo projetada.

A figura 5.21 mostra um exemplo de como os atributos de uma entidade serão registrados no dicionário de dados. A descrição do atributo nome é proveniente da Teoria do Domínio, e a partir da descrição do atributo tipo, poderia haver uma interação com os especialistas para verificar se classificação econômica poderia ser um conceito importante na sub-teoria ecologia de insetos.

Dicionário de Dados

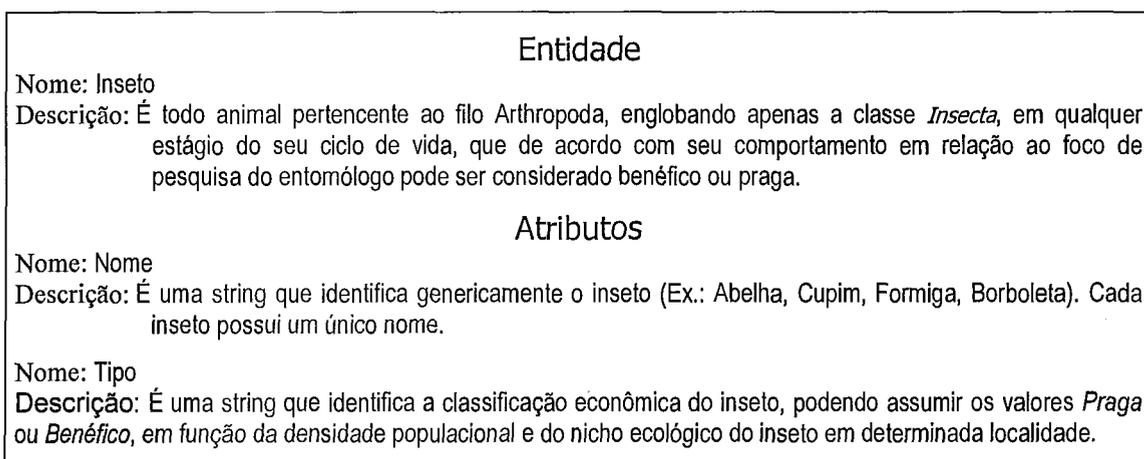


Figura 5.21 – Dicionarização dos Atributos da Entidade Localidade

A atividade V do procedimento - *Modelagem das Estruturas de Generalização* - é feita considerando as entidades identificadas na atividade I - *Modelagem de Objetos*. O projetista pode observar que a Teoria do Domínio, através da sub-teoria definida na seção 5.2.2.4, define que os conceitos Ordem, Gênero e Espécie são subtipos (classificações) do conceito *Categoria Taxonômica* e, pelo que estabelece o mapeamento da seção 4.4, esta estrutura também existe na modelagem ER, tendo como característica em comum o atributo identificador nome científico.

Outra possível generalização está relacionada com as entidades Apêndice Torácico e Apêndice Móvel que apresentam em comum os atributos nome, localização, função, quantidade, tamanho e tipo, e como entidade geral teremos a entidade *Apêndice* (proveniente da Teoria do Domínio).

De acordo com a sub-teoria definida na seção 5.2.2.5 e com o mapeamento da seção 4.4 as entidades trabalho científico, espécime, coleção e classificação seriam especializações de uma entidade geral *produto atividades* (não proveniente da Teoria do Domínio) com o atributo nome em comum, mas para o caso desta base não será necessário este mapeamento. Da mesma forma, a entidade controle biológico seria uma especialização de uma entidade geral *atividade de controle*.

A dicionarização das estruturas de generalização também deve ser feita como parte da execução desta atividade. A figura 5.22 mostra um exemplo de uma entrada no dicionário de dados de uma das estruturas de generalização identificadas. Nela podemos observar o atributo atuação do agente, que não é proveniente diretamente da Teoria do Domínio, pois foi identificado como necessário, na execução desta atividade.

Dicionário de Dados

Generalização/Especialização	
Entidade Genérica: Atividade de Controle	
Descrição: É todo procedimento executado para manter a população de insetos pragas abaixo do nível de dano econômico. Esta entidade identifica o tipo de controle que será aplicado à população de insetos, cuja densidade atinja o nível de dano.	
Atributos	
Nome: Nome	
Descrição: Atributo identificador, que determina o controle que será aplicado à população de insetos.	
Entidade Especializada: Controle Biológico	
Descrição: É um tipo de controle que utiliza organismos vivos (predadores, parasitas e competidores) ou produtos biológicos (hormônio) para controlar pragas agrícolas ou de outro tipo.	
Atributos	
Nome: Nome do Agente	
Descrição: Atributo identificador da entidade que determina o agente utilizado como inimigo natural ou agente causador de doença na população de inseto a ser controlada.	
Nome: Atuação do Agente	
Descrição: Determina o efeito causado pelo agente utilizado como inimigo natural ou agente causador de doença na população de inseto a ser controlada.	

Figura 5.22 – Dicionarização de uma estrutura de Generalização/Especialização

Os relacionamentos em um modelo, além da existência de associações naturais entre as entidades, podem surgir em função das exigências e necessidades de recuperação de informações por parte do usuário que ocasionam a existência de outras associações entre as entidades.

Conforme o que foi estabelecido na seção 4.3 e pelo mapeamento da seção 4.4, os relacionamentos correspondem às relações identificadas na Teoria do Domínio e são estabelecidos entre as entidades simples, genéricas e especializadas já identificadas.

Como resultado da execução da atividade VI do procedimento - *Modelagem dos Relacionamentos entre Objetos* - teremos o conjunto de relacionamentos que respondem às questões da abordagem OQD definida para a base de dados, com suas características que serão documentadas no dicionário de dados.

Na tabela 5.4 mostramos o resultado da execução da atividade, mapeando os relacionamentos com as questões definidas. Os relacionamentos provenientes diretamente da Teoria do Domínio são: 2, 3, 9, 12, 13 e 19 (sub-teoria ecologia de insetos); 4, 6 e 7 (sub-teoria anatomia de insetos); 1 e 18 (sub-teoria biologia de insetos); 5, 8 e 20 (sub-teoria classificação de insetos); 10, 15, 16 e 17 (sub-teoria processo de trabalho do entomólogo). Os relacionamentos 11 e 14 não estão representados na Teoria do Domínio, e poderiam originar na mesma, um axioma ou integração entre as sub-teorias ecologia, classificação e processo de trabalho do entomólogo.

Ao executarmos a atividade VII do procedimento - *Modelagem das Estruturas de Agregação* - apesar de não termos identificado nenhum relacionamento ternário ou de ordem superior que originasse uma estrutura de agregação, identificamos que o relacionamento existente entre controle e população é estabelecido entre a agregação do relacionamento *população – ocupa – localidade*, uma vez que o controle de uma população é feito na localidade que a mesma ocupa, e isto não é proveniente diretamente da Teoria do Domínio, tendo sido identificado este relacionamento, durante a execução desta atividade. A figura 5.23 mostra a estrutura de agregação identificada.

Tabela 5.4 – Relacionamentos que respondem às questões estabelecidas

Relacionamentos Identificados	Questões Respondidas
1. Inseto possui Ciclo de Vida	a e j
2. Inseto parte-de População	a, c, d, e e j
3. População ocupa Localidade	a, c, d e j
4. Espécie identifica Inseto	b, c, d, e, f, g e i
5. Generalização das entidades Ordem, Gênero e Espécie originando a entidade Categoria Taxonômica.	b, f, i e j
6. Generalização das entidades Apêndice Torácico e Apêndice Móvel originando a entidade Apêndice	b e f
7. Apêndice caracteriza Espécie	b e f
8. Classificação identifica Inseto	b, f, g, h, i e j
9. Inseto executa Nicho Ecológico	c e d
10. Generalização da entidade Controle Biológico originando a entidade Atividade de Controle	c
11. Controle atua População	c
12. População parte-de Comunidade	e
13. Comunidade parte-de Ecossistema	e
14. Trabalho Científico descreve Espécie	g
15. Espécime representa Inseto	h
16. Espécime possui Classificação	h
17. Espécime parte-de Coleção	h
18. Reprodução parte-de Ciclo de Vida	j
19. Comunidade designa Inseto	e
20. Categoria Taxonômica parte-de Classificação	b, f, g, h, i e j

A execução da atividade VIII do procedimento - *Dicionarização de Relacionamentos e seus Atributos* - resulta na descrição dos relacionamentos e seus atributos bem como das estruturas de agregação identificadas, que irão compor o dicionário de dados. Esta dicionarização, como todas as outras, considera o conteúdo da Teoria do Domínio (Anexo 1). A figura 5.24 mostra um exemplo de como serão as entradas no dicionário de dados para estes construtores, onde ambas as descrições foram elaboradas durante a execução da atividade.

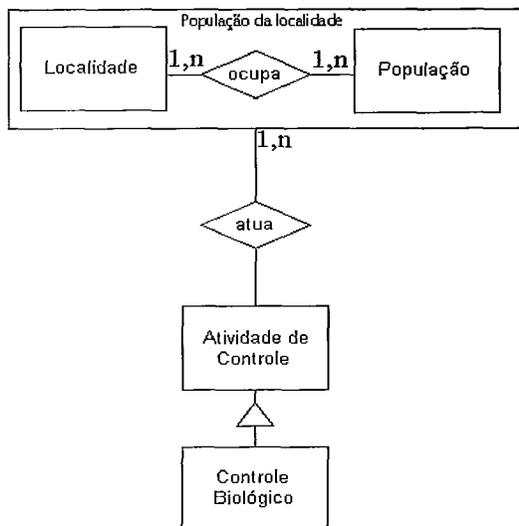


Figura 5.23 – Identificação da estrutura de agregação

Dicionário de Dados

Relacionamento	
Nome: caracteriza	
Descrição: Relacionamento independente entre as entidades gerais, Categoria taxonômica e Apêndice, significando que o apêndice, como uma característica morfológica do inseto, caracteriza as categorias taxonômicas nas quais o inseto está inserido. Sua função é estabelecer a ligação entre as entidades, sempre que o sistemata tiver observado que o apêndice caracteriza a categoria taxonômica.	
OBS.: A descrição deve estabelecer: Sua função / O que ele representa / Quais são as regras de seu estabelecimento / Quais são as exceções a seu estabelecimento / Quando ocorre / Quando pode deixar de existir	
Cardinalidade Mínima: 1	Cardinalidade Máxima: N
Atributos	
Para este relacionamento não foi identificado nenhum atributo.	
Agregação	
Nome: População da localidade	
Descrição: Agrega as entidades <i>População</i> e <i>Localidade</i> , relacionadas através do relacionamento <i>ocupa</i> , significando que o controle de uma população é feito na localidade que ela ocupa.	
A descrição deve estabelecer: Quais as condições em que existe o relacionamento e as entidades agregadas.	

Figura 5.24 – Dicionarização de relacionamentos e estruturas de agregação

O produto da execução da atividade IX do procedimento - *Representação no Diagrama* - é apresentado na figura 5.25, que ilustra a integração dos construtores identificados e dicionarizados, de forma que o diagrama gerado seja montado da forma mais clara possível, com base no entendimento que o projetista tem do domínio, através do que está disponível na Teoria do Domínio e das novas informações identificadas com a execução do procedimento.

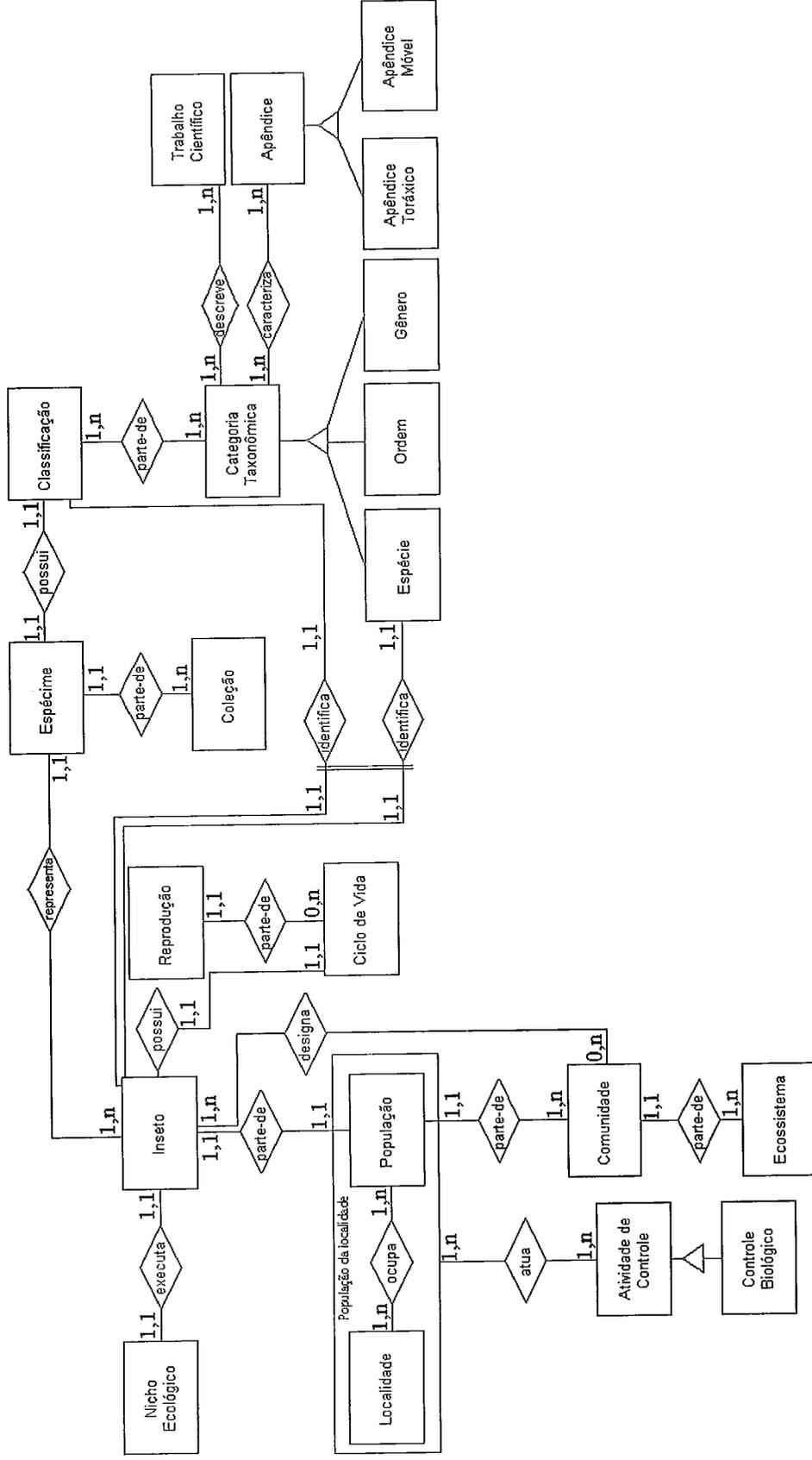


Figura 5.25 – Diagrama ER inicial da base de dados

A partir do modelo inicial (diagrama + dicionário de dados), executa-se a atividade X do procedimento - *Refinamento do Modelo Conceitual* - em busca de um modelo o mais próximo da realidade do domínio que está sendo observado, de forma que o mesmo seja completo, de fácil leitura, correto, que expresse e represente todos os requisitos de forma natural e seja facilmente adaptável a mudanças de requisitos.

Mesmo com a definição de um procedimento a ser seguido (seção 4.3.1), do estabelecimento de como ontologias apóiam a execução do procedimento (tabela 4.2) e do mapeamento estabelecido na seção 4.4, o saber interpretar o que um dado caracteriza ou a que este dado se refere, é importante para o resultado correto do modelo entidade-relacionamento. Todo tipo de modelagem sempre refletirá a interpretação de um ser humano, da realidade que o cerca.

Neste contexto, a real utilização da Teoria do Domínio é permitir um bom entendimento desta realidade, possibilitando agilizar a construção do MER em função deste entendimento prévio, facilitando e diminuindo a interação projetista-especialista.

5.5 – Considerações Finais

Neste capítulo apresentamos a Teoria do Domínio de Entomologia, construída segundo o processo definido por OLIVEIRA (1999). Especializamos o processo padrão definido na seção 3.2 para que seja possível o uso do conhecimento disponível na forma de Teoria do Domínio, pelas atividades que dele necessitem.

Tanto o conhecimento do processo especializado quanto da Teoria do Domínio, foram inseridos na infraestrutura da Estação TABA para que pudéssemos construir o ADSOD *Insecta*, seguindo as etapas definidas por OLIVEIRA (1999) ao estender a Estação TABA para possibilitar a definição e construção de ADSOD para qualquer domínio desejado.

Para demonstrar o apoio do conhecimento do domínio na execução das atividades do processo padrão definido, onde quer que este conhecimento possa ser utilizado para facilitar sua execução e melhorar a qualidade dos produtos resultantes, apresentamos a execução da atividade Projeto Conceitual, onde também foi possível observamos a possibilidade de evolução do conhecimento na Teoria do Domínio, através da identificação de novos conceitos, relacionamentos e axiomas.

No próximo capítulo apresentamos nossas conclusões, resultantes pela realização deste trabalho e as perspectivas para a sua continuidade em realizações futuras.

Capítulo 6

Conclusões e Perspectivas Futuras

Ao longo dos doze anos em que desenvolvemos nossas atividades em uma instituição de C&T, podemos notar a necessidade de construção de bases de dados para armazenar os dados gerados em seus projetos de pesquisa, como forma de contribuir para uma maior utilização e cruzamento das informações disponíveis, além de permitirem a realização de pesquisas sobre os dados armazenados.

A Embrapa, mesmo sendo detentora de valioso acervo de conhecimento e tecnologia, tem encontrado dificuldades para organizá-lo e colocá-lo à disposição da sociedade, devido a diversos fatores que são comuns nas organizações e que influenciam negativamente a qualidade do produto final e a produtividade no desenvolvimento de software:

- a não existência de um processo de desenvolvimento de software definido de forma a apoiar a realização das atividades pela equipe de desenvolvimento, e que possa também ser controlado, medido e melhorado;
- profissionais responsáveis pela construção de software não capacitados e com baixo grau de conhecimento dos domínios para os quais precisam desenvolver os produtos;
- a diversidade e especificidade das áreas englobadas pelos projetos de pesquisa.

Visando superar essas dificuldades e diminuir os riscos envolvidos na execução das tarefas relacionadas com a construção de bases de dados de pesquisa em domínios com os quais o desenvolvedor não tem familiaridade, optamos por colocar em prática as recomendações feitas por OLIVEIRA (1999) através da definição de um ADSOD que se propõe a apoiar o desenvolvimento de software em domínios específicos através do uso do conhecimento deste domínio, durante todo o processo de desenvolvimento, auxiliando os desenvolvedores no entendimento do domínio do problema.

Tudo o que foi proposto ao longo deste trabalho, visa solucionar basicamente dois problemas principais:

- De Comunicação
 - Possibilitar que pessoas de áreas muito diferentes - desenvolvedores e especialistas - desenvolvam um trabalho em conjunto.
 - Permitir que novos profissionais de informática, que entrem na organização, consigam entender sobre o domínio modelado mais rapidamente.
- De Desenvolvimento
 - Disponibilizar rapidamente para os grupos de pesquisa, bases de dados de pesquisa, de forma que possam trabalhar efetivamente com as mesmas.
 - Permitir o desenvolvimento de um produto de grande utilidade com pouco recurso e de forma rápida.

Podemos destacar como contribuições deste trabalho:

- a definição de um processo de desenvolvimento de bases de dados de pesquisa, baseado na norma NBR ISO 12207 (1998), que pode ser adaptado para uso em qualquer organização através da inclusão das atividades que retratem sua cultura organizacional;
- a organização do conhecimento na forma de uma Teoria do Domínio para Entomologia, com base em ontologia do domínio;
- a organização de um roteiro que considera diretrizes para uso da ontologia na modelagem conceitual de dados;
- a avaliação efetiva da ontologia a partir de características de qualidade definidas na literatura;
- a construção, através da estação TABA, de um servidor de conhecimento para Entomologia, permitindo a execução da atividade Investigação do Domínio ao longo do processo de desenvolvimento;
- a instanciação de *Insecta*, um ADSOD para o domínio de Entomologia, definido para apoiar futuros projetos para construção de bases de dados neste domínio na EMBRAPA;

Através da realização deste trabalho, contribuímos também na validação dos resultados de pesquisas anteriores na área de ambientes de desenvolvimento de software, realizadas na COPPE/UFRJ. Isto nos permitiu observar a necessidade pela realização de outros trabalhos, para que desenvolvedores disponham de um ambiente que apóie, da forma mais completa possível, a execução das atividades presentes no processo de desenvolvimento de software considerando o conhecimento do domínio.

Especificamente para a construção de bases de dados, podemos identificar como futuras pesquisas:

- a) a definição e implementação de uma ferramenta para a geração automática do MER (DER+Dicionário) a partir da ontologia de domínio, com base no que foi estabelecido no capítulo 4;
- b) a verificação mais consistente da utilização da ontologia de domínio apoiando a execução das atividades de Projeto Lógico e Projeto Físico;
- c) a formalização e codificação dos axiomas, visando avaliar de que forma podem ser gerados pela ferramenta na forma de restrições de integridade;
- d) a evolução da Teoria do Domínio a partir de sua utilização na construção de várias bases de dados, permitindo a identificação da necessidade de novos conceitos, que poderão significar a necessidade de inclusão de novo conhecimento, melhoria do conhecimento já modelado ou mesmo a necessidade de integração com outras ontologias existentes.

A realização deste trabalho, além de verificar a existência de uma solução viável, para problemas que dificultam o êxito no desenvolvimento de software para apoio à pesquisa na Embrapa, permitiu também a validação e consolidação da proposta de OLIVEIRA (1999) para ADSOD, enfocando o uso de ontologias de domínio para um tipo de desenvolvimento específico, a construção de bases de dados de pesquisas.

Referências Bibliográficas

- BASIL, V.R., CALDIERA, G., ROMBACH, H.D., “Goal Question Metric Paradigm”, *Encyclopedia of Software Engineering*, Vol. 2, pp 528-532, 1994.
- BATINI, C., CERI, S., NAVATHE, S.B., *Conceptual Database Design: an Entity-Relationship Approach*, The Benjamin/Cummings Publishing Company Inc., edwood City, 496p, 1992.
- BENCH-CAPON, T., “Why Database AND Expert Systems ?”, *In: Proceedings of the 8th International Workshop on Database and Expert Systems Applications – DEXA’97*, pp 2-5, 1997.
- BORROR, D.J., DeLONG, D.M., *Introdução ao Estudo dos Insetos*, São Paulo, Ed. da Universidade de São Paulo, 1984.
- BOWERS, D.S., *From Data to Database*, Van Nostrand Reinhold (International), 1998.
- BRAY, T., PAOLI, J., SPERBERG, C., *Extensible Markup Language (XML) 1.0 (Second Edition)*. W3C Recommendation. Out. 2000. Disponível em: <http://www.w3.org/TR/REC-xml>>. Acesso em: 10/11/2000.
- BRICKLEY, D., GUHA, R.V., *Resource Description Framework (RDF) Schema Specification*. W3C Proposed Recommendation. Mar., 2000. Disponível em: <http://www.w3.org/TR/2000/CR-rdf-schema-20000327>>. Acesso em: 05/04/2000.
- CASTRO, A.M.G. de, LIMA, S.M.V, CARVALHO, J.R.P. de., *Planejamento de C&T – Sistemas de Informação Gerencial*, Brasília-DF, SPI-Embrapa, 1999.
- CHANDRASEKARAN, B., JOSEPHSON, J.R., BENJAMINS, V.R., “What are Ontologies, and Why do we Need Them ?” *IEEE Intelligent Systems*, Jan-Fev, pp 20-26, 1999.
- CHEN, P.P., “The Entity-Relationship Model - Toward a Unified View of Data”, *ACM Transactions on Database Systems*, Vol. 1, No. 1, Mar., pp 9-36, 1976.

- CHEN, P.P, THALHEIM, B., WONG, L.Y., “Future Directions of Conceptual Modeling”,
In: Conceptual Modeling: Current Issues and Future Directions (eds), Springer-verlag,
Berlin, Lecturing Notes in Computer Sciences, No. 1565, pp 294-308, 1998.
- CHEN, P.P., “From Ancient Egyptian Language to Future Conceptual Modeling”, *In:*
Conceptual Modeling: Current Issues and Future Directions (eds), Springer-Verlag,
Berlin, Lecturing Notes in Computer Sciences, No. 1565, pp 57-66, 1998.
- CORCHO, O.; GÓMEZ-PÉREZ, A., “Evaluating Knowledge Representation and Reasoning
Capabilities of Ontology Specification Languages”, *In: Workshop on Applications of
Ontologies and Problem-Solving Methods – ECAI-00*, Berlim, Alemanha, Ago., 2000.
- COUGO, P.S., *Modelagem Conceitual e Projeto de Banco de Dados*, 4ed., Rio de Janeiro,
Campus, 284p. 1997.
- CRANEFIELD, S; PURVIS, M., “UML as an Ontology Modelling Language”, *In:*
Proceedings of the IJCAI-99 Workshop on Intelligent Information Integration, 16th
International Joint Conference on AI. Estocolmo, Suécia, Jul. 1999.
- DATE, C.J., *Introdução a Sistemas de Banco de Dados*. Tradução [da 7^a Edição Americana]
Vandenberg Dantas de Souza, Rio de Janeiro, Campus, 2000.
- DOWELL, M.L., STEPHENS, L.M., BONNELL, R.D., “Using a Domain Knowledge
Ontology as a Semantic Gateway among Databases”, *In: Workshop on Basic
Ontological Issues in Knowledge Sharing*, IJCAI, Montreal, Canadá, 1995.
- EMAM, K. E., DROUIN, J. AND MELO W., *SPICE – The Theory and Practice of Software
Process Improvement and Capability Determination*, IEEE Computer Society Press,
1998.
- FALBO, R., MENEZES, C., ROCHA, A.R.C. da., “Using Knowledge Servers to Promote
Knowledge Integration in Software Engineering Environments”, *In: Proceedings of the
11th Software Engineering and Knowledge Engineering Conference*, pp. 170-174,
Kaiserslautern, Alemanha, Jun., pp 16-19, 1999a.

- FALBO, R., MENEZES, C., ROCHA, A.R.C. da., “Assist-Pro: Um Assistente Inteligente para Apoiar a Definição de Processos de Software”, *Anais do XIII Simpósio Brasileiro de Engenharia de Software*, pp 147-162, Florianópolis, SC, Brasil, Out., 1999b.
- FALBO, R., *Integração de Conhecimento em um Ambiente de Desenvolvimento de Software*, Tese de D. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1998.
- FARQUHAR, A., FIKES, R., RICE, J., “The Ontolingua Server: A Tool for Collaborative Ontology Construction”, *Proceedings of KAW96*, Banff, Canada, 1996.
- FENSEL, D., HORROCKS, I., HARMELEN, F.V. *et al.*, “OIL in a Nutshell”, *Workshop on Applications of Ontologies and Problem-Solving Methods – ECAI-00*, Berlim, Alemanha, Ago., pp 4.1-4.12, 2000.
- FISCHER, G., “Domain-Oriented Design Environments, Automated Software Engineering”, *The International Journal of Automated Reasoning and Artificial Intelligence in Software Engineering*, Vol 1, N 2, Jun., pp 177-203, 1994.
- FIKES, R., FARQUHAR, A., “Distributed Repositories of Highly Expressive Reusable Ontologies”, *IEEE Intelligent Systems & their applications*, v. 14, n. 2 (Mar/Abr), pp 73-79, 1999.
- FOURO, A.M.M., WERNER, C.M.L., *Modelos de Domínio e Ontologias: similares, diferentes ou complementares*, Publicações Técnicas, COPPE/UFRJ, ES 7/99, Rio de Janeiro-RJ, Brasil, Dez., 1999.
- FOURO, A.M.M., WERNER, C.M.L., “Modelos de Domínio ou Ontologias ?”, *Revista Tecnologia da Informação*, v. 2, n. 1, pp 17-24, Dez., 2000.
- GALLO, D., NAKANO, O., SILVEIRA NETO, S. *et al.*, *Manual de Entomologia Agrícola*, 2ed., São Paulo, Ed. Agronômica Ceres, 649p. il. 24p. cor, 1988.

- GALOTTA, C., ROCHA, A. R., OLIVEIRA, K. M., *Ambientes de Desenvolvimento de Software Orientado a Domínio – Definição Inicial de uma Ontologia para Acústica Submarina*, Publicações Técnicas, COPPE/UFRJ, ES 496/99, Rio de Janeiro, RJ, Brasil, 1999.
- GENESERETH, M., FIKES, R., *Knowledge Interchange Format*. Technical Report. Computer Science Department. Stanford University. Logic-92-1. 1992.
- GÓMEZ-PÉREZ, A., BENJAMINS, V.R., “Overview of Knowledge Sharing and Reuse Componentes: Ontologies and Problem-Solving Methods”, *In: Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving methods (KRR5)*, Estocolmo, Suécia, pp 1-1 – 1-15, Ago., 1999.
- GREGERSEN, H., JENSEN, C.S., “Temporal Entity-Relationship Models - a Survey”, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11 No. 3, pp 464-497, Maio/Junho, 1999.
- GUARINO, N., “Understanding, Building, and Using Ontologies”, *International Journal of Human-Computer Studies - special issue on Using Explicit Ontologies in KBS Development*, Vol. 46, No. 2/3, pp 293-310, 1997.
- GUARINO, N., “Formal Ontology and Information System”, *In: Guarino, N. (ed) Formal Ontology in Information System*, pp 3-15, IOS Press, 1998.
- GUIZZARDI, G., *Uma Abordagem Metodológica de Desenvolvimento Para e Com Reuso, Baseada em Ontologias Formais de Domínio*. Tese de M. Sc., UFES, Vitória, ES, Brasil, 2000.
- HEUSER, C. A., *Projeto de Banco de Dados*, 4ed., Porto Alegre, Ed. Sagra Luzzato, 204p, 2001.
- ISO/IEC 9126/NBR 13596, *Tecnologia de Informação – Avaliação de Produto de Software – Características de Qualidade e Diretrizes para seu uso*, ABNT – Associação Brasileira de Normas Técnicas, 1996.

- KARP, R., CHAUDHRI, V., THOMERE, J., *XOL: An XML-Based Ontology Exchange Language*. Aug., 1999. Disponível em: <<http://www.ai.sri.com/pkarp/xol/>>. Acesso em: 20/10/1999.
- KIFER, M., LAUSEN, G., WU, J., “Logical Foundations of Object-Oriented and Frame-Based Languages”, *Journal of the ACM*, 1995.
- KOLP, M., ZIMÁNYI, E., “Enhanced ER to relational mapping and interrelational normalization”, *Information and Software Technology*, Vol. 42, pp 1057-1073, 2000.
- KNUBLAUCH, H., ROSE, T., “Round-Trip Engineering of Ontologies for Knowledge-Based Systems”, In: *Proceedings of The 12th International Conference on Software Engineering & Knowledge Engineering – SEKE*, pp 239-247, Jul., 2000.
- LARA, F.M., *Princípios de Entomologia*, 3ed., São Paulo, Ícone, 1992.
- LASSILA, O., SWICK, R. R. *Resource Description Framework (RDF) Model and Syntax Specification*. W3C Proposed Recommendation. Feb., 1999. Disponível em: <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>>. Acesso em: 05/03/1999.
- LÓPEZ, M.F., GÓMEZ-PÉREZ, A., SIERRA, J.P., *et al.*, “Building a Chemical Ontology Using Methontology and the Ontology Design Environment”, *IEEE Intelligent Systems*, pp 37-46, Jan/Fev, 1999.
- LUKE, S., HEFLIN, J., *SHOE 1.01. Proposed Specification*. SHOE Project. Apr. 2000. Disponível em: <<http://www.cs.umd.edu/projects/plus/SHOE/SPEC.HTML>>. Acesso em: 10/05/2000.
- MACHADO, L.F.C., *Modelo para Definição, Especialização e Instanciação de Processos de Software*, Tese de M. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2000.
- MACHADO, F.N.R., ABREU, M.P., *Projeto de Banco de Dados - Uma Visão Prática*, 5ed., São Paulo, Érica, 297p. 1996.

- MAIDANTCHIK, C., *Modelo de Processo de Gerência de Software para Equipes Geograficamente Dispersas*, Tese de D. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1999.
- MCLELLAN, T., *Data Modelling: Finding the Perfect Fit*, 1995. Disponível em: <<http://www.islandnet.com/~tmc/html/articles/datamodl.htm>>. Acesso em: 20/10/1999.
- MELLO, R.S., HEUSER, C.A., “Aplicação de Ontologias a Dados Semi-Estruturados”, In: *CLEI 2000*. México, Actas, ITESM, 2000. Disponível em: <<http://www.inf.ufrgs.br/~ronaldo/a000139.pdf>>. Acesso em: 20/11/2000.
- MILTON, S., KAZMIERCZAK, E., “Enriching the Ontological Foundations of Modelling in Information Systems”, In: *Proceedings of the Information Systems Foundations Workshop Ontology, Semiotics and Practice*, 1999.
- MOTTA, E., *Reusable Components for Knowledge Modelling*. IOS Press. Amsterdam, 1999.
- MOURA, L. M. V., ROCHA, A. R. C. da, *Ambientes de Desenvolvimento de Software*, Publicações Técnicas COPPE/UFRJ, ES-271/92, Rio de Janeiro, Brasil, 1992.
- NBR ISO/IEC 12207, *Tecnologia de Informação – Processos de Ciclo de Vida de Software*, Associação Brasileira de Normas Técnicas, Rio De Janeiro, Brasil, 1998.
- OLIVEIRA, K.M., SANTOS, G., ZLOT, F., *et al.*, “Construção de Ambientes de Desenvolvimento de Software Orientados a Domínio na Estação TABA”, *Anais do Terceiro Workshop Ibero-americano de Engenharia de Requisitos e Ambientes de Software – IDEAS’00*, Cancun, México, Abr., 2000a.
- OLIVEIRA, K.M., MENEZES, C., ROCHA, A.R.C. da., *et al.*, “Ambiente de Desenvolvimento de Software Orientado a Domínio”, *Anais do XIV Simpósio Brasileiro de Engenharia de Software*, pp 275-290, João Pessoa, PB, Out., 2000b.
- OLIVEIRA, K.M., *Modelo para Construção de Ambientes de Desenvolvimento de Software Orientados a Domínio*, Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1999.

- OLIVEIRA, K.M., ROCHA, A.R.C., TRAVASSOS, G.H. *et al.*, “Using Domain-Knowledge in Software Development Environments”, *In: Proceedings of the 11th International Conference on Software Engineering and Knowledge Engineering*, pp. 180-187, Kaiserslauter, Alemanha, Jun. 1999a.
- OLIVEIRA, K. M., ROCHA, A. R., TRAVASSOS, G. H. *et al.*, “O uso da Teoria do Domínio no Processo de Desenvolvimento de Software”, *In: Anais da X Conferencia Internacional de Tecnologia de Software*, pp 223-235, Curitiba, Brasil, Mai., 1999b.
- PAPAVERO, N., *Fundamentos práticos de taxonomia zoológica: coleções, bibliografia, nomenclatura*. Museu paraense Emílio Goeldi e Sociedade Brasileira de Zoologia. Belém-PA, 1983.
- PARSONS, J., WAND, Y., “Using Objects for Systems Analysis”, *Communications of the ACM*, Vol 40, No. 12, pp. 104-110. Dec, 1997.
- PAULK, M. C., WEBER, C. V., CURTIS, B., *et al.* (eds), *The Capability Maturity Model: Guidelines for Improving the Software Process*, Carnegie Mellon University, Software Engineering Institute, Addison-Wesley Longman Inc. 1997.
- PINHO, S.F.C., *Avaliação da Qualidade de Dados pela não Conformidade*, Tese de Mestrado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, Abr., 2001.
- PINTO, H.S., GÓMEZ-PÉREZ, A., MARTINS, J.P., “Some Issues on Ontology Integration”, *In: Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5)*, 16th International Joint Conference on AI, Estocolmo, Suécia, Agosto, 1999.
- REYNAUD, C., TORT, F., “Use of Expertise Ontologies in the Knowledge Engineering Process”, *Proceedings of 8th International Conference on Tools with Artificial Intelligence*, ICTAI’96, Nov., 1996.

- ROCHA, A.R.C., MAIDANTCHIK, C., OLIVEIRA, K.M., *et al.*, *Experience in Defining, Using and Improving Software Process*. Publicações Técnicas, COPPE/UFRJ, ES-507/99, set., Rio de Janeiro, RJ, Brasil, 1999.
- ROCHA, A.R.C., AGUIAR, T.C., SOUZA, J.M., “Taba: A Heuristic Workstation for Software development”, *COMPEURO 90*, Tel Aviv, Israel, Mai., 1990.
- RONSEBERG, G.H., *The Methods of Ontology*. EJAP, Indiana University. Disponível em: <<http://ejap.louisiana.edu/EJAP/1997.spring/rosenberg976.html>>. Acesso em: 20/10/1999.
- SANTOS, E., *Os Insetos (Vida e Costumes)*, Belo Horizonte, Itatiaia Limitada, Coleção Zoologia Brasília, Vol 9, 1982.
- SANTOS, E., *Os Insetos*, Belo Horizonte, Itatiaia Limitada, Coleção Zoologia Brasília, Vol. 10, 1985.
- SILVA, C.G.M. da., *Netuno: Um Ambiente de Desenvolvimento de Software Orientado ao Domínio de Acústica Submarina*, Tese de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2000.
- SMITH, B. *Ontology: Philosophical and Computational*. draft: Jul., 2000. Disponível em: <<http://ontology.buffalo.edu/smith/articles/ontologies.htm>>. Acesso em: 10/08/2000.
- SILVEIRA NETO, S. *et al.*, *Manual de Ecologia dos Insetos*, São Paulo, Ed. Agronômica Ceres, 1976.
- STAAB, S., MAEDCHE, A., “Ontology Engineering beyond the Modeling of Concepts and Relations”, *Workshop on Applications of Ontologies and Problem-Solving Methods – ECAI-00*, Berlim, Alemanha, Ago., pp 8.1-8.6, 2000.
- STORER, T.I. *et al.*, *Zoologia Geral*, 6ed., São Paulo, Editora Nacional (Biblioteca Universitária; série 3ª; Ciências Puras), Vol. 8, 1984.

- THOMPSON, H., BEECH, D., MALONEY, M., *et al.*, *XML Schema Part 1: Structures*.
May., 2001. Disponível em: <<http://www.w3.org/TR/xmlschema-1/>>. Acesso em:
10/05/2001.
- TRAVASSOS, G. H., *O Modelo de Integração de Ferramentas da Estação TABA*, Tese de
D. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1994.
- TRAVASSOS, G. H., ROCHA, A.R.C.da, “Um Modelo para Construção e Integração de
Ferramentas”, *VIII Simpósio Brasileiro de Engenharia de Software*, Curitiba, Paraná,
Out., 1994a.
- TRAVASSOS, G. H., ROCHA, A.R.C.da, “O Modelo de Integração de Ferramentas da
Estação TABA”, *Proceedings of JAIIO*, Buenos Aires, Argentina, 1994b.
- USCHOLD, M., GRUNINGER, M., “Ontologies: principles, methods and applications”, *The
Knowledge Engineering Review*, Vol 11:2, pp 93-136, 1996.
- USCHOLD, M., CLARK, P., HEALY, *et al.*, *An Experiment in Ontology Reuse*. 1998.
Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW98/uchold/>>. Acesso em:
10/05/1999.
- VALENTE, A., *Legal Knowledge Engineering - A modelling Approach*, Amsterdam,
Holanda, IOS Press, 1995.
- VALENTE, A., RUSS, T., MacGREGOR, R., *et al.*, “Building and (Re)Using an Ontology
of Air Campaign Planning”, *IEEE Intelligent Systems (Jan/Fev)*, pp 27-36, 1999.
- WAND, Y., STOREY, V. C., WEBER, R., “An Ontological Analysis of the Relationship
Construct in Conceptual Modeling”, *ACM Transaction on Database Systems*, Vol 24,
No. 4, Dez., pp 494-528, 1999.
- WAND, Y., “Ontology as a foundation for meta-modelling and method engineering”,
Information and Software Technology, No 38, pp 281-287, 1996.

YERROS. *About Conceptual Modeling*, 2001. Disponível em:
<<http://yeroos.qant.ucl.ac.br/dummies/concmod.html>>. Acesso em: 15/08/2001.

ZLOT, F. SANTOS, G. *Definição e Instanciação de Ambientes na Estação TABA*. Projeto Final de Curso, UFRJ, Rio de Janeiro, RJ, Brasil, 1999.

Referências Bibliográficas

- BASILI, V.R., CALDIERA, G., ROMBACH, H.D., “Goal Question Metric Paradigm”, *Encyclopedia of Software Engineering*, Vol. 2, pp 528-532, 1994.
- BATINI, C., CERI, S., NAVATHE, S.B., *Conceptual Database Design: an Entity-Relationship Approach*, The Benjamin/Cummings Publishing Company Inc., edwood City, 496p, 1992.
- BENCH-CAPON, T., “Why Database AND Expert Systems ?”, *In: Proceedings of the 8th International Workshop on Database and Expert Systems Applications – DEXA’97*, pp 2-5, 1997.
- BORROR, D.J., DeLONG, D.M., *Introdução ao Estudo dos Insetos*, São Paulo, Ed. da Universidade de São Paulo, 1984.
- BOWERS, D.S., *From Data to Database*, Van Nostrand Reinhold (International), 1998.
- BRAY, T., PAOLI, J., SPERBERG, C., *Extensible Markup Language (XML) 1.0 (Second Edition)*. W3C Recommendation. Out. 2000. Disponível em: <http://www.w3.org/TR/REC-xml>. Acesso em: 10/11/2000.
- BRICKLEY, D., GUHA, R.V., *Resource Description Framework (RDF) Schema Specification*. W3C Proposed Recommendation. Mar., 2000. Disponível em: <http://www.w3.org/TR/2000/CR-rdf-schema-20000327>. Acesso em: 05/04/2000.
- CASTRO, A.M.G. de, LIMA, S.M.V, CARVALHO, J.R.P. de., *Planejamento de C&T – Sistemas de Informação Gerencial*, Brasília-DF, SPI-Embrapa, 1999.
- CHANDRASEKARAN, B., JOSEPHSON, J.R., BENJAMINS, V.R., “What are Ontologies, and Why do we Need Them ?” *IEEE Intelligent Systems*, Jan-Fev, pp 20-26, 1999.
- CHEN, P.P., “The Entity-Relationship Model - Toward a Unified View of Data”, *ACM Transactions on Database Systems*, Vol. 1, No. 1, Mar., pp 9-36, 1976.

- CHEN, P.P, THALHEIM, B., WONG, L.Y., “Future Directions of Conceptual Modeling”,
In: Conceptual Modeling: Current Issues and Future Directions (eds), Springer-verlag,
Berlin, Lecturing Notes in Computer Sciences, No. 1565, pp 294-308, 1998.
- CHEN, P.P., “From Ancient Egyptian Language to Future Conceptual Modeling”, *In:*
Conceptual Modeling: Current Issues and Future Directions (eds), Springer-Verlag,
Berlin, Lecturing Notes in Computer Sciences, No. 1565, pp 57-66, 1998.
- CORCHO, O., GÓMEZ-PÉREZ, A., “Evaluating Knowledge Representation and Reasoning
Capabilities of Ontology Specification Languages”, *In: Workshop on Applications of
Ontologies and Problem-Solving Methods – ECAI-00*, Berlim, Alemanha, Ago., 2000.
- COUGO, P.S., *Modelagem Conceitual e Projeto de Banco de Dados*, 4ed., Rio de Janeiro,
Campus, 284p. 1997.
- CRANEFIELD, S; PURVIS, M., “UML as an Ontology Modelling Language”, *In:*
Proceedings of the IJCAI-99 Workshop on Intelligent Information Integration, 16th
International Joint Conference on AI. Estocolmo, Suécia, Jul. 1999.
- DATE, C.J., *Introdução a Sistemas de Banco de Dados*. Tradução [da 7ª Edição Americana]
Vandenberg Dantas de Souza, Rio de Janeiro, Campus, 2000.
- DOWELL, M.L., STEPHENS, L.M., BONNELL, R.D., “Using a Domain Knowledge
Ontology as a Semantic Gateway among Databases”, *In: Workshop on Basic
Ontological Issues in Knowledge Sharing*, IJCAI, Montreal, Canadá, 1995.
- EMAM, K. E., DROUIN, J. AND MELO W., *SPICE – The Theory and Practice of Software
Process Improvement and Capability Determination*, IEEE Computer Society Press,
1998.
- FALBO, R., MENEZES, C., ROCHA, A.R.C. da., “Using Knowledge Servers to Promote
Knowledge Integration in Software Engineering Environments”, *In: Proceedings of the
11th Software Engineering and Knowledge Engineering Conference*, pp. 170-174,
Kaiserslautern, Alemanha, Jun., pp 16-19, 1999a.

- FALBO, R., MENEZES, C., ROCHA, A.R.C. da., “Assist-Pro: Um Assistente Inteligente para Apoiar a Definição de Processos de Software”, *Anais do XIII Simpósio Brasileiro de Engenharia de Software*, pp 147-162, Florianópolis, SC, Brasil, Out., 1999b.
- FALBO, R., *Integração de Conhecimento em um Ambiente de Desenvolvimento de Software*, Tese de D. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1998.
- FARQUHAR, A., FIKES, R., RICE, J., “The Ontolingua Server: A Tool for Collaborative Ontology Construction”, *Proceedings of KAW96*, Banff, Canada, 1996.
- FENSEL, D., HORROCKS, I., HARMELEN, F.V. *et al.*, “OIL in a Nutshell”, *Workshop on Applications of Ontologies and Problem-Solving Methods – ECAI-00*, Berlim, Alemanha, Ago., pp 4.1-4.12, 2000.
- FISCHER, G., “Domain-Oriented Design Environments, Automated Software Engineering”, *The International Journal of Automated Reasoning and Artificial Intelligence in Software Engineering*, Vol 1, N 2, Jun., pp 177-203, 1994.
- FIKES, R., FARQUHAR, A., “Distributed Repositories of Highly Expressive Reusable Ontologies”, *IEEE Intelligent Systems & their applications*, v. 14, n. 2 (Mar/Abr), pp 73-79, 1999.
- FOURO, A.M.M., WERNER, C.M.L., *Modelos de Domínio e Ontologias: similares, diferentes ou complementares*, Publicações Técnicas, COPPE/UFRJ, ES 7/99, Rio de Janeiro-RJ, Brasil, Dez., 1999.
- FOURO, A.M.M., WERNER, C.M.L., “Modelos de Domínio ou Ontologias ?”, *Revista Tecnologia da Informação*, v. 2, n. 1, pp 17-24, Dez., 2000.
- GALLO, D., NAKANO, O., SILVEIRA NETO, S. *et al.*, *Manual de Entomologia Agrícola*, 2ed., São Paulo, Ed. Agronômica Ceres, 649p. il. 24p. cor, 1988.

- GALOTTA, C., ROCHA, A. R., OLIVEIRA, K. M., *Ambientes de Desenvolvimento de Software Orientado a Domínio – Definição Inicial de uma Ontologia para Acústica Submarina*, Publicações Técnicas, COPPE/UFRJ, ES 496/99, Rio de Janeiro, RJ, Brasil, 1999.
- GENESERETH, M., FIKES, R., *Knowledge Interchange Format*. Technical Report. Computer Science Department. Stanford University. Logic-92-1. 1992.
- GÓMEZ-PÉREZ, A., BENJAMINS, V.R., “Overview of Knowledge Sharing and Reuse Componentes: Ontologies and Problem-Solving Methods”, *In: Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving methods (KRR5)*, Estocolmo, Suécia, pp 1-1 – 1-15, Ago., 1999.
- GREGERSEN, H., JENSEN, C.S., “Temporal Entity-Relationship Models - a Survey”, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11 No. 3, pp 464-497, Maio/Junho, 1999.
- GUARINO, N., “Understanding, Building, and Using Ontologies”, *International Journal of Human-Computer Studies - special issue on Using Explicit Ontologies in KBS Development*, Vol. 46, No. 2/3, pp 293-310, 1997.
- GUARINO, N., “Formal Ontology and Information System”, *In: Guarino, N. (ed) Formal Ontology in Information System*, pp 3-15, IOS Press, 1998.
- GUIZZARDI, G., *Uma Abordagem Metodológica de Desenvolvimento Para e Com Reuso, Baseada em Ontologias Formais de Domínio*. Tese de M. Sc., UFES, Vitória, ES, Brasil, 2000.
- HEUSER, C. A., *Projeto de Banco de Dados*, 4ed., Porto Alegre, Ed. Sagra Luzzato, 204p, 2001.
- ISO/IEC 9126/NBR 13596, *Tecnologia de Informação – Avaliação de Produto de Software – Características de Qualidade e Diretrizes para seu uso*, ABNT – Associação Brasileira de Normas Técnicas, 1996.

- KARP, R., CHAUDHRI, V., THOMERE, J., *XOL: An XML-Based Ontology Exchange Language*. Aug., 1999. Disponível em: <<http://www.ai.sri.com/pkarp/xol/>>. Acesso em: 20/10/1999.
- KIFER, M., LAUSEN, G., WU, J., “Logical Foundations of Object-Oriented and Frame-Based Languages”, *Journal of the ACM*, 1995.
- KOLP, M., ZIMÁNYI, E., “Enhanced ER to relational mapping and interrelational normalization”, *Information and Software Technology*, Vol. 42, pp 1057-1073, 2000.
- KNUBLAUCH, H., ROSE, T., “Round-Trip Engineering of Ontologies for Knowledge-Based Systems”, In: *Proceedings of The 12th International Conference on Software Engineering & Knowledge Engineering – SEKE*, pp 239-247, Jul., 2000.
- LARA, F.M., *Princípios de Entomologia*, 3ed., São Paulo, Ícone, 1992.
- LASSILA, O., SWICK, R. R. *Resource Description Framework (RDF) Model and Syntax Specification*. W3C Proposed Recommendation. Feb., 1999. Disponível em: <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>>. Acesso em: 05/03/1999.
- LÓPEZ, M.F., GÓMEZ-PÉREZ, A., SIERRA, J.P., *et al.*, “Building a Chemical Ontology Using Methontology and the Ontology Design Environment”, *IEEE Intelligent Systems*, pp 37-46, Jan/Fev, 1999.
- LUKE, S., HEFLIN, J., *SHOE 1.01. Proposed Specification*. SHOE Project. Apr. 2000. Disponível em: <<http://www.cs.umd.edu/projects/plus/SHOE/SPEC.HTML>>. Acesso em: 10/05/2000.
- MACHADO, L.F.C., *Modelo para Definição, Especialização e Instanciação de Processos de Software*, Tese de M. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2000.
- MACHADO, F.N.R., ABREU, M.P., *Projeto de Banco de Dados - Uma Visão Prática*, 5ed., São Paulo, Érica, 297p. 1996.

- MAIDANTCHIK, C., *Modelo de Processo de Gerência de Software para Equipes Geograficamente Dispersas*, Tese de D. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1999.
- MCLELLAN, T., *Data Modelling: Finding the Perfect Fit*, 1995. Disponível em: <http://www.islandnet.com/~tmc/html/articles/datamodl.htm>. Acesso em: 20/10/1999.
- MELLO, R.S., HEUSER, C.A., “Aplicação de Ontologias a Dados Semi-Estruturados”, In: *CLEI 2000*. México, Actas, ITESM, 2000. Disponível em: <http://www.inf.ufrgs.br/~ronaldo/a000139.pdf>. Acesso em: 20/11/2000.
- MILTON, S., KAZMIERCZAK, E., “Enriching the Ontological Foundations of Modelling in Information Systems”, In: *Proceedings of the Information Systems Foundations Workshop Ontology, Semiotics and Practice*, 1999.
- MOTTA, E., *Reusable Components for Knowledge Modelling*. IOS Press. Amsterdam, 1999.
- MOURA, L. M. V., ROCHA, A. R. C. da, *Ambientes de Desenvolvimento de Software*, Publicações Técnicas COPPE/UFRJ, ES-271/92, Rio de Janeiro, Brasil, 1992.
- NBR ISO/IEC 12207, *Tecnologia de Informação – Processos de Ciclo de Vida de Software*, Associação Brasileira de Normas Técnicas, Rio De Janeiro, Brasil, 1998.
- OLIVEIRA, K.M., SANTOS, G., ZLOT, F., *et al.*, “Construção de Ambientes de Desenvolvimento de Software Orientados a Domínio na Estação TABA”, *Anais do Terceiro Workshop Ibero-americano de Engenharia de Requisitos e Ambientes de Software – IDEAS’00*, Cancun, México, Abr., 2000a.
- OLIVEIRA, K.M., MENEZES, C., ROCHA, A.R.C. da., *et al.*, “Ambiente de Desenvolvimento de Software Orientado a Domínio”, *Anais do XIV Simpósio Brasileiro de Engenharia de Software*, pp 275-290, João Pessoa, PB, Out., 2000b.
- OLIVEIRA, K.M., *Modelo para Construção de Ambientes de Desenvolvimento de Software Orientados a Domínio*, Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1999.

- OLIVEIRA, K.M., ROCHA, A.R.C., TRAVASSOS, G.H. *et al.*, “Using Domain-Knowledge in Software Development Environments”, *In: Proceedings of the 11th International Conference on Software Engineering and Knowledge Engineering*, pp. 180-187, Kaiserlauter, Alemanha, Jun. 1999a.
- OLIVEIRA, K. M., ROCHA, A. R., TRAVASSOS, G. H. *et al.*, “O uso da Teoria do Domínio no Processo de Desenvolvimento de Software”, *In: Anais da X Conferencia Internacional de Tecnologia de Software*, pp 223-235, Curitiba, Brasil, Mai., 1999b.
- PAPAVERO, N., *Fundamentos práticos de taxonomia zoológica: coleções, bibliografia, nomenclatura*. Museu paraense Emílio Goeldi e Sociedade Brasileira de Zoologia. Belém-PA, 1983.
- PARSONS, J., WAND, Y., “Using Objects for Systems Analysis”, *Communications of the ACM*, Vol 40, No. 12, pp. 104-110. Dec, 1997.
- PAULK, M. C., WEBER, C. V., CURTIS, B., *et al.* (eds), *The Capability Maturity Model: Guidelines for Improving the Software Process*, Carnegie Mellon University, Software Engineering Institute, Addison-Wesley Longman Inc. 1997.
- PINHO, S.F.C., *Avaliação da Qualidade de Dados pela não Conformidade*, Tese de Mestrado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, Abr., 2001.
- PINTO, H.S., GÓMEZ-PÉREZ, A., MARTINS, J.P., “Some Issues on Ontology Integration”, *In: Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5)*, 16th International Joint Conference on AI, Estocolmo, Suécia, Agosto, 1999.
- REYNAUD, C., TORT, F., “Use of Expertise Ontologies in the Knowledge Engineering Process”, *Proceedings of 8th International Conference on Tools with Artificial Intelligence*, ICTAI'96, Nov., 1996.

- ROCHA, A.R.C., MAIDANTCHIK, C., OLIVEIRA, K.M., *et al.*, *Experience in Defining, Using and Improving Software Process*. Publicações Técnicas, COPPE/UFRJ, ES-507/99, set., Rio de Janeiro, RJ, Brasil, 1999.
- ROCHA, A.R.C., AGUIAR, T.C., SOUZA, J.M., “Taba: A Heuristic Workstation for Software development”, *COMPEURO 90*, Tel Aviv, Israel, Mai., 1990.
- RONSEBERG, G.H., *The Methods of Ontology*. EJAP, Indiana University. Disponível em: <<http://ejap.louisiana.edu/EJAP/1997.spring/rosenberg976.html>>. Acesso em: 20/10/1999.
- SANTOS, E., *Os Insetos (Vida e Costumes)*, Belo Horizonte, Itatiaia Limitada, Coleção Zoologia Brasília, Vol 9, 1982.
- SANTOS, E., *Os Insetos*, Belo Horizonte, Itatiaia Limitada, Coleção Zoologia Brasília, Vol. 10, 1985.
- SILVA, C.G.M. da., *Netuno: Um Ambiente de Desenvolvimento de Software Orientado ao Domínio de Acústica Submarina*, Tese de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2000.
- SMITH, B. *Ontology: Philosophical and Computational*. draft: Jul., 2000. Disponível em: <<http://ontology.buffalo.edu/smith/articles/ontologies.htm>>. Acesso em: 10/08/2000.
- SILVEIRA NETO, S. *et al.*, *Manual de Ecologia dos Insetos*, São Paulo, Ed. Agronômica Ceres, 1976.
- STAAB, S., MAEDCHE, A., “Ontology Engineering beyond the Modeling of Concepts and Relations”, *Workshop on Applications of Ontologies and Problem-Solving Methods – ECAI-00*, Berlim, Alemanha, Ago., pp 8.1-8.6, 2000.
- STORER, T.I. *et al.*, *Zoologia Geral*, 6ed., São Paulo, Editora Nacional (Biblioteca Universitária; série 3ª; Ciências Puras), Vol. 8, 1984.

- THOMPSON, H., BEECH, D., MALONEY, M., *et al.*, *XML Schema Part 1: Structures*. May., 2001. Disponível em: <<http://www.w3.org/TR/xmlschema-1/>>. Acesso em: 10/05/2001.
- TRAVASSOS, G. H., *O Modelo de Integração de Ferramentas da Estação TABA*, Tese de D. Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1994.
- TRAVASSOS, G. H., ROCHA, A.R.C.da, “Um Modelo para Construção e Integração de Ferramentas”, *VIII Simpósio Brasileiro de Engenharia de Software*, Curitiba, Paraná, Out., 1994a.
- TRAVASSOS, G. H., ROCHA, A.R.C.da, “O Modelo de Integração de Ferramentas da Estação TABA”, *Proceedings of JAIIO*, Buenos Aires, Argentina, 1994b.
- USCHOLD, M., GRUNINGER, M., “Ontologies: principles, methods and applications”, *The Knowledge Engineering Review*, Vol 11:2, pp 93-136, 1996.
- USCHOLD, M., CLARK, P., HEALY, *et al.*, *An Experiment in Ontology Reuse*. 1998. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW98/uchold/>>. Acesso em: 10/05/1999.
- VALENTE, A., *Legal Knowledge Engineering - A modelling Approach*, Amsterdam, Holanda, IOS Press, 1995.
- VALENTE, A., RUSS, T., MacGREGOR, R., *et al.*, “Building and (Re)Using an Ontology of Air Campaign Planning”, *IEEE Intelligent Systems (Jan/Fev)*, pp 27-36, 1999.
- WAND, Y., STOREY, V. C., WEBER, R., “An Ontological Analysis of the Relationship Construct in Conceptual Modeling”, *ACM Transaction on Database Systems*, Vol 24, No. 4, Dez., pp 494-528, 1999.
- WAND, Y., “Ontology as a foundation for meta-modelling and method engineering”, *Information and Software Technology*, No 38, pp 281-287, 1996.

YERROS. *About Conceptual Modeling*, 2001. Disponível em:
<<http://yeroos.qant.ucl.ac.br/dummies/concmod.html>>. Acesso em: 15/08/2001.

ZLOT, F. SANTOS, G. *Definição e Instanciação de Ambientes na Estação TABA*. Projeto Final de Curso, UFRJ, Rio de Janeiro, RJ, Brasil, 1999.

Anexo 1

Descrição de Conceitos da Teoria do Domínio de Entomologia

Tabela A1.1 – Conceitos da Subteoria Anatomia de inseto

Conceito	Sinônimo	Descrição	SubConceito	Instâncias	Propriedades de Instância
INSETO		Animal da classe insecta, pertencente ao filo Arthropoda.		Abelha, Besouro, Borboleta, Cigarra, Cupim, Grilo, Joanhina, Libélula, Mosca, Pernilongo, Pulga etc.	Nome
MORFOLOGIA		Forma e estrutura do inseto	Morfologia Interna, Morfologia Externa	Morfologia de Abelhas, Morfologia de Besouros etc.	
MORFOLOGIA INTERNA	Anatomia Interna	São todas as partes que compõem internamente o corpo dos insetos.		Anatomia interna de grilos, Anatomia de cupins etc.	
ÓRGÃO	Órgão interno	Cada uma das partes componentes de um sistema interno, que tem uma função específica ou um conjunto de funções.		Edeago, ovário, testículo, estatocisto, tímpano, intestino, cérebro, coração	Nome, Função
SISTEMA	Aparelho	Grupo de órgãos internos, que em conjunto exercem uma função especial num corpo.		Glandular, Reprodutor, Nervoso, Excretor, Circulatório, Digestivo, Muscular, Trocas Gasosas	Nome, Função

Conceito	Sinônimo	Descrição	SubConceito	Instâncias	Propriedades de Instância
ENDOESQUELETO		Estrutura de sustentação no interior do corpo.			
MORFOLOGIA EXTERNA	Anatomia Externa	São todas as partes que compõem externamente o corpo dos insetos.			
CORPO		Conjunto de todas as regiões que compõem a estrutura física do inseto.			Forma, Tamanho
EXOESQUELETO	Tegumento, Carapaça, Couraça	É a estrutura de sustentação externa (esqueleto), que reveste o corpo do inseto.			Cor
CAMADA		Cada um dos revestimentos que compõem o exoesqueleto.		Cutícula, Epiderme, Membrana Basal	Nome
SUPERFÍCIE		É a cobertura ou parede do exoesqueleto.			Cor
SUTURA		Um sulco externo, semelhante a uma linha na superfície, ou uma área membranosa estreita entre os escleritos.		Apódema, Carena, Claval, Coronal, Esternocostal	Nome, Localização
ESCLERITO		Placa endurecida (esclerosada) da superfície, marginada por suturas.		Tergo, Noto, Esterno, Pleurito, Escutelo, Gula, Fronte, Clípeo	Nome, Localização, Forma
REGIÃO	Tagma	Cada uma das partes ou seções em que se pode considerar dividido o corpo dos insetos.	Cabeça, Tórax, Abdome		Nome, Tamanho
CABEÇA		É a região anterior do corpo dos insetos consistindo de uma cápsula quitinizada.			Forma
TÓRAX		Região média do corpo de um inseto, que se caracteriza por apresentar os apêndices locomotores.			Forma
ABDOME		É a terceira região do corpo do inseto, situada atrás do tórax.	Abdome Sésil, Abdome Livre, Abdome Pedunculado		Cor, Quantidade de segmentos

Conceito	Sinônimo	Descrição	SubConceito	Instâncias	Propriedades de Instância
ABDOME SÉSSIL		É o abdome que se liga ao tórax em toda a extensão de seu primeiro segmento.			
ABDOME LIVRE		É o abdome cuja área transversal não se liga totalmente ao tórax.			
ABDOME PEDUNCULADO		É o abdome que possui o segundo segmento apresentando uma forte e longa constrição.			
APÊNDICE		É uma peça ou parte suplementar de qualquer região do corpo do inseto com função específica.	Apêndice da Cabeça, Apêndice Torácico, Apêndice Abdominal	Procto, Cerco	Nome, Localização, Função, Quantidade, Cor, Tamanho, Forma
APÊNDICE DA CABEÇA		É toda e qualquer peça ou parte suplementar localizada na cabeça do inseto, com função específica.	Fixo, Móvel		
APÊNDICE FIXO		É toda e qualquer peça ou parte suplementar fixa à cabeça do inseto.		Ocelo, Olho	
APÊNDICE MÓVEL		É toda e qualquer peça ou parte suplementar localizada na cabeça do inseto, que pode mover-se.		Aparelho Bucal, Antena	Tipo
APÊNDICE TORÁCICO		É cada um dos apêndices localizados no tórax.	Locomotor		
APÊNDICE LOCOMOTOR		Apêndice torácico que permite a locomoção terrestre, aquática ou aérea.		Perna, Asa	Tipo
APÊNDICE ABDOMINAL		É cada um dos apêndices localizados no abdome.		Cerco, Urogonfo, Filamento Mediano, Gnathos	
SEGMENTO		Cada uma das subdivisões das regiões.	Segmento da Cabeça, Segmento Torácico, Segmento Abdominal		Nome, Localização

Conceito	Sinônimo	Descrição	SubConceito	Instâncias	Propriedades de Instância
SEGMENTO DA CABEÇA		É cada uma das subdivisões da cabeça.	Área Intersutural		
ÁREA INTERSUTURAL	Área da cabeça	É cada um dos segmentos da cabeça delimitados por suturas.			Forma, Tamanho, Tipo
SEGMENTO TORÁXICO		É cada um dos segmentos que compõe o tórax.			Forma, Cor
PROTÓRAX		É o segmento torácico que está unido à cabeça.			
MESOTÓRAX		É o segmento torácico localizado entre o protórax e o metatórax.			
METATÓRAX		É o terceiro segmento torácico que se liga ao abdome.			
SEGMENTO ABDOMINAL	Urômero, Urito	Cada uma das partes, bem definidas, em que se divide o abdome dos insetos.			
SEGMENTO PRÉ-GENITAL	Visceral	É o conjunto formado pelos urômeros de 1-7 nas fêmeas e 1-8 nos machos.		Propódeo	
SEGMENTO GENITAL	Genitália	É um dos segmentos abdominal terminal modificado para copulação ou postura de ovos. São os órgãos sexuais externos, correspondendo ao nono segmento nos machos e oitavo segmento nas fêmeas.			
SEGMENTO PÓS-GENITAL		Compreende, principalmente, o décimo e o décimo primeiro urômeros.			

Tabela A1.2 - Propriedades de Instâncias dos Conceitos da Anatomia de inseto

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Nome	INSETO	Identificação pela qual os insetos são conhecidos vulgarmente.	String		
Nome	ÓRGÃO	Identificação do órgão.	String		
Função	ÓRGÃO	Função desempenhada pelo órgão dentro do sistema que faz parte.	Texto		
Nome	SISTEMA	Designação do sistema.	String		
Função	SISTEMA	Função desempenhada pelo sistema na estrutura (corpo) do inseto.	Texto		
Forma	CORPO	Formato da estrutura exterior do corpo dos insetos.	Texto		
Tamanho	CORPO	Medida do comprimento do corpo dos insetos.	Numérico		cm ou mm
Cor	EXOESQUELETO	Coloração do exoesqueleto dos insetos.	String		
Nome	CAMADA	Identificação de cada um dos revestimentos que compõem o exoesqueleto.	String	Cutícula, Epiderme, Membrana Basal	
Cor	SUPERFÍCIE	Coloração da superfície que reveste o exoesqueleto.	String		
Nome	SUTURA	Designação de cada sutura que compõe a superfície.	String		
Localização	SUTURA	Lugar onde a sutura se encontra na superfície.	Texto		
Nome	ESCLERITO	Designação de cada esclerito que compõe a superfície.	String		
Localização	ESCLERITO	Lugar onde o esclerito se encontra na superfície.	Texto		
Forma	ESCLERITO	Formato da estrutura do esclerito.	Texto		
Nome	REGIÃO	Identificação de cada uma das regiões do corpo.	String	Cabeça, Tórax, Abdome	
Tamanho	REGIÃO	Medida do comprimento de cada uma das regiões do corpo.	Numérico		cm ou mm
Forma	CABEÇA	Formato da estrutura da cabeça.	Texto		
Forma	TÓRAX	Formato da estrutura do tórax.	Texto		
Cor	ABDOME	Coloração do abdome dos insetos.	String		

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Quantidade de segmentos	ABDOME	Número total de segmentos que constituem o abdome.	Numérico	6 a 11	Unidade
Nome	APÊNDICE	Identificação de cada um dos apêndices existentes do corpo do inseto.	String		
Localização	APÊNDICE	Lugar do corpo onde se encontra o apêndice.	String		
Função	APÊNDICE	Função desempenhada pelo apêndice no corpo do inseto.	Texto		
Quantidade	APÊNDICE	Número total de um determinado apêndice no corpo.	Numérico		Unidade
Cor	APÊNDICE	Coloração do apêndice.	String		
Tamanho	APÊNDICE	Medida do comprimento do apêndice.	Numérico		cm ou mm
Forma	APÊNDICE	Formato do apêndice.	Texto		
Tipo	APÊNDICE MÓVEL	Classificação dos apêndices móveis de acordo com sua forma, localização ou função.	String		
Tipo	APÊNDICE LOCOMOTOR	Classificação dos apêndices locomotores de acordo com sua forma, função ou constituição.	String		
Nome	SEGMENTO	Identificação de cada um dos segmentos existentes no corpo do inseto.	String		
Localização	SEGMENTO	Lugar do corpo, onde se encontra o segmento.	String		
Forma	ÁREA INTERSUTURAL	Formato da estrutura da área intersutural.	Texto		
Tamanho	ÁREA INTERSUTURAL	Medida do tamanho da área intersutural.	Numérico		cm ou mm
Tipo	ÁREA INTERSUTURAL	Classificação da área intersutural de acordo com sua localização.	String		
Forma	SEGMENTO TORÁXICO	Formato da estrutura do segmento torácico.	Texto		
Cor	SEGMENTO TORÁXICO	Pigmentação do segmento torácico.	String		

Tabela A1.3 – Conceitos Subteoria Biologia de inseto

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
CICLO DE VIDA		Conjunto de acontecimentos semelhantes que todos os indivíduos de uma espécie experimentam ao longo da sua vida, desde o nascimento até à morte.			Tipo
REPRODUÇÃO		Função biológica pela qual os insetos reproduzem a sua espécie, quase sempre em indivíduos distintos (macho e fêmea), que se unem para dar início a novo ser.			Tipo, Quantidade de ovos
DESENVOLVIMENTO		É o progresso de um estágio a outro, havendo mudanças de forma e hábito, de tal modo que o seguinte é sempre mais perfeito que o anterior.			Tipo de Metamorfose
ESTÁGIO	Estádio, Fase	É cada estágio particular que ocorre no desenvolvimento, entre mudas sucessivas.	Ovo, Ninfa, Larva, Pupa, Adulto		Número médio de dias
OVO		Célula que resulta da fecundação dos gametas. Primeiro estágio do desenvolvimento dos insetos.		Ovo do gafanhoto	Forma, Cor, Tamanho
NINFA		É o estágio imaturo do desenvolvimento dos insetos que não tem estágio pupal, caracterizado pela imobilidade completa.			Tamanho
LARVA	Lagarta	Primeiro estágio do desenvolvimento dos insetos, no qual sofrem metamorfose completa, depois de saírem do ovo, caracterizado por crescimento intenso; é o embrião que se torna livre, abandonando normalmente os ovos ou o organismo progenitor.			Tamanho, Tipo
PUPA		Estágio intermediário entre a larva e o adulto, em insetos com metamorfose completa, um estágio sem alimentação e usualmente inativo. É a segunda fase pós-embrionária e que se caracteriza por aparente dormência.			Tipo
ADULTO		Última fase do ciclo de vida dos insetos, significando que atingiu pleno desenvolvimento.			

Tabela A1.4 - Propriedades de Instâncias dos Conceitos da Biologia de inseto

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Tipo	CICLO DE VIDA	Classificação do ciclo de vida sob o aspecto da existência ou não de um período de inatividade.	String	Homodinâmico, Heterodinâmico	
Tipo	REPRODUÇÃO	Classificação da reprodução sob o aspecto da forma como a reprodução ocorre.	String	Fecundação, Partenogênese, Pedogênese, Poliembrião	
Quantidade de ovos	REPRODUÇÃO	Número total de ovos que o inseto põe a cada reprodução.	Numérico		Unidade
Tipo de Metamorfose	DESENVOLVIMENTO	Classificação da metamorfose ocorrida no desenvolvimento do inseto.	String	Ametabolia, Hemimetabolia, Hipermetabolia, Hipometabolia, Holometabolia, Paurometabolia	
Número médio de dias	ESTÁGIO	Número médio de dias que o inseto permanece em determinado estágio.	Numérico		Unidade
Forma	OVO	Forma geométrica do ovo.	String		
Cor	OVO	Pigmentação do ovo.	String		
Tamanho	OVO	Medida do comprimento do ovo.	Numérico		cm ou mm
Tamanho	NINFA	Medida do comprimento do inseto no estágio de ninfa.	Numérico		cm ou mm
Tamanho	LARVA	Medida do comprimento do inseto no estágio de larva.	Numérico		cm ou mm
Tipo	LARVA	Classificação da larva sob o aspecto de sua forma.	String	Campodeiforme, Carabiforme, Curculioniforme, Elateriforme, Eruciforme, Escarabeiforme, Limaciforme, Vermiforme	
Tipo	PUPA	Classificação da pupa de acordo com seu recobrimento por casulos.	String	Exarada, Obtecta, Coarctata	

Tabela A1.5 – Conceitos da Subteoria Ecologia de inseto

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
ECOSSISTEMA	Bioma	É a unidade funcional da ecologia, sendo constituído pelo conjunto de comunidades que ocupam um meio ambiente, contendo recursos suficientes para poder assegurar a conservação da vida e ocorrem numa localidade.		Estuário, Rio, Lago, Brejo, Tundra, Montanha, Floresta, Campo	Nome, Tipo
MEIO AMBIENTE		É o conjunto de fatores ecológicos (condições de temperatura e umidade) de um local, que agem de forma permanente sobre os seres vivos, influenciando os nichos ecológicos dos organismos em geral.			
COMUNIDADE		É o agrupamento natural de populações de diversos insetos, que ocorre em um meio ambiente, com capacidade de sobrevivência e sustentação própria.		Comunidade de Abelhas, comunidade de Cupins, etc.	Nome
NICHO ECOLÓGICO		É o papel que um inseto executa e que é realizado dentro da comunidade, implicando em todos os componentes químicos, físicos e biológicos, o período do dia em que se alimenta, a temperatura, a umidade, as partes da localidade que usa, a forma como se reproduz e a forma como se comporta.			
POPULAÇÃO		É um grupo de insetos da mesma espécie, ocupando uma localidade, capazes de se reproduzir entre si.			Densidade, Nível de equilíbrio, nível de dano, limiar econômico
LOCALIDADE		É o lugar determinado que a população ocupa e que determina a sua distribuição.			Nome
DISTRIBUIÇÃO		É a disposição das diversas populações de insetos, definida pelas diversas localidades ocupadas pelas mesmas.			

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
COMPORTAMENTO		É o conjunto de reações, observáveis objetivamente, de um inseto que age em resposta a um estímulo ou através de seus instintos. É um dos componentes do nicho ecológico executado pelo inseto.			
RESPOSTA		É o resultado de impulsos, que são despertados pelos estímulos ou acionados pelo instinto e que se deslocam pelo sistema nervoso.			Tipo
ESTÍMULO		É um elemento do meio ambiente ou do meio interno suscetível de provocar, em certas condições, uma resposta de um sistema ou ser vivo. É um fenômeno biofísico capaz de desencadear a experiência de uma sensação pelo ser vivo.			Tipo
INSTINTO		Mecanismo pelo qual os insetos podem desenvolver padrões complexos de comportamento sem aprendizagem ou esforço consciente.			
FATOR ECOLÓGICO	Fator do meio ambiente	É qualquer elemento do meio ambiente capaz de influenciar diretamente os seres vivos, através de seus nichos ecológicos e tendo ascendência uns sobre os outros.	Alimento, Tempo		Nome
ALIMENTO		É o componente do meio ambiente que nutre e sustenta os seres vivos.			Tipo
TEMPO		É o conjunto de fatores físicos de uma localidade, num período menor que um ano, ou seja, é o estado da atmosfera num momento e lugar.			

Tabela A1.6 - Propriedades de Instâncias dos Conceitos da Ecologia de Inseto

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Nome	ECOSSISTEMA	Identificação do ecossistema onde o inseto está inserido.	String		
Tipo	ECOSSISTEMA	É a classificação geral dos ecossistemas.	String	Aquático, Terrestre	
Nome	COMUNIDADE	Identificação da comunidade onde o inseto executa seu nicho ecológico.	String		
Densidade	POPULAÇÃO	É o número de indivíduos de uma população por unidade de área.	Numérico		
Nível de equilíbrio	POPULAÇÃO	É a densidade média da população durante um longo período de tempo, onde não ocorreram mudanças no meio ambiente.	Numérico		
Nível de dano	POPULAÇÃO	É a menor densidade populacional que causa prejuízo econômico, numa localidade.	Numérico		
Limiar econômico	POPULAÇÃO	É a densidade populacional na qual, medidas devem ser tomadas para evitar que a população atinja o nível de dano.	Numérico		
Nome	LOCALIDADE	Identificação do local onde a população de inseto ocorre.	String		
Tipo	RESPOSTA	É a classificação da resposta dada pelos insetos em função da ocorrência de estímulos ou de instintos.	String	Reflexo, Taxia, Chinese	
Tipo	ESTÍMULO	É a classificação do estímulo que pode provocar uma resposta por parte dos insetos.	String		
Nome	FATOR ECOLÓGICO	É a identificação do fator ecológico que influencia o nicho ecológico dos insetos.	String	Tempo, Alimento	
Tipo	ALIMENTO	É a classificação dos alimentos consumidos pelos insetos.	String		

Tabela A1.7 - Subteoria Classificação de Inseto

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
CLASSIFICAÇÃO	Taxonomia, Sistemática	Agrupamento dos insetos com base, principalmente, nas características morfológicas que tenham em comum.			
CATEGORIA TAXONÔMICA	Categoria Sistemática	É cada um dos grupos em que são classificados os insetos.	Reino, Filo, Classe, Ordem, Família, Gênero, Espécie		Nome Científico
REINO		É a maior das categorias de classificação biológica que divide, atualmente, todos os organismos vivos e fósseis em cinco reinos.		Monera (bactérias), Protista, Fungo, Vegetal, Animal	
FILO	Ramo	Categoria de classificação biológica mais abrangente dentro de um reino. O filo é formado por classes relacionadas.		Protozoa, Porífera, Coelenterata, Arthropoda, Chordata	
SUBFILO		Maior subdivisão de um filo, contendo um grupo de classes aparentadas.		Hexapoda	
CLASSE		Categoria de classificação biológica entre a ordem e o filo. A classe é formada por ordens relacionadas.		Insecta, Arachnida, Crustacea, Xiphosura	
SUBCLASSE		Maior subdivisão de uma classe, contendo um grupo de ordens aparentadas.		Apterygota, Pterygota, Exopterygota	
ORDEM		Categoria de classificação biológica entre a família e a classe. Uma ordem é formada por um grupo de famílias relacionadas.		Hymenoptera, Thysanura, Odonata, Orthoptera, Isoptera	

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
SUBORDEM		Maior subdivisão de uma ordem, contendo um grupo de famílias aparentadas.		Apocrita, Acridoidea, Tettigonoidea	
FAMÍLIA		Categoria de classificação biológica entre o gênero e a ordem. É a subdivisão de uma ordem, subordem ou superfamília. A família é formada por um grupo de gêneros, tribos ou subfamílias relacionadas.		Formicidae, Apidae, Chrysomelidae, Vespidae	
SUPERFAMÍLIA		É um grupo de famílias de parentesco próximo.		Formicoidea, Aeshnoidea, Tettigonoidea, Grylloidea	
SUBFAMÍLIA		Maior subdivisão de uma família, contendo um grupo de tribos ou gêneros aparentados.		Ponerinae, Acridinae, Triatominae, Eumeninae	
TRIBO		Subdivisão de uma subfamília, contendo um grupo de gêneros aparentados.		Ectatommini	
GÊNERO		É a categoria de classificação biológica entre a espécie e a família. O gênero é composto por espécies estruturalmente relacionadas.		Ectatomma, Anacroneuria, Monalonion, Phassus	
SUBGÊNERO		Maior subdivisão de um gênero, contendo um grupo de espécies aparentadas.		Ectatomma, Epicharis (Hoplepicharis)	
ESPÉCIE		É a categoria abaixo do gênero, é a unidade fundamental da classificação biológica. É formada por organismos que podem se acasalar e gerar descendentes plenamente funcionais.		Edentatum, Papilio glaucus Linneu, Leptinotarsa decimlineata (Say)	Nome popular
SUBESPÉCIE	Raça geográfica, Variedade	Subdivisão de uma espécie, usualmente uma raça geográfica de uma espécie.		Muticum, Melipona rufiventris flavolineata Friese	

Tabela A1.8 - Propriedades de Instâncias dos Conceitos da Classificação de Inseto

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Nome Científico	CATEGORIA TAXONÔMICA	Identificação de toda e qualquer categoria taxonômica	String		
Nome Popular	ESPÉCIE	Identificação regional de uma espécie.	String		

Tabela A1.9 - Subteoria Processo de Trabalho do Entomólogo

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
ATIVIDADE	Tarefa	É toda ação executada pelos recursos humanos, transformando insumos em produtos, com a utilização dos recursos necessários à sua execução.	Pré-Identificação, Identificação, Manejo, Controle		Nome
PRÉ-IDENTIFICAÇÃO		É toda tarefa que precisa ser executada antes da atividade de identificação.	Coleta, Tratamento, Armazenamento		
COLETA		É a tarefa realizada pelo coletor, utilizando material e procedimento específico de acordo com a espécie a ser coletada.		Coleta de insetos voadores	
TRATAMENTO		É tarefa, composta de várias atividades, tais como montagem, rotulagem e etiquetagem, que objetivam tratar adequadamente o material coletado.		Tratamento de insetos voadores	
ARMAZENAMENTO		É a tarefa que visa armazenar adequadamente o material tratado.			

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
IDENTIFICAÇÃO	Classificação	Processo de determinar uma forma biológica por descrição, original ou secundária, ou por comparação com o tipo ou outro exemplar previamente classificado, determinando a denominação dos espécimes (nome científico), para que possam ser introduzidos ordenadamente nas coleções.			Espécie, Autor, Ano da Identificação, Determinador
MANEJO		É a tarefa que busca lidar racionalmente com a criação de insetos benéficos e de interesse econômico.			
CONTROLE		É o processo de supervisão utilizado para manter as pragas abaixo do nível de dano econômico.	Controle Mecânico, Controle Químico, Controle Integrado, Controle Biológico, Controle Cultural, Controle Comportamental, Controle Genético		Nome
CONTROLE MECÂNICO		É um dos métodos de controle de pragas, que consiste do emprego de medidas mecânicas de controle.		Catação manual de pragas, Formação de barreiras ou sulcos, Uso de armadilha	
CONTROLE QUÍMICO		É um dos métodos de controle de pragas, que consiste no uso de compostos químicos através de uma formulação.	Inseticida, Formulação		

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
INSETICIDA		É todo e qualquer composto químico que aplicado direta ou indiretamente sobre os insetos, em concentrações adequadas, provocam a sua morte.			Nome, Composição, Toxicidade
FORMULAÇÃO		É a maneira de transformar um produto técnico numa forma apropriada de uso. Ou seja, a formulação permite que o produto seja utilizado de modo conveniente.			
CONTROLE INTEGRADO		É o método de controle de pragas, que consiste da utilização de meios (técnicas), representados pelos diferentes métodos de controle e que podem, inclusive, serem integrados com inseticidas, desde que esta integração seja feita de forma harmoniosa.			
CONTROLE BIOLÓGICO		É a utilização de organismos vivos para controlar pragas agrícolas ou de outro tipo, com a utilização de outros organismos vivos (predadores, parasitas e competidores) ou de produtos biológicos, tais como hormônios.	Inimigo Natural, Patógeno		Nome do agente
INIMIGO NATURAL		É o controle que utiliza agentes de mortalidade biótica. Estes agentes são os insetos entomófagos (predador e parasito).			
PATÓGENO		É o controle que utiliza agentes que causam doenças nos insetos. São os responsáveis pelo controle microbiano, visando a manutenção da população das pragas a níveis não prejudiciais.			

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
CONTROLE CULTURAL		É um dos métodos de controle de pragas e consiste em se empregar certas práticas culturais para o controle, baseando-se em conhecimentos ecológicos e biológicos das pragas.		Rotação de cultura, Aração do solo, Poda, Adubação, Irrigação	
CONTROLE COMPORTAMENTAL		É um dos métodos de controle de pragas e se baseia nos estudos da fisiologia dos insetos.		Controle com hormônios, Uso de atraentes e repelentes, Esterilização de insetos	
CONTROLE GENÉTICO		É um dos métodos de controle de pragas e consiste na modificação genética dos insetos.			
PRODUTO		É tudo que é gerado em consequência da realização das atividades.	Chave de Classificação, Coleção, Trabalho Científico, Receituário de Controle		Nome
CHAVE DE CLASSIFICAÇÃO	Chave analítica	É um dos produtos resultantes da atividade de identificação, servindo como meio para identificar todos os tipos de organismos, inclusive os insetos.		Chave para os gêneros de Eburini	
COLEÇÃO	Coleção taxonômica	É o agrupamento ordenado de espécimes mortos ou partes corporais dos mesmos, devidamente montados e preservados para estudo, visando atender a diversas finalidades, ou seja, desde o interesse amador de colecionadores até o interesse puramente científico de profissionais dedicados à sistemática entomológica.			Localização

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
ESPÉCIME		Amostra de inseto devidamente identificada, montada e conservada como parte de uma coleção.			
TRABALHO CIENTÍFICO		É qualquer trabalho publicado descrevendo espécies novas, chaves analíticas ou revisões de qualquer categoria taxonômica.			Título
RECEITUÁRIO DE CONTROLE		É o resultado da atividade de controle que estabelece o método a ser empregado e os procedimentos a serem utilizados para controlar a praga identificada.			
RECURSO		Equivale a tudo o que é necessário para a execução de uma atividade.	Procedimento, Código de Nomenclatura, Recurso Humano		Nome
PROCEDIMENTO		É toda e qualquer conduta bem estabelecida e ordenada para a realização das atividades.			
CÓDIGO DE NOMENCLATURA		É a norma a ser seguida quando da atribuição de nomes aos espécimes, de forma que o nome seja único e distinto, promovendo a estabilidade e universalidade dos nomes científicos.			
RECURSO HUMANO		São todas as pessoas que desempenham determinado papel na execução das atividades.	Curador, Entomólogo, Sistemata, Coletor		Nome
ENTOMÓLOGO		É todo profissional especialista no estudo dos insetos sob todos os seus aspectos.			
SISTEMATA	Determinador	É todo entomólogo com especialidade em sistemática (estudo dos princípios gerais da classificação científica).			

Conceito	Sinônimo	Descrição	Sub-Conceito	Instâncias	Propriedades de Instâncias
CURADOR		É o profissional responsável pela execução de todas as atividades administrativas de uma coleção, tais como: cuidar das decisões para o bom manejo das coleções, avaliação das necessidades e condições de empréstimo do material, procedimentos e adoção de métodos de catalogação, levantamentos ou tombamento, doações e permutas.			
COLETOR		É a pessoa responsável pela execução da atividade de coleta.			

Tabela A1.10 - Propriedades de Instâncias dos conceitos do Processo de Trabalho do Entomólogo

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Nome	ATIVIDADE	Identificação das atividades executadas no processo de trabalho do entomólogo.	String		
Espécie	IDENTIFICAÇÃO	Nome científico da espécie que está sendo identificada.	String		
Autor	IDENTIFICAÇÃO	Nome do autor da identificação da primeira espécie.	String		
Ano da Identificação	IDENTIFICAÇÃO	Ano em que foi executada a identificação da espécie.	Numérico		
Determinador	IDENTIFICAÇÃO	Nome da pessoa que executou a identificação da espécie.	String		
Nome	CONTROLE	Identificação dos controles que podem ser recomendados através dos receituários de controle.	String		
Nome	INSETICIDA	Identificação do inseticida utilizado no controle químico.	String		
Composição	INSETICIDA	Especificação dos elementos químicos que compõem o inseticida.	Texto		
Toxidade	INSETICIDA	Especificação do nível de toxidade do inseticida.	Texto		
Nome agente	CONTROLE BIOLÓGICO	Identificação do agente utilizado no controle biológico, como inimigo natural ou como agente causador de doença no inseto a ser controlado.	String		

Propriedades de Instância	Conceito	Descrição	Tipo	Faixa de valores	Unidade
Nome	PRÓDUTO	Identificação do produto gerado ou utilizado na realização das atividades do processo de trabalho do entomólogo.	String		
Localização	COLEÇÃO	Local onde a coleção está instalada.	String		
Título	TRABALHO CIENTÍFICO	Identificação do trabalho científico produzido.	String		
Nome	RECURSO	Identificação do recurso utilizado na realização das atividades do processo de trabalho do entomólogo.	String		
Nome	RECURSO HUMANO	Identificação do recurso humano que atua na realização das atividades do processo de trabalho do entomólogo.	String		

Anexo 2

Avaliação da Teoria do Domínio de Entomologia

Este questionário visa avaliar a construção da teoria do Domínio de Entomologia. O objetivo é identificar se a teoria está definida de forma clara e objetiva, se está consistente, se é geral e completa dentro do seu escopo, e se é concisa.

Caracterização do Especialista:

Nome (opcional):		e-mail (opcional):		
ÁREA DE ATUAÇÃO		FORMAÇÃO		
<i>Empresa</i>		<i>Universidade</i>		<i>Nível</i>
Empresário	Professor	Doutorado	<input type="checkbox"/> Eng de Software <input type="checkbox"/> Computação/Informática <input type="checkbox"/> Outro	
Gerente de Informática	Pesquisador	Mestrado	<input type="checkbox"/> Eng de Software <input type="checkbox"/> Computação/Informática <input type="checkbox"/> Outro	
Gerente da Qualidade	Consultor	Especialização	<input type="checkbox"/> Eng de Software <input type="checkbox"/> Computação/Informática <input type="checkbox"/> Outro	
Gerente de Projeto	Aluno de Doutorado	Graduação	<input type="checkbox"/> Eng de Software <input type="checkbox"/> Computação/Informática <input type="checkbox"/> Outro	
Analista de Sistemas	Aluno de Mestrado	Muito obrigada por sua colaboração !		
Outro:	Aluno de Graduação			
Tempo de Atuação na Área: Anos				

Definições das Características Existentes na Teoria

1. **Clareza:** evidência da existência de uma comunicação efetiva do significado dos termos definidos, das descrições feitas, e se as definições e descrições são objetivas e independente de contexto social
2. **Consistência:** evidência da existência, da consistência semântica das definições e descrições; se as definições e descrições são consistentes com o mundo real e consistentes umas com as outras; se as definições e descrições não são sentenças contraditórias e que podem ser inferidas usando outras definições.
3. **Generalidade:** evidência da existência da possibilidade do compartilhamento da teoria como um todo entre diferentes atividades.
4. **Completeness:** evidência de que o domínio está precisamente e exatamente bem delimitado.
5. **Concisão:** evidência de que a teoria como um todo captura apenas a informação necessária ao seu propósito, e de que suas definições são realmente concisas.
6. **Robustez:** evidência e clareza de que pequenas mudanças não afetam o conjunto de definições já avaliadas.

INSTRUÇÕES		Grau de Existência	Valor
<p>A avaliação consiste de três partes.</p> <p>Na primeira parte, temos os itens existentes na teoria e que deverão ser avaliados quanto ao grau de existência das características de qualidade descritas a seguir.</p> <p>Determine o Grau de Existência (segundo os valores da tabela ao lado) de cada um dos itens apresentados.</p>		Nenhuma	0
		Insuficiente	1
		Razoável	2
		Boa	3
		Excelente	4

Itens para avaliação		Características Existentes na Teoria					
		Clareza	Consistência	Generalidade	Compleitude	Concisão	Robustez
1	O texto introdutório, descrevendo a motivação e a forma de construção da teoria do domínio						
2	A descrição da Identificação do Propósito da Teoria do Domínio						
3	As questões de competência gerais definidas						
4	As subteorias originárias das questões de competência gerais						
5	O texto explicativo e a Figura 1 (inter-relacionamento das subteorias)						
6	A especificação da subteoria anatomia de insetos						
7	O texto referente a conceituação da subteoria anatomia de insetos						
8	As questões de competência da subteoria anatomia de insetos						
9	A árvore de conceitos da subteoria anatomia de insetos (Figura 2)						
10	Os axiomas definidos para a subteoria anatomia de insetos						
11	A descrição dos conceitos da subteoria anatomia de insetos						
12	A descrição das propriedades de instâncias dos conceitos da subteoria anatomia de insetos						
13	A especificação da subteoria biologia de insetos						
14	O texto referente a conceituação da subteoria biologia de insetos						
15	As questões de competência da subteoria biologia de insetos						
16	A árvore de conceitos da subteoria biologia de insetos (Figura 3)						
17	Os axiomas definidos para a subteoria biologia de insetos						
18	A descrição dos conceitos da subteoria biologia de insetos						
19	A descrição das propriedades de instâncias dos conceitos da subteoria biologia de insetos						
20	A especificação da subteoria ecologia de insetos						
21	O texto referente a conceituação da subteoria ecologia de insetos						

Grau de Existência	Valor
Nenhuma	0
Insuficiente	1
Razoável	2
Boa	3
Excelente	4

Itens para avaliação		Características Existentes na Teoria					
		Clareza	Consistência	Generalidade	Completude	Concisão	Robustez
22	As questões de competência da subteoria ecologia de insetos						
23	A árvore de conceitos da subteoria ecologia de insetos (Figura 4)						
24	Os axiomas definidos para a subteoria ecologia de insetos						
25	A descrição dos conceitos da subteoria ecologia de insetos						
26	A descrição das propriedades de instâncias dos conceitos da subteoria ecologia de insetos						
27	A especificação da subteoria classificação de insetos						
28	O texto referente a conceituação da subteoria classificação de insetos						
29	As questões de competência da subteoria classificação de insetos						
30	A árvore de conceitos da subteoria classificação de insetos (Figura 5)						
31	Os axiomas definidos para a subteoria classificação de insetos						
32	A descrição dos conceitos da subteoria classificação de insetos						
33	A descrição das propriedades de instâncias dos conceitos da subteoria classificação de insetos						
34	A especificação da subteoria processo de trabalho do entomólogo						
35	O texto referente a conceituação da subteoria processo de trabalho do entomólogo						
36	As questões de competência da subteoria processo de trabalho do entomólogo						
37	A árvore de conceitos da subteoria processo de trabalho do entomólogo (Figura 6)						
38	Os axiomas definidos para a subteoria processo de trabalho do entomólogo						
39	A descrição dos conceitos da subteoria processo de trabalho do entomólogo						
40	A descrição das propriedades de instâncias dos conceitos da subteoria processo de trabalho do entomólogo						
41	As tarefas identificadas para o domínio de entomologia						
42	A relação tarefa - subteoria definida						

INSTRUÇÕES

Na segunda parte, temos a relação de itens existentes na teoria e que deverão ser avaliados quanto a sua concordância da existência da relação.
 Determine se a relação existe ou não, de forma clara, para cada um dos itens apresentados.

Relação de complementação entre os itens		Avaliação	
		SIM	NÃO
1	A árvore de conceitos da subteoria anatomia de insetos (Figura 2) está respondendo as questões de competência estabelecidas para a subteoria		
2	Os axiomas definidos para a subteoria anatomia de insetos está complementando as informações existentes na árvore de conceitos (Figura 2) em relação às questões de competência estabelecidas para a subteoria		
3	A árvore de conceitos da subteoria biologia de insetos (Figura 3) está respondendo as questões de competência estabelecidas para a subteoria		
4	Os axiomas definidos para a subteoria biologia de insetos está complementando as informações existentes na árvore de conceitos (Figura 3) em relação às questões de competência estabelecidas para a subteoria		
5	A árvore de conceitos da subteoria ecologia de insetos (Figura 4) está respondendo as questões de competência estabelecidas para a subteoria		
6	Os axiomas definidos para a subteoria ecologia de insetos está complementando as informações existentes na árvore de conceitos (Figura 4) em relação às questões de competência estabelecidas para a subteoria		
7	A árvore de conceitos da subteoria classificação de insetos (Figura 5) está respondendo as questões de competência estabelecidas para a subteoria		
8	Os axiomas definidos para a subteoria classificação de insetos está complementando as informações existentes na árvore de conceitos (Figura 5) em relação às questões de competência estabelecidas para a subteoria		
9	A árvore de conceitos da subteoria processo de trabalho do entomólogo (Figura 6) está respondendo as questões de competência estabelecidas para a subteoria		
10	Os axiomas definidos para a subteoria processo de trabalho do entomólogo está complementando as informações existentes na árvore de conceitos (Figura 6) em relação às questões de competência estabelecidas para a subteoria		
11	Você concorda com a relação estabelecida entre as tarefas identificadas e as subteorias definidas		

12	Em qual (is) situação (ões) você consegue visualizar a utilização da Teoria do Domínio de Entomologia ?
----	---

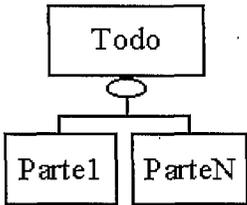
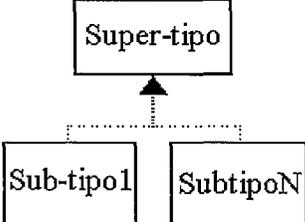
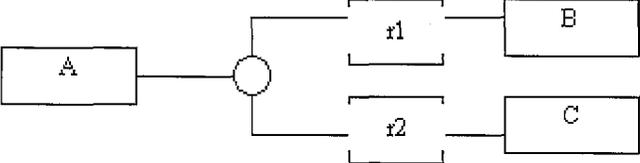
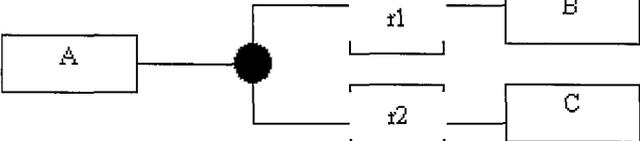
13	Se você conseguiu visualizar uma ou mais utilizações para a Teoria do Domínio de Entomologia, de que forma ela poderia ser utilizada ?
----	--

Anexo 3

LINGO: Linguagem Gráfica para descrever Ontologias

Este anexo apresenta as bases da linguagem gráfica para descrever ontologias, definida por FALBO (1998). LINGO uma linguagem gráfica capaz de representar a conceituação de um domínio, é constituída apenas por notações capazes de capturar certos axiomas de forma implícita. Ou seja, ao utilizar as notações de LINGO estamos, de fato, descrevendo o conjunto de axiomas que elas representam.

NOTAÇÃO	SIGNIFICADO	AXIOMAS EMBUTIDOS
<div style="border: 1px solid black; padding: 5px; display: inline-block; margin-bottom: 10px;">conceito</div> <div style="border: 1px solid black; padding: 5px; display: inline-block; margin-bottom: 10px;">relação</div>	São as primitivas básicas capazes de representar a conceituação de um domínio	—
<div style="display: flex; align-items: center; justify-content: center;"> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;">Atividade</div> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;">geração</div> <div style="text-align: center; margin-right: 10px;"> $1,1$ \downarrow </div> <div style="border: 1px solid black; padding: 5px;">Produto</div> </div>	Relação binária entre os conceitos de atividade e produto, em um contexto de manufatura, indicando que atividades geram produtos. Cardinalidades são usadas para mostrar quantas instâncias de um conceito podem participar da relação. Uma vez que a cardinalidade (0..n) não impõe nenhum axioma, ela não é representada. Também são válidas as relações de ordem superior, tais como relações ternárias.	O uso da cardinalidade (1,1) induz os seguintes axiomas: $(\forall p) (\text{produto}(p) \rightarrow (\exists a) (\text{geração}(p, a))$ (reflexo da cardinalidade mínima 1) $(\forall p, a_1, a_2) (\text{geração}(p, a_1) \wedge \text{geração}(p, a_2) \rightarrow a_1 = a_2)$ (reflexo da cardinalidade máxima 1)
<div style="display: flex; align-items: center; justify-content: center;"> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;">Atividade</div> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;">pré-atividade</div> <div style="text-align: center; margin-right: 10px;">dependência</div> <div style="border: 1px solid black; padding: 5px; margin-right: 10px;">pós-atividade</div> </div>	Relações entre instâncias de um mesmo conceito com uso de papéis.	$(\forall a_1, a_2) (\text{preatividade}(a_1, a_2) \leftrightarrow \text{positividade}(a_2, a_1))$

NOTAÇÃO	SIGNIFICADO	AXIOMAS EMBUTIDOS
	<p>Relação de composição (todo-parte). Ao utilizar uma relação todo-parte, estamos importando e aplicando uma teoria abstrata de composição de elementos ao conteúdo da ontologia em desenvolvimento. Trata-se de uma relação entre instâncias de conceitos, e como tal, cardinalidades devem ser representadas. Para representar este tipo de relação, utiliza-se uma linha cheia com uma pequena elipse junto ao todo.</p>	$(\forall x,y) (partede(x,y) \rightarrow \neg partede(y,x))$ $(\forall x,y,z) (partede(x,y) \wedge partede(y,z) \rightarrow partede(x,z))$ $(\forall x,y) (disjunto(x,y) \leftrightarrow \neg (\exists z) (partede(z,x) \wedge partede(z,y)))$ $(\forall x) (atômico(x) \leftrightarrow \neg (\exists y) (partede(y,x)))$
	<p>Associação sub-tipo-de, que se dá entre conceitos e não entre suas instâncias. Então, utiliza-se uma linha pontilhada para representá-la, com uma seta apontando para o super-tipo.</p>	$(\forall x, y) (subtipo(x, y) \rightarrow supertipo(y, x))$
	<p>Expressa condicionantes entre relações. Para expressar uma condicionante entre as relações r1 e r2, dizendo que se uma instância de A está relacionada com uma instância de B, então ela não pode estar relacionada com uma instância de C.</p>	<p>Ao utilizarmos esta notação, estamos assumindo que:</p> $(\forall a) ((a \in A) \wedge ((\exists b) (b \in B) \wedge r1(a,b)) \rightarrow \neg ((\exists c) (c \in C) \wedge r2(a,c)))$ $(\forall a) ((a \in A) \wedge ((\exists c) (c \in C) \wedge r2(a,c)) \rightarrow \neg ((\exists b) (b \in B) \wedge r1(a,b)))$
	<p>Estabelece uma condicionante de obrigatoriedade entre relações: se uma instância de A está relacionada com uma instância de B, então ela obrigatoriamente tem de estar relacionada com uma instância de C.</p>	$(\forall a) ((a \in A) \wedge ((\exists b) (b \in B) \wedge r1(a,b)) \rightarrow ((\exists c) (c \in C) \wedge r2(a,c)))$ $(\forall a) ((a \in A) \wedge ((\exists c) (c \in C) \wedge r2(a,c)) \rightarrow ((\exists b) (b \in B) \wedge r1(a,b)))$

LINGO permite a incorporação de *notações* para outros tipos de condicionantes, pois a proposta apresentada não deve ser considerada definitiva e, a linguagem deve ser considerada aberta, suscetível a extensões para capturar outras necessidades da modelagem no nível ontológico.