

**Universidade Federal de São Carlos**  
**Centro de Ciências Exatas e de Tecnologia**  
**Departamento de Computação**  
**Programa de Pós-Graduação em Ciência da Computação**

*Chem-Risk: Avaliação e Gerenciamento de Risco  
Químico em Alimentos Empregando Técnicas de  
Lógica Nebulosa e Ontologias*

*Walter Coelho Pereira de Magalhães Junior*

*Orientadora: Marilde Terezinha Prado Santos*

**São Carlos – SP**  
**Mai 2011**

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária da UFSCar**

M188cr

Magalhães Junior, Walter Coelho Pereira de.

Chem-risk : avaliação e gerenciamento de risco químico em alimentos empregando técnicas de lógica nebulosa e ontologias / Walter Coelho Pereira de Magalhães Junior. -- São Carlos : UFSCar, 2011.

147 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2011.

1. Data mining (Mineração de dados). 2. Ontologia. 3. Fuzzy logic. I. Título.

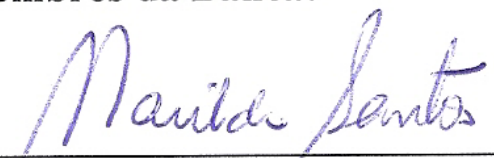
CDD: 005.741 (20ª)

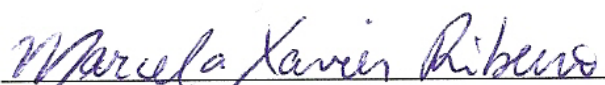
# “Chem-Risk: Avaliação e Gerenciamento de Risco Químico em Alimentos Empregando Técnicas de Lógica Nebulosa e Ontologias”

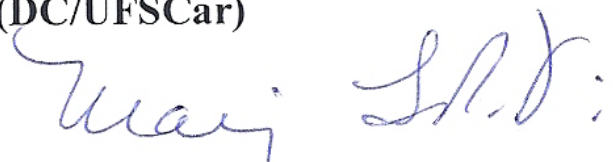
WALTER COELHO PEREIRA DE MAGALHÃES JÚNIOR


Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação

Membros da Banca:

  
\_\_\_\_\_  
Profa. Dra. Marilde Terezinha Prado Santos  
(DC/UFSCar)

  
\_\_\_\_\_  
Profa. Dra. Marcela Xavier Ribeiro  
(DC/UFSCar)

  
\_\_\_\_\_  
Profa. Dra. Marina Teresa Pires Vieira  
(UNIMEP)

  
\_\_\_\_\_  
Prof. Dr. Marcelo Bonnet Alvarenga  
(EMBRAPA/Gado de Leite/Juiz de Fora)

## ***Agradecimentos***

***A Deus,***

*Por todo amor, conhecimento e luz,  
Por tudo que nos permite aprender e realizar,  
Que todo fruto seja sempre ao bem a todos.*

***Ao Mestre Jesus, aos Protetores e Trabalhadores da Luz,***  
*Pelos ensinamentos sublimes  
Pela oportunidade de aprimoramento,  
Pela presença constante junto à nossa caminhada.*

***Aos meus queridos Pais,***  
*Que recebam esta singela homenagem.*

***A minha esposa Vanessa,***  
*Eterna companheira de luta, dedicação e perseverança.*

***A professora Marilde, Marcelo Bonnet e Leandro Feijó,***  
*Pelo abençoado inconformismo, orientações, incentivos e apoio constante.*

***Ao GBD, Professores e Companheiros de Estudo na UFSCar ,***  
*Pela luta, apoio e companherismo.*

***A todos que de alguma forma contribuíram com este trabalho, em especial a,***  
*Cristiane Yaguinuma, Daniel Krugler, Eduardo Cirilo, Vinícius Toledo, Hiromiti Nakagawa  
Waldomiro Barioni, Stanley Robson de Medeiros, Silvia Maria Fonseca, Luiz Manoel Silva,  
Maria Angélica Leite, Marcos Hott, Humberto Brandão, Glauco Rodrigues, Milton Morenz  
Victor Lima, Mario Junior, Leandro Rubiale, Inês Rodrigues, Margarida Salomão  
Marco Cristo, Francisco Louzada Neto, Pedro Ferreira Filho, Liane Biehl Printes*

## Resumo

---

Esta pesquisa apresenta a abordagem Chem-Risk a qual é aplicada ao domínio mundial da segurança química alimentar, visando contornar um problema frequentemente encontrado: dados com pouca semântica, apresentando baixa qualidade, que levam a resultados de baixa qualidade, pouco confiáveis e, portanto, inadequados para corretamente expressarem a realidade do domínio analisado. Em atendimento a esta finalidade, a abordagem Chem-Risk cria uma Plataforma Base Flexível para adequadamente pré-processar os resultados das análises laboratoriais em alimentos agropecuários, permitindo que inúmeras outras variáveis, aqui consideradas como fatores de risco alimentar, sejam consideradas nestas análises. Gerenciados por esta plataforma, conceitos da lógica *fuzzy* e ontologias viabilizam a incorporação dos novos fatores de risco, aprimoram a qualidade dos dados e produzem um conhecimento técnico superior, àquele oferecido pelas abordagens tradicionalmente empregadas no contexto das políticas públicas brasileiras, que visam avaliar os riscos químicos e maximizar a segurança química dos alimentos agropecuários. Devido à complexidade e escopo requeridos para a ontologia, uma nova abordagem de engenharia ontológica é apresentada. As técnicas de mineração de dados aqui empregadas desempenham um papel fundamental, auxiliando as autoridades governamentais e especialistas de risco a extraírem maiores níveis de conhecimentos não-triviais, a partir do conhecimento superior fornecido pela incorporação da lógica *fuzzy* e ontologia, durante a etapa de pré-processamento dos dados. Os resultados alcançados na etapa de mineração revelam uma semântica mais fidedigna à realidade da segurança química alimentar, complementam esta realidade e ampliam o nível de conhecimento especializado em favor de decisões não triviais. Por outro lado, se não houvesse adequado pré-processamento de dados a nível semântico, os resultados, apesar de aparentemente corretos, estariam encobrendo bases falsas e, portanto, estariam completamente equivocados perante a realidade da segurança química mundial dos alimentos agropecuários. Acreditamos que o que torna o nosso estudo de caso mais notável é que ele ajuda a melhorar a saúde da população mundial e, por extensão, salvar nossas próprias vidas.

## Abstract

---

This research presents the Chem-Risk approach which is applied to chemical food safety domain, in order to circumvent a problem often encountered: data with little semantic or presenting low quality results that lead to results of low quality, unreliable and therefore inadequate to the reality of domain analyzed. In response to this purpose, the approach creates a Flexible Base Platform to properly pre-process the results of laboratory tests in agricultural food, allowing so many other variables, here considered as food risk factors, are observed in this analysis. Managed by the platform, the concepts of fuzzy logic and ontologies enable the incorporation of new risk factors, improve data quality and produce a superior technical expertise to that offered by traditional approaches used in the context of Brazilian public policies that aim to assess chemical risks and maximize the chemical safety of agricultural food. Due to the complexity and scope required for the ontology, a new approach to ontological engineering is presented. Data mining techniques used here play a key role, helping government officials and experts to extract higher risk levels of non-trivial knowledge, from the superior knowledge provided by the incorporation of fuzzy logic and ontology during the preprocessing step. The results obtained in mining step reveal a more faithfully semantic to the reality of food chemical safety, complement this reality and increase the level of expertise in favor of non-trivial decisions. By the other hand, if there were no proper pre-processing at the semantic level, the results, although apparently correct, would be concealing its false basis and, and thus presenting results completely wrong before the reality of chemical safety of agricultural food worldwide. We believe that what makes our real case study more remarkable is that it helps to improve the health of global population and, by extension, save our own lives.

## Lista de Figuras

---

|   |    |
|---|----|
| Figura 1 – Processo de Gerência de Risco - Baseado em (ALENCAR; SCHMITZ, 2006) .....  | 8  |
| Figura 2 – Análise de Risco em Dados Contínuos .....  | 12 |
| Figura 3 – Funções de Pertinência Representativas do Conceito Adolescente por Conjunto Crisp ( $\mu_A$ ) Versus Conjunto <i>Fuzzy</i> ( $\mu_B$ ) na forma Trapezoidal (SANDRI; CORREA, 1999) ..... | 19 |
| Figura 4 – Representação Gráfica dos Conceitos Jovem, Adulto e Idoso.....   | 20 |
| Figura 5 – Representação Gráfica de Função de Pertinência Trapezoidal .....   | 20 |
| Figura 6 – Tipos de Representação de Ontologias (USCHOLD; GRÜNINGER, 2004) .....  | 26 |
| Figura 7 – Exemplo de uma Ontologia Simples (YAGUINUMA, 2007).....  | 27 |
| Figura 8 – Esquema Conceitual de Operação do Projeto OMCS-Br e sua Interação com Domínio de Conhecimento Associações Perigosas de Analitos - APA .....  | 31 |
| Figura 9 – Formato das Regras Utilizado na Linguagem SWRL.....  | 34 |
| Figura 10 – Atividades do Processo de Descoberta de Conhecimento em Bases de Dados .....  | 38 |
| Figura 11 – O Processo de Descoberta de Conhecimento em Bases de Dados Baseado em (TAN; STEINBACH; KUMAR, 2009) .....   | 38 |
| Figura 12 – Principais Tarefas de Mineração de Dados .....  | 41 |
| Figura 13 – Abordagem Vetor de Risco Ponderado Sugerida ao MAPA .....   | 46 |
| Figura 14 – Visão Geral da Abordagem Chem-Risk.....   | 47 |
| Figura 15 – Principais Etapas que Compõem a Plataforma Base Flexível - Macro-Etapa 1 da Abordagem Chem-Risk .....   | 50 |
| Figura 16 – Representação Gráfica da Etapa 1.1 da Plataforma Base Flexível .....  | 51 |
| Figura 17 – Conjuntos <i>fuzzy</i> Associados às Faixas de Risco para as Concentrações de Analitos Detectadas na Amostra de Matriz Agropecuária Analisada .....                                     | 53 |
| Figura 18 – Mapeamento Gráfico do Resultado Percentual (RP=35%) aos Conjuntos <i>Fuzzy</i> .....  | 55 |
| Figura 19 – Representação Gráfica da Etapa 1.2 da Plataforma Base Flexível .....  | 59 |
| Figura 20 – Representação Gráfica da Etapa 1.3 da Plataforma Base Flexível .....  | 61 |

|   |     |
|---|-----|
| Figura 21 – Conjuntos <i>Fuzzy</i> Pré-estabelecidos para Faixas de Risco de Amostras Conformes..   | 62  |
| Figura 22 – Mapeamento Gráfico do Percentual do Conformidade (PC=24.05%) aos Conjuntos <i>Fuzzy</i> Pré-estabelecidos para Faixas de Risco de Amostras Conformes.....             | 64  |
| Figura 23 – Conjuntos <i>Fuzzy</i> Pré-estabelecidos para Faixas de Risco de Amostras Não Conformes .....   | 65  |
| Figura 24 – Mapeamento Gráfico do Percentual de Não Conformidade (PNC=4.65%) aos Conjuntos <i>Fuzzy</i> Pré-estabelecidos para Faixas de Risco de Amostras Não Conformes.....     | 66  |
| Figura 25 – Representação Gráfica da Etapa 1.4 da Abordagem Chem-Risk.....  | 70  |
| Figura 26 – Representação Gráfica da Etapa 1.5 da Abordagem Chem-Risk.....  | 72  |
| Figura 27 – Representação Gráfica da Etapa 1.6 da Abordagem Chem-Risk.....  | 74  |
| Figura 28 – Sequência de Etapas, Sub-Etapas e Fluxo das Informações Trabalhadas pela Plataforma Base Flexível para Avaliação de Riscos Químicos Agropecuários- Macro-Etapa 1..    | 75  |
| Figura 29 – Representação Gráfica da Sub-Etapa 1.1.3 da Abordagem Chem-Risk .....   | 82  |
| Figura 30 – Avaliação da Concentração do Analito (ACA) Clortetraciclina – Abordagem Tradicional.....  | 83  |
| Figura 31 – Associação do Resultado Percentual (RP) aos Conjuntos <i>crisp</i> para Inferência do Nível de Risco da Concentração (NRC) - Abordagem Vetor de Risco Ponderado ..... | 84  |
| Figura 32 – Método para Avaliação de Amostra (AA) - Abordagem Tradicional.....  | 87  |
| Figura 33 – Método para Avaliação do NRA - Abordagem Vetor de Risco Ponderado .....   | 87  |
| Figura 34 – Questão de Competência (QC1) e Caso de Uso Correspondente .....   | 97  |
| Figura 35 – Atividade de Construção do Léxico de Referência (LR) Baseada no Método UPON (DE NICOLA; MISSIKOFF; NAVIGLI, 2009) .....   | 99  |
| Figura 36 – Processo de Produção e Incorporação do Léxico OCMS-Br ao Léxico de Referência (LR) – Abordagem ACATO (Estágio I).....   | 100 |
| Figura 37 – Processo de Produção e Incorporação do Léxico de Dados ao Léxico de Referência (LR) – Abordagem ACATO (Estágio I).....  | 102 |
| Figura 38 – Processo de Produção do Léxico OCMS-Br, Léxico dos Dados e Incorporação ao Léxico de Referência (LR) – Abordagem ACATO (Estágio I e II).....                          | 103 |



|  |     |
|--|-----|
| Figura 39 – Diagramas UML Modelando uma Parte do Cenário da Aplicação: a) Diagrama de Caso de Uso; b) Diagrama de Atividade; c) Diagrama de Classe .....                         | 107 |
| Figura 40 – Processo Identificado (PR2) para a Aplicação APA com Estrutura Parte-de .....  | 110 |
| Figura 41 – Exemplos de Relacionamentos Ontológicos para Aplicação APA .....   | 110 |
| Figura 42 – Fragmento de Código em OWL-DL da Ontologia Associações Perigosas de Analitos APA .....   | 111 |
| Figura 43 – Fragmento da Ontologia Associação Perigosa de Analitos – APA Ilustrando a Associação Perigosa APA-003 para Matriz Agropecuária Leite .....                           | 114 |
| Figura 44 – Resultados Comparativos dos Níveis de Riscos Químicos das Concentrações de Analitos - Abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk .....              | 117 |
| Figura 45 – Resultados Comparativos dos Níveis de Riscos Químicos de Amostras - Abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk .....                                | 119 |
| Figura 46 – Resultados Comparativos dos Níveis de Riscos Químicos de Lotes de Amostras - Abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk .....                       | 121 |
| Figura 47 – Migração de Níveis de Riscos de Amostras Aferidos na Abordagem Vetor de Risco Ponderado para os Novos Agrupamentos de Risco Sugeridos pela Abordagem Chem-Risk ..... | 122 |
| Figura 48 – Técnica de Associação Sugerindo Novas Associações Perigosas de Analitos - APA .....  | 123 |
| Figura 49 – Técnica de Associação Preditiva Sugerindo Inúmeras Tendências de Novas Associações Perigosas de Analitos – APA Envolvendo o Analito Clortetraciclina .....           | 124 |
| Figura 50 – Resultado da Aplicação da Técnica de Classificação para as Concentrações de Analitos (CA) – Algoritmo J48 (WEKA 3.6.1) .....   | 125 |
| Figura 51 – Técnica de Agrupamento de Dados Utilizando o Algoritmo Simple K-Means Sugerindo Novos Agrupamentos de Dados para a Abordagem Chem-Risk .....                         | 126 |
| Figura 52 – Ferramenta de Visualização para Série Histórica das Análises de Risco Associadas ao Analito “Aflatoxina M1” .....  | 128 |
| Figura 53 – Exemplo de Esquema Estrela Inicial para o Data Warehouse - ARQA .....  | 134 |
| .....  |     |
| Figura A.1 – Diagrama ER para o Banco de Dados Análise de Risco Químico Agropecuário - ARQA .....  | 142 |

|   |     |
|---|-----|
| Figura A.2 – Esquema Conceitual do Banco de Dados Análise de Risco Químico Agrppecuário - ARQA.....   | 142 |
| Figura A.3 – Script CriarTabelas.sql para Geração de Tabelas, Relacionamentos e Atributos – Banco de Dados Análise de Risco Químico Agropecuário - ARQA ..... | 144 |

## Lista de Tabelas

---

|   |     |
|---|-----|
| Tabela 1 – Análise de Risco em Dados Discretos .....  | 11  |
| Tabela 2 – Fragmento da Tabela Concentrações de Analitos – CA .....   | 52  |
| Tabela 3 – Conceitual das Tabelas LMR / TMC para Analitos .....   | 52  |
| Tabela 4 – Fragmento da Tabela de Analitos da Amostra Exibindo o NRCA Inferido a partir dos Conjuntos <i>Fuzzy</i> .....  | 55  |
| Tabela 5 – Tendências da Concentração do Analito – TDCA .....   | 56  |
| Tabela 6 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCA e TDCA .....  | 56  |
| Tabela 7 – Prazos Estimados de Ajustamento da Concentração do Analito (CA) aos Autocontroles do MAPA – PZCA .....   | 57  |
| Tabela 8 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCA, TDCA e PZCA .....  | 57  |
| Tabela 9 – Tabela de Custos Estimados de Ajustamento da Concentração do Analito (CA) aos Autocontroles do MAPA – CTCA .....   | 58  |
| Tabela 10 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCA, TDCA, PZCA e CTCA .....   | 58  |
| Tabela 11 – Fragmento da Tabela Analitos da Amostra Apresentando o Somatório ( $\Sigma$ ) dos Níveis de Risco NRCA e NRAA .....   | 59  |
| Tabela 12 – Fragmento da Tabela Analitos da Amostra Apresentando o Sistema de Alertas <i>Fuzzy</i> Baseado no Nível de Risco Agregado do Analito - NRAA .....                         | 60  |
| Tabela 13 – Níveis de Toxicidade de Analitos Não Conformes - NTA .....  | 67  |
| Tabela 14 – Fragmento da Tabela Analitos da Amostra Apresentando o Sistema de Alertas <i>Fuzzy</i> Re-hierarquizado para Analitos Não Conformes Detectados na Amostra Analisada ..... | 67  |
| Tabela 15 – Fragmento da Tabela Associações Perigosas de Analitos – APA .....   | 69  |
| Tabela 16 – Fragmento da Tabela Níveis de Intervenção e Ações Regulatórias–NIAR .....   | 71  |
| Tabela 17 – Avaliações de Risco Químico em Carne Bovina–Músculo–Amostra 5 / Lote 229 .....  | 116 |
| Tabela 18 – Avaliações de Risco Químico em Amostras de Leite .....  | 118 |
| Tabela 19 – Avaliações de Risco Químico em Lotes de Amostras de Mel .....   | 120 |

## Lista de Quadros

---

|   |     |
|---|-----|
| Quadro 1 – Recorte do Quadro V - Programa Nacional de Controle de Resíduos e Contaminantes em Leite – PNCRC / Animal / 2010 ..... | 14  |
| Quadro 2 – Recorte do Anexo I – Agrotóxicos Monitorados e Limites Máximos de Resíduos - PNCRC / Vegetal / 2010 .....              | 15  |
| Quadro 3 – Análise de Risco Periódica: Analitos e Amostras de Leite por Proprietário .....  | 124 |
| Quadro 4 – Análise de Risco Acumulada: Analitos e Amostras de Leite por Proprietário .....  | 124 |

## Lista de Abreviaturas e Siglas

---

|         |   |
|---------|---|
| AA      | Avaliação da Amostra  |
| ACA     | Avaliação da Concentração do Analito                        |
| ACATO   | Abordagem Colaborativa para Avaliação de Termos Ontológicos |
| APA     | Associações Perigosas de Analitos                           |
| ARQA    | Análise de Risco Químico Agropecuário                       |
| BPA     | Boas Práticas Agrícolas                                     |
| BPC     | Boas Práticas de Comercialização                            |
| BPD     | Boas Práticas de Distribuição                               |
| BPF     | Boas Práticas de Fabricação                                 |
| BPH     | Boas Práticas de Higiene                                    |
| BPM     | Boas Práticas de Manipulação                                |
| BPP     | Boas Práticas Pecuárias                                     |
| CA      | Concentração do Analito                                     |
| CCRC    | Coordenação de Controle de Resíduos e Contaminantes         |
| CTCA    | Custo de Ajustamento da Concentração do Analito             |
| DCBD    | Descoberta de Conhecimento em Bases de Dados                |
| DOU     | Diário Oficial da União                                     |
| EMBRAPA | Empresa Brasileira de Pesquisa Agropecuária                 |
| ETA     | Efeito Tóxico Aditivo                                       |
| ETP     | Efeito Tóxico de Pontenciação                               |
| ETR     | Efeito Tóxico Redutor                                       |

|           |  |
|-----------|--|
| ETS       | Efeito Tóxico Sinérgico                        |
| FAO       | Food and Agriculture Organization              |
| GR        | Glossário de Referência                        |
| ICA       | Intervalo de Conformidade da Amostra           |
| INCA      | Intervalo de Não Conformidade da Amostra       |
| IPPC      | International Plant Protection Convention      |
| ISO       | International Organization for Standardization |
| LA        | Léxico da Aplicação                            |
| LCA       | Local de Colheita da Amostra                   |
| LD        | Léxico do Domínio                              |
| LmC_NRAA  | Limite Mínimo de Conformidade do NRAA          |
| LMC_NRAA  | Limite Máximo de Conformidade do NRAA          |
| LmNC_NRAA | Limite Mínimo de Não Conformidade do NRAA      |
| LMNC_NRAA | Limite Máximo de Não Conformidade do NRAA      |
| LmC_NRAO  | Limite Mínimo de Conformidade do NRAO          |
| LMC_NRAO  | Limite Máximo de Conformidade do NRAO          |
| LmNC_NRAO | Limite Mínimo de Não Conformidade do NRAO      |
| LMNC_NRAO | Limite Máximo de Não Conformidade do NRAO      |
| LmC_NRCA  | Limite Mínimo de Conformidade do NRCA          |
| LMC_NRCA  | Limite Máximo de Conformidade do NRCA          |
| LmNC_NRCA | Limite Mínimo de Não Conformidade do NRCA      |
| LMNC_NRCA | Limite Máximo de Não Conformidade do NRCA      |
| LMDR      | Limite Mínimo de Desempenho Requerido          |

|         |  |
|---------|--|
| LMR     | Limite Máximo de Resíduo   |
| LR      | Limite de Referência (PNCRC/MAPA) ou Léxico de Referência (UPON) |
| MAPA    | Ministério da Agricultura, Pecuária e Abastecimento              |
| NIAR    | Nível de Intervenção e Ações Regulatórias                        |
| NRA     | Nível de Risco da Amostra  |
| NRAA    | Nível de Risco Agregado do Analito                               |
| NRAF    | Nível de Risco da Amostra Final                                  |
| NRAO    | Nível de Risco da Amostra Original                               |
| NRAP    | Nível de Risco da Associação Perigosa                            |
| NRC     | Nível de Risco da Concentração                                   |
| NRCA    | Nível de Risco da Concentração do Analito                        |
| NRL     | Nível de Risco do Lote   |
| NRLA    | Nível de Risco do Lote de Amostra                                |
| NTA     | Níveis de Toxicidade dos Analitos não conformes                  |
| OCMS-Br | Projeto Open Mind Common Sense - Brasil                          |
| OIE     | Organização Internacional de Epizootias                          |
| OMS     | Organização Mundial da Saúde                                     |
| ONU     | Organização das Nações Unidas                                    |
| OWL     | Web Ontology Language  |
| OWL-DL  | Web Ontology Language - Description Logic                        |
| PC      | Percentual de Conformidade                                       |
| PNC     | Percentual de Não Conformidade                                   |
| PNCRC   | Plano Nacional de Controle de Resíduos e Contaminantes           |

|        |   |
|--------|---|
| PZCA   | Prazo de Ajustamento da Concentração do Analito |
| QC     | Questão de Competência                          |
| RL     | Risco do Lote                                   |
| RP     | Resultado Percentual                            |
| RuleML | Rule Markup Language                            |
| SDA    | Secretaria de Defesa Agropecuária               |
| SIF    | Serviço de Inspeção Federal                     |
| SWRL   | Semantic Web Rule Language                      |
| SQWRL  | Semantic Query-Enhanced Web Rule Language       |
| TDCA   | Tendência da Concentração do Analito            |
| TMC    | Teor Máximo de Contaminante                     |
| UML    | Unified Modeling Language                       |
| UP     | Unified Process                                 |
| UPON   | Unified Process for Ontology                    |



# Sumário

---

## Parte I – Apresentação

|  |           |
|--|-----------|
| <b>Capítulo 1. Introdução.....</b>                                   | <b>2</b>  |
| 1.1 Motivação .....  | 2         |
| 1.2 Objetivo da Pesquisa .....                                       | 3         |
| 1.3 Organização do Trabalho.....                                     | 4         |
| <b>Capítulo 2. Análise de Risco .....</b>                            | <b>6</b>  |
| 2.1 Considerações Iniciais .....                                     | 6         |
| 2.2 Conceitos Associados ao Termo Risco .....                        | 6         |
| 2.2.1 Fator de Risco.....  | 6         |
| 2.2.2 Risco .....  | 6         |
| 2.2.3 Percepção da Existência de Riscos.....                         | 7         |
| 2.2.4 Gerência de Risco.....   | 7         |
| 2.3 Conceitos Praticados pelo PNCRC / MAPA.....                      | 8         |
| 2.3.1 Resíduos e Contaminantes .....                                 | 9         |
| 2.3.2 Limite Máximo de Resíduo - LMR.....                            | 9         |
| 2.3.3 Teor Máximo de Contaminante - TMC.....                         | 9         |
| 2.3.4 Limite Mínimo de Desempenho Requerido - LMDR .....             | 9         |
| 2.3.5 Limite de Referência para Tomada de Ação Regulatória.....      | 10        |
| 2.4 Considerações Finais .....                                       | 10        |
| <b>Capítulo 3. Análise de Risco Praticada no Âmbito do MAPA.....</b> | <b>11</b> |
| 3.1 Considerações Iniciais .....                                     | 11        |

|     |   |    |
|-----|---|----|
| 3.2 | Análise de Risco em Dados Discretos.....                    | 11 |
| 3.3 | Análise de Risco em Dados Contínuos.....                    | 12 |
| 3.4 | Plano Nacional de Controle de Resíduos e Contaminantes..... | 13 |
| 3.5 | Considerações Finais .....                                  | 15 |

## Parte II – Referencial Teórico

|  |   |    |
|--|---|----|
| <b>Capítulo 4. Lógica Nebulosa ou <i>Fuzzy</i> .....</b>                     | <b>18</b>   |    |
| 4.1  | Considerações Iniciais .....                          | 18 |
| 4.2  | Conceitos Básicos.....                                | 19 |
| 4.3  | Limitações Relevantes ao Contexto da Pesquisa.....    | 21 |
| 4.4  | Considerações Finais .....                            | 22 |
| <b>Capítulo 5. Ontologias.....</b>   | <b>23</b>   |    |
| 5.1  | Considerações Iniciais .....                          | 23 |
| 5.2  | Definições e Conceitos Básicos.....                   | 23 |
| 5.3  | Tipos e Representações .....                          | 24 |
| 5.4  | Engenharia de Ontologias.....                         | 28 |
| 5.4.1  | Método Unified Process for Ontology (UPON) .....      | 29 |
| 5.4.2  | Projeto Open Mind Common Sense (OMCS-Br).....         | 30 |
| 5.5  | Limitações Relevantes ao Contexto da Pesquisa.....    | 31 |
| 5.6  | Ambiente da Semantical Web Rule Language (SWRL) ..... | 33 |
| 5.7  | Considerações Finais .....                            | 35 |
| <b>Capítulo 6. Descoberta de Conhecimento em Bases de Dados – DCBD .....</b> | <b>37</b>   |    |
| 6.1  | Considerações Iniciais .....                          | 37 |
| 6.2  | O Processo DCBD .....                                 | 37 |

|     |  |    |
|-----|--|----|
| 6.3 | Pré-Processamento de Dados .....                   | 39 |
| 6.4 | Mineração de Dados .....                           | 39 |
| 6.5 | Pós-Processamento de Dados .....                   | 41 |
| 6.6 | Limitações Relevantes ao Contexto da Pesquisa..... | 42 |
| 6.7 | Considerações Finais .....                         | 42 |

## **Parte III – Trabalho de Pesquisa**

|   |  |    |
|---|--|----|
| <b>Capítulo 7. Abordagem Chem-Risk.....</b>                 | <b>45</b>  |    |
| 7.1   | Considerações Iniciais .....   | 45 |
| 7.2   | Macro-Etapas.....  | 47 |
| 7.2.1   | Macro-Etapa 1 – Pré-Processamento de Dados .....                     | 48 |
| 7.2.2   | Macro-Etapa 2 – Seleção // Formatação de Dados.....                  | 74 |
| 7.2.3   | Macro-Etapa 3 – Seleção // Formatação de Resultados .....            | 76 |
| 7.3   | Considerações Finais .....   | 77 |
| <b>Capítulo 8. Resultados do Estudo de Viabilidade.....</b> | <b>78</b>  |    |
| 8.1   | Considerações Iniciais .....   | 78 |
| 8.2   | Principais Contribuições.....  | 78 |
| 8.2.1   | Medidas Qualitativas de Desempenho .....                             | 79 |
| 8.2.2   | Medidas Quantitativas de Desempenho .....                            | 81 |
| 8.3   | Análise da Incorporação da Lógica <i>Fuzzy</i> .....                 | 82 |
| 8.3.1   | Estimativa do Nível de Risco da Concentração do Analito (NRCA) ..... | 82 |
| 8.3.2   | Estimativa do Nível de Risco da Amostra Original (NRAO) .....        | 86 |
| 8.3.3   | Conclusões Acerca da Incorporação da Lógica <i>Fuzzy</i> .....       | 89 |
| 8.4   | Análise da Incorporação da Ontologia APA .....                       | 90 |

|  |            |
|--|------------|
| 8.4.1 Engenharia da Ontologia APA: a Abordagem ACATO .....                   | 91         |
| 8.4.2 Estimativa do Nível de Risco da Associação Perigosa (NRAP) .....       | 114        |
| 8.5 Resultados da Incorporação da Lógica <i>Fuzzy</i> e Ontologias .....     | 115        |
| 8.5.1 Resultados das Avaliações de Risco nas Concentrações de Analitos ..... | 115        |
| 8.5.2 Resultados das Avaliações de Risco em Amostras .....                   | 117        |
| 8.5.3 Resultados das Avaliações de Risco em Lotes de Amostras .....          | 120        |
| 8.6 Resultados da Aplicação das Técnicas de Mineração de Dados .....         | 121        |
| 8.6.1 Migração dos Níveis de Risco das Amostras .....                        | 122        |
| 8.6.2 Aplicação da Técnica de Associação e Associação Preditiva .....        | 122        |
| 8.6.3 Aplicação da Técnica de Classificação .....                            | 125        |
| 8.6.4 Aplicação da Técnica de Agrupamento .....                              | 126        |
| 8.7 Análises estratégicas .....  | 127        |
| 8.8 Considerações Finais .....   | 128        |
| <b>Capítulo 9. Conclusões .....</b>  | <b>130</b> |
| 9.1 Considerações Iniciais .....   | 130        |
| 9.2 Sinopse dos Resultados Alcançados pela Abordagem Chem-Risk .....         | 131        |
| 9.3 Trabalhos Futuros .....  | 133        |
| <b>Referências .....</b>   | <b>135</b> |
| <b>Apêndice A. Banco de Dados ARQA .....</b>                                 | <b>141</b> |
| A.1 Considerações Iniciais .....   | 141        |
| A.2 Diagrama ER .....  | 141        |
| A.3 Esquema Conceitual .....   | 142        |
| A.4 Script CriarTabelas.sql .....  | 143        |
| A.5 Análise das Atualizações Estruturais .....                               | 146        |

# **Parte I – Apresentação**

# Capítulo 1

## Introdução

---

### 1.1 Motivação

Como uma questão de interesse mundial, melhorar a segurança química alimentar tem atraído a atenção de muitos governos e organizações no mundo inteiro, particularmente a atenção da Comissão do *Codex Alimentarius* – OMS / ONU que estabelece os índices mundiais de segurança química alimentar, referenciados nesta pesquisa como limites de referência para resíduos e contaminantes em produtos agrícolas. Como membro signatário das Nações Unidas, em atendimento às recomendações do *Codex Alimentarius*, o governo brasileiro instituiu o Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC, o qual é monitorado pela Secretaria de Defesa Agropecuária – SDA, do Ministério da Agricultura, Pecuária e Abastecimento – MAPA. Refletindo o melhor conhecimento técnico-científico disponível e estratégia nacional para alcançar reconhecimento internacional, atualmente o PNCRC permite que o Brasil ofereça garantias de segurança química, em nível mundial, para os produtos agrícolas produzidos internamente. Inúmeras amostras de matrizes agrícolas, tais como, músculo bovino, rim suíno, fígado de aves, urina de equino, leite, mel, arroz, alface, tomate, batata, uva, morango, peixe, camarão, etc., são coletadas diretamente das fontes produtivas em todo o país e analisadas por laboratórios credenciados pelo PNCRC. Em conformidade com as recomendações do *Codex*, o PNCRC estabelece quais analitos devem ser analisados de acordo com a matriz agropecuária monitorada.

Entretanto, foi observado que durante estas análises basicamente a concentração de cada analito (resíduo ou contaminante) estava sendo considerada, ignorando o fato de que, de acordo com (BONNET & SILVA, 2007), a produção agropecuária mundial está associada a um complexo de variáveis relacionadas às questões ambientais, sócio-político-econômicas, químicas, biológicas e físicas que são características dos sistemas vivos. Conjuntamente consideradas, essas variáveis formam um quadro de incertezas associado ao processo de produção agrícola brasileiro e mundial, as quais necessitam ser melhor conhecidas e controladas, visando mitigá-las em favor

de um processo produtivo mais eficaz, saudável, que apresente menores riscos à saúde humana e animal.

Devido à existência destas variáveis não consideradas, afirmamos existir perda de informações importantes que denunciam a baixa qualidade dos dados utilizados nas abordagens tradicionais, comumente empregadas para análise de riscos no contexto das políticas públicas brasileiras que visam maximizar a segurança química dos alimentos. Estes dados com pouca semântica afetam negativamente a qualidade do resultado final das avaliações dos riscos inerentes ao consumo dos produtos agrícolas produzidos no país. Por esta razão, a abordagem Chem-Risk considera inúmeros outros fatores de risco, além da concentração do analito, que incidem tanto ao nível dos analitos como ao nível das amostras das matrizes analisadas. Entretanto, para que estas variáveis adicionais (fatores de risco) estejam corretamente incorporadas ao processo de avaliação dos riscos, novos tipos de análises complexas são necessários, envolvendo a participação de diversos especialistas de risco e peritos agropecuários. Assim, para torná-los viáveis, a abordagem Chem-Risk aplica técnicas administrativas (KEPNER; TREGOE, 1991), de análise de risco (ALENCAR, SCHMITZ, 2006) e computacionais inerentes à lógica *fuzzy* (ZADEH, 1965) (SANDRI & CORREA, 1999) (KLIR & YUAN, 1995), Open Mind Common Sense Project (LIU & SINGH, 2004) e ontologias (GRUBER, 2009) (USCHOLD & GRÜNINGER, 2004), sincronizadas por uma plataforma base flexível para pré-processamento e mineração dos dados (CHAKRABARTI et al, 2009), (TAN; STEINBACH; KUMAR, 2009), (HAN; KAMBER, 2006) (WITTEN & FRANK, 2005) gerados a partir da massa de testes laboratoriais em amostras de matrizes agrícolas, realizadas no âmbito do PNCRC / MAPA.

## **1.2 Objetivo da Pesquisa**

Em conformidade com o contexto acima explicitado, esta pesquisa visou criar uma plataforma base flexível com capacidade para:

- Em um primeiro momento, permitir a incorporação de novos fatores de risco ao nível dos analitos e amostras analisadas, promovendo um adequado pré-processamento de dados por meio da aplicação de técnicas administrativo-computacionais a fim de que os dados, ou seja, as avaliações de risco químico em alimentos agropecuários, estivessem

mais exatos, precisos e apresentando uma semântica mais fidedigna à realidade da segurança química dos alimentos agrícolas brasileiros;

- Em um segundo momento, permitir nova preparação de dados visando submissão aos processos de mineração, a fim de que:
  - A melhor qualidade dos dados gerada na etapa de pré-processamento pudesse ser comprovada;
  - Os resultados obtidos na etapa de mineração de dados pudessem complementar e ampliar o nível de conhecimento superior, proporcionado anteriormente pela aplicação da lógica *fuzzy* e ontologias, em favor de um conhecimento e decisões não-triviais e inéditas por parte das autoridades governamentais brasileiras.

Em atendimento a este objetivo foi desenvolvida uma nova abordagem para avaliação de riscos químicos ao nível dos analitos, amostras e lotes de amostras de matrizes agropecuárias, denominada Chem-Risk, a ser aplicada no contexto das políticas públicas brasileiras que visam maximizar a segurança química dos alimentos agrícolas brasileiros.

### **1.3 Organização do Trabalho**

A presente dissertação está organizada em três partes, a saber:

- I - Apresentação, que inclui os capítulos de 1 a 3;
- II - Referencial Teórico, que inclui os capítulos de 4 a 6;
- III – Trabalho de Pesquisa, que inclui os capítulos 7, 8 e 9.

Os capítulos estão organizados conforme explicado a seguir.

O Capítulo 2 apresenta diversos conceitos relacionados à Análise de Risco, tendo por objetivo facilitar o entendimento de termos técnicos citados ao longo desta monografia.

O Capítulo 3 apresenta como é a análise de risco praticada no âmbito do Ministério da Agricultura, Pecuária e Abastecimento – MAPA, do Governo Federal Brasileiro, apresentando ao final, um recorte do Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC / Animal / Vegetal, promovido pela Coordenação de Controle de Resíduos e Contaminantes – CCRC, da Secretaria de Defesa Agropecuária – SDA / MAPA.



O Capítulo 4 apresenta conceitos de lógica difusa ou *fuzzy*, os quais contribuem para o aprimoramento da qualidade dos dados e resultados alcançados neste trabalho.

O Capítulo 5 apresenta conceitos relacionados às ontologias. Tais conceitos são empregados para, adicionalmente, aprimorar a qualidade dos dados e resultados aferidos, permitindo aumentar a semântica e expressividade da abordagem proposta.

O Capítulo 6 discute os conceitos referentes ao processo de Descoberta de Conhecimento em Bases de Dados – DCBD, ou mais comumente, *Knowledge Discovery in Databases - KDD*.

O Capítulo 7 discute a abordagem Chem-Risk, considerada o foco principal deste trabalho de pesquisa, apresentando como a abordagem pode prover conhecimento não-trivial e estratégico aos tomadores de decisão do PNCRC / MAPA.

O Capítulo 8 apresenta os resultados do estudo de viabilidade, a partir da aplicação da abordagem Chem-Risk sobre a massa de resultados de exames laboratoriais em amostras de matrizes agropecuárias monitoradas no país pelo PNCRC e mundialmente pela Comissão do Codex Alimentarius – FAO/OMS/ONU.

O Capítulo 9 apresenta as conclusões alcançadas por esta pesquisa e as sugestões de trabalhos futuros.

Finalmente, o Apêndice A apresenta o banco de dados Análise de Risco Químico Agropecuário - ARQA em maior detalhamento técnico.

# Capítulo 2

## Análise de Risco

---

### 2.1 Considerações Iniciais

Neste capítulo são apresentados os conceitos básicos referentes à Análise de Risco que são fundamentais para compreensão dos termos técnicos utilizados nesta pesquisa.

### 2.2 Conceitos Associados ao Termo Risco

#### 2.2.1 Fator de risco

Fator de risco é qualquer evento que possa prejudicar, de forma parcial ou totalmente, as chances de sucesso de um projeto ou processo, por exemplo, o processo de produção dos alimentos agropecuários brasileiros, no sentido de realizarem o que foi proposto dentro do prazo e fluxo de caixa estabelecidos.

#### 2.2.2 Risco

Risco é a probabilidade de um fator de risco assumir um valor prejudicial e, assim, prejudicar parcialmente ou totalmente as chances de sucesso de um projeto ou processo.

Falar sobre risco consiste em identificar:

- Os fatores (de risco);
- A probabilidade desses fatores assumirem valores prejudiciais;
- As consequências, caso esses fatores assumam valores prejudiciais.

O termo risco faz referência a eventos cuja realização é incerta. Para tratamento de eventos, cuja realização é certa, muitas vezes, basta que se empregue bom senso na análise. Entretanto, eventos, cuja realização é incerta, exigem conhecimento científico, disciplina, método e técnica. Com relação aos alimentos, os analistas de risco associam o termo risco à existência de perigos. Um perigo pode ser um agente biológico, químico ou físico que esteja presente no alimento e que seja capaz de provocar efeitos adversos à saúde, por exemplo:

- Perigos biológicos: bactérias que causam doenças (patógenos);

- Perigos químicos (resíduos e contaminantes): pesticidas, toxinas, drogas veterinárias, antibióticos, carrapaticidas, hormônios, etc.;
- Perigos físicos: prego, caco de vidro, pedra, pau, plástico, etc.

No contexto desta pesquisa, o termo risco encontra-se em função da probabilidade e severidade de um perigo químico ocasionar efeitos adversos à saúde humana ou animal.

### **2.2.3 Percepção da Existência de Riscos**

Análise de risco ou percepção da existência de riscos consiste em um conjunto de quatro atividades básicas (ALENCAR; SCHMITZ, 2006):

- Identificar os fatores de risco;
- Avaliar os impactos nos objetivos do processo e as probabilidades de sua ocorrência;
- Elaborar planos de contenção e contingência, os quais, respectivamente, traduzem ações ou procedimentos que objetivam eliminar ou reduzir as chances dos fatores de risco prejudicarem o processo, ou as ações que visam minimizar o impacto sobre o processo, caso, inevitavelmente, um ou mais fatores de risco assumam valores prejudiciais;
- Redefinir o planejamento do processo em função desta análise.

#### **Base Legal:**

Nacionalmente, a aplicação da Análise de Risco está contemplada pelo Decreto 5741/2106, que regulamenta a Lei 9712/98 e disciplina o capítulo de defesa agropecuária da Lei 8171 (Lei Agrícola). No âmbito internacional, a aplicação da Análise de Risco é disciplinada pela comissão do CODEX ALIMENTARIUS/FAO/OMS, pela OIE (Organização Internacional de Epizootias) e pela IPPC (Convenção Internacional de Proteção de Plantas).

### **2.2.4 Gerência de Risco**

Conjunto de atividades (processo) que visa maximizar os impactos positivos e minimizar os impactos negativos dos fatores de risco, mantendo-os em uma faixa de variabilidade e incerteza aceitáveis. O processo de Gerência de Risco é formado por dois sub-processos: na fase de planejamento envolve a Análise de Risco, na fase de execução, o Controle de Risco. A figura 1 ilustra o processo de Gerência de Risco baseado em (ALENCAR; SCHMITZ, 2006):

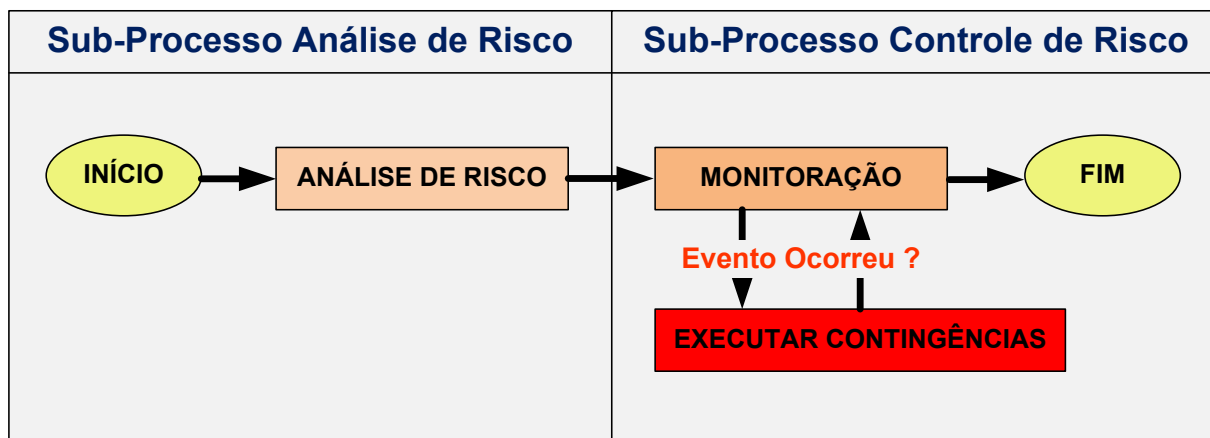


Figura 1 - Processo de Gerência de Risco - Baseado em (ALENCAR; SCHMITZ, 2006)

### 2.3 Conceitos Praticados pelo PNCRC / MAPA

No exercício de suas atribuições, o Ministério da Agricultura, Pecuária e Abastecimento - MAPA, de acordo com a Instrução Normativa nº 14 de 25/05/2009, instituiu novas abordagens e conceitos, visando compor um novo Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC em 2009, os quais encontram-se direta ou indiretamente referenciados nesta proposta, por estarem intrinsecamente relacionados à disciplina análise de risco. Entre as características assinaladas envolvendo os limites de referência (teores e limites residuais máximos), discutidos a seguir, destaca-se que são promovidos de forma nacional e rigorosamente observados no comércio internacional de alimentos, entre mercados consumidores cada vez mais exigentes, que exigem teores e limites em índices cada vez mais baixos. Atualmente, em alguns índices, observam-se divergências entre o *Codex Alimentarius* e a União Européia. Outra característica importante a ser observada nestes índices é que, caso os mesmos ainda não estejam estabelecidos para determinado resíduo ou contaminante, torna-se necessário utilizar a melhor evidência científica ou administrativa para realizar o gerenciamento do risco potencial, estabelecendo, assim, um Limite de Referência para a Tomada de Ação Regulatória em Limite Máximo de Resíduo (LMR) ou Teor Máximo de Contaminante (TMC), ditos “administrativos”. Estes e outros conceitos praticados no âmbito do PNCRC / MAPA serão explicados em seguida.

### **2.3.1 Resíduos e Contaminantes**

Especificamente para produtos agropecuários, entende-se como resíduos certos tipos de substâncias passíveis de serem encontradas nestes alimentos, tais como, antibióticos, carrapaticidas, hormônios, promotores de crescimento, etc. Analogamente, entende-se como contaminantes, as substâncias do tipo: nitratos, micotoxinas (aflatoxinas, ocratoxina A, patulina e toxinas Fusarium), metais pesados (chumbo, cádmio, mercúrio, metilmercúrio), dioxinas, entre outros.

### **2.3.2 Limite Máximo de Resíduo – LMR**

O Limite Máximo de Resíduo é definido como a quantidade máxima de resíduo oficialmente aceita no alimento, em decorrência de aplicação adequada em fase específica, desde sua produção até o consumo, expressa em partes do resíduo por milhão de partes de alimento (em litro ou peso) (ppm ou mg/kg). Estes limites são estabelecidos internacionalmente pela Comissão do *Codex Alimentarius*. Esse colegiado foi criado em 1963, num esforço conjunto da FAO (Food and Agriculture Organization) e Organização Mundial de Saúde (OMS), com o objetivo de proteger a saúde dos consumidores, assegurar transações comerciais justas e estabelecer um programa padronizado para controle de alimentos.

### **2.3.3 Teor Máximo de Contaminante – TMC**

Analogamente à União Européia, o MAPA fixa os teores máximos para certos contaminantes, de maneira a reduzir a presença desses contaminantes em determinados gêneros alimentícios em níveis tão baixos quanto razoavelmente possível, segundo as boas práticas de fabrico ou agrícolas. O objetivo é obter um nível elevado de proteção à saúde pública, em especial para os grupos sensíveis da população: crianças, lactentes, pessoas alérgicas, etc.

### **2.3.4 Limite Mínimo de Desempenho Requerido – LMDR**

O Limite Mínimo de Desempenho Requerido corresponde ao teor mínimo de uma substância a ser analisada em uma amostra que deve ser detectável e passível de confirmação. Pretende-se harmonizar o desempenho analítico de métodos relativos a substâncias para as quais não se encontre definido um limite permitido.

### **2.3.5 Limite de Referência para Tomada de Ação Regulatória**

Em função da presença do resíduo ou contaminante ser permitida ou proibida, e existir ou não LMR ou TMC respectivamente estabelecidos, o MAPA estabelece limites de referência para a quantidade observada dos mesmos em produtos comercializados dentro e fora do país.

## **2.4 Considerações Finais**

O agronegócio constitui o maior negócio brasileiro, sendo responsável por 37% dos empregos no país e 42,5% das exportações brasileiras, gerando divisas na ordem de 65,8 bilhões de dólares, que colocam o setor agropecuário como principal sustentador da balança comercial nacional, tornando o Brasil uma das maiores potências agropecuárias do planeta (MAPA, 2009). O país tem se destacado por sua liderança e elevada competitividade na produção de alimentos, posição esta alcançada principalmente devido a substanciais esforços em pesquisa, desenvolvimento e inovação, associado ao livre empreendedorismo, aos esforços governamentais, notadamente do Ministério da Agricultura, Pecuária e Abastecimento - MAPA, e às condições naturais favoráveis à farta e eficiente produção de alimentos. Entretanto, se por um lado estão largamente superadas as questões quantitativas da produção de alimentos no país, a continuidade sustentável da liderança e competitividade alcançadas ainda apresenta um desafio premente: a crescente exigência sobre a adequação dos sistemas de controle de qualidade e segurança dos alimentos em um mundo cada vez mais complexo. É nesse aspecto de segurança e qualidade que o país deverá enfrentar um de seus principais desafios nas próximas décadas, seja no âmbito do comércio nacional ou internacional de alimentos. Entretanto, os fatos indicam que o conceito de segurança é inseparável da questão probabilística, devendo assim a segurança ser examinada à luz de evidências científicas e, em particular, probabilísticas, no âmbito da Análise de Risco.

# Capítulo 3

## Análise de Risco Praticada no Âmbito do MAPA

### 3.1 Considerações Iniciais

As análises laboratoriais, coordenadas pelo PNCRC / MAPA, são desenvolvidas somente por laboratórios credenciados localizados em várias regiões do país. Uma vez que o foco deste trabalho está baseado nos resultados das análises de risco envolvendo resíduos e contaminantes, serão exemplificadas aplicações da análise de risco praticada no âmbito do MAPA, envolvendo dados discretos e contínuos. Os dados apresentados nas seções 3.2 e 3.3 são exclusivamente para fins didáticos e, por esta razão, podem não refletir os índices ou limites apresentados na seção 3.4.

### 3.2 Análise de Risco em Dados Discretos

Um exemplo simples de análise de risco praticada em dados discretos (enumeráveis) é a contagem de coliformes fecais (CF) em amostras de leite. Neste caso, obtêm-se uma Distribuição Binomial B(n,p) apresentando as características apresentadas na tabela 1:

TABELA 1 - Análise de Risco em Dados Discretos

| Amostras de Leite<br>(2 laticínios por região) | Contagem   |          |          |
|--|------------|----------|----------|
|  |            | ≤ 20     | > 20     |
| 1  | 10         | (-)      |          |
| 2  | 30         |          | (+)      |
| 3  | 34         |          | (+)      |
| 4  | 25         |          | (+)      |
| 5  | 22         |          | (+)      |
| 6  | 78         |          | (+)      |
| 7  | 12         | (-)      |          |
| 8  | 23         |          | (+)      |
| 9  | 10         | (-)      |          |
| 10   | 0          | (-)      |          |
| <b>Resultado</b>                               | <b>...</b> | <b>4</b> | <b>6</b> |

**LMR = 20  
definido por  
especialista**

Neste exemplo, a estimativa do risco ao amostrarmos um laticínio e encontrarmos valores de coliformes fecais acima do permitido é de 60%, ou seja, a probabilidade de encontrarmos um laticínio “não conforme” de acordo com o parâmetro CF. Observe que o desvio associado à porcentagem 60% é de 15% (0.15492) aproximado, obtido através da fórmula:  $\sigma = \sqrt{P(1-P)/n}$ . Desta forma, o intervalo de risco para o desvio padrão acima estaria entre 45 e 75 %.

### 3.3 Análise de Risco em Dados Contínuos

Analogamente ao exemplo acima, um exemplo simples de análise de risco praticada em dados contínuos pode ser a medição de promotores de crescimento em amostras de leite. Neste exemplo fictício apresentado na figura 2, para que o leite esteja fisiologicamente dentro dos padrões em relação à porcentagem de promotores, deve ter uma média ( $\mu$ ) de 22% e um desvio ( $\sigma$ ) de 2.

| Amostras | Medição |
|----------|---------|
| 01       | 20      |
| 02       | 28      |
| 03       | 23      |
| 04       | 20      |
| 05       | 21      |
| 06       | 27      |
| 07       | 25      |
| ...      | ...     |
| 100      | 23      |

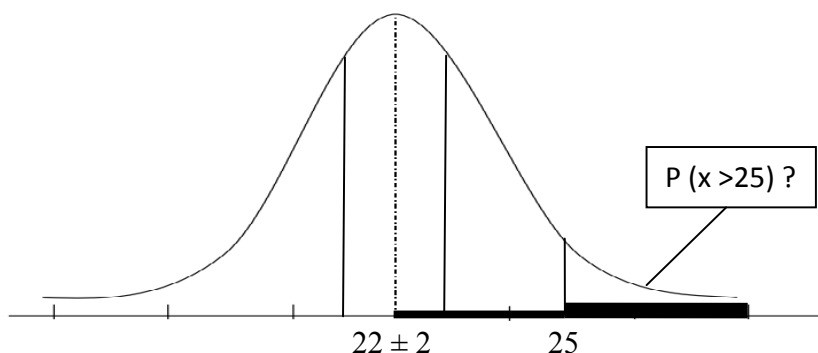


Figura 2 - Análise de Risco em Dados Contínuos

Observe que:  $\sigma = \sqrt{(x_i - \mu) / n}$

$$x \approx N(\mu, \sigma^2) \Rightarrow Z = (x - \mu) / \sigma \Rightarrow Z \approx N(0, 1) \therefore \text{Tabela Normal Z ou Normal padrão}$$

Questão:

Qual é o risco de coletarmos uma amostra de leite de um mercado em São Carlos-SP e encontrarmos mais do que 25% de promotores de crescimento? Ou seja  $P(x > 25) = ?$

Solução:

$$P((x - \mu) / \sigma > (25 - \mu) / \sigma) = P(Z > (25 - 22) / 2) = P(Z > 1.5) = 0.5 - P(Z \leq 1.5) = 0.5 - 0.4332 = 0,0668 \approx 7\%$$

Resposta: A estimativa do risco de encontrarmos uma amostra, apresentando mais que 25% de



promotores de crescimento no leite, é de 7%, para um mercado em São Carlos-SP.

### **3.4 Plano Nacional de Controle de Resíduos e Contaminantes**

Com vigência e atualizações anuais, o Plano Nacional de Controle de Resíduos e Contaminantes - PNCRC do Ministério da Agricultura, Pecuária e Abastecimento – MAPA, monitora a presença de resíduos de medicamentos veterinários e contaminantes em produtos de origem animal, quais sejam, carnes (bovina, aves, suína e eqüina), leite, mel, ovos e pescado, estendendo o monitoramento de agrotóxicos e outros contaminantes aos produtos de origem vegetal, tais como, as produções de mamão, maçã, abacaxi, alface, amendoim, arroz, banana, batata, castanha do brasil, limão, lima ácida, manga, melão, milho, morango, pimenta-do-reino, tomate, uva, entre outros.

As análises laboratoriais, realizadas somente nos laboratórios oficiais credenciados pelo PNCRC, levam em consideração as recomendações do *Codex Alimentarius* (fórum internacional de regularização de alimentos), estabelecidas pela Organização das Nações Unidas (ONU) e pela Organização Mundial de Saúde (OMS). Todos os estabelecimentos registrados no Serviço de Inspeção Federal (SIF) participam de sorteios semanais para coleta de amostras que serão examinadas no âmbito do PNCRC, inclusive aqueles habilitados para emitir certificado sanitário internacional. Importante ressaltar que somente por meio do PNCRC é possível ampliar as garantias de inocuidade dos produtos de origem animal e vegetal consumidos no Brasil e exportados. O trabalho é reconhecido internacionalmente pela União Européia e por países como Estados Unidos, China e Rússia, que realizam auditorias anuais, enviando missões que tem como objetivo avaliar a eficácia do Plano no monitoramento de resíduos e contaminantes dos produtos brasileiros exportados para esses mercados.

Segundo a Instrução Normativa nº 10, publicada no Diário Oficial da União (DOU), o número de substâncias de possíveis resíduos de medicamentos veterinários e contaminantes em produtos de origem animal aumentou de 130, em 2007, para 248, em 2008. Já em 2010 no PNCRC/Animal foram selecionados 163 diferentes analitos devendo permanecer próximo a 200 até o final de 2011. No PNCRC/Vegetal foram monitorados 154 analitos em 2010 devendo atingir a marca de 230 analitos monitorados até o final de 2011. No total, até o final de 2011, serão monitorados aproximadamente 430 analitos diferentes, sendo que um mesmo analito ou

grupo destes poderá estar relacionado ao monitoramento também de outras matrizes. No PNCRC/Animal, atualmente são 125 analitos pré-determinados para a carne bovina, 45 para o mel, 56 para pescado e camarão, 43 para o leite, etc. Para apreciação e melhor compreensão, o Quadro 1 apresenta um recorte do Quadro V do PNCRC / Animal / 2010 ilustrando os 43 analitos pré-determinados para a matriz leite, os grupos aos quais pertencem e os respectivos limites de referência a serem observados. Na sequência, o Quadro 2 apresenta um recorte do Anexo I referente ao PNCRC / Vegetal / 2010.

**QUADRO 1 – Recorte do Quadro V - Programa Nacional de Controle de Resíduos e Contaminantes em Leite – PNCRC / Animal / 2010**

*(<http://extranet.agricultura.gov.br/sislegis/loginAction.do?method=exibirTela>)*

**ANEXO  
QUADRO V - PROGRAMA DE CONTROLE DE RESÍDUOS E CONTAMINANTES EM LEITE – PNCRC/2010  
Plano de Controle de Resíduos e Contaminantes – Leite**

| Grupo                              | Analito             | Matriz | LIMITE DE REFERÊNCIA* | Nº de Itens de ensaio |       |             |    |
|------------------------------------|---------------------|--------|-----------------------|-----------------------|-------|-------------|----|
|                                    |                     |        | (µg/L)                |                       |       |             |    |
| Micotoxinas                        | Aflatoxina M1       | LEITE  | 0,5                   | 100                   |       |             |    |
| Pesticidas Organoclorados e PCBs** | Aldrin              | LEITE  | 8                     | 45                    |       |             |    |
|                                    | Alfa-HCH            |        | 4                     |                       |       |             |    |
|                                    | Lindane             |        | 10                    |                       |       |             |    |
|                                    | HCB                 |        | 10                    |                       |       |             |    |
|                                    | Dieldrin            |        | 8                     |                       |       |             |    |
|                                    | Endrin              |        | 2                     |                       |       |             |    |
|                                    | Heptacloro (d)      |        | 4                     |                       |       |             |    |
|                                    | DDT e Metabólitos   |        | 40                    |                       |       |             |    |
|                                    | Clordane (e)        |        | 2                     |                       |       |             |    |
|                                    | Mirex               |        | 2 *(IV)               |                       |       |             |    |
|                                    | Metoxicloro         |        | 10                    |                       |       |             |    |
|                                    | PCBs                |        | 2 *(IV)               |                       |       |             |    |
|                                    | Antiparasitários    |        | Abamectina (e)        |                       | LEITE | 10 *(II)    | 70 |
| Doramectina                        |                     | 15     |                       |                       |       |             |    |
| Eprinomectina                      |                     | 20     |                       |                       |       |             |    |
| Ivermectina (f)                    |                     | 10     |                       |                       |       |             |    |
| Moxidectina                        |                     | 10     |                       |                       |       |             |    |
| Albendazol                         |                     | 100    |                       |                       |       |             |    |
| Antimicrobianos                    | Clortetraciclina    | LEITE  | 100                   | 150                   |       |             |    |
|                                    | Oxitetraciclina     |        |                       |                       |       |             |    |
|                                    | Tetraciclina        |        |                       |                       |       |             |    |
|                                    | Sulfatiazol (b)     |        |                       |                       |       |             |    |
|                                    | Sulfametazina (b)   | LEITE  | 100                   | 75                    |       |             |    |
|                                    | Sulfadimetoxina (b) |        |                       |                       |       |             |    |
|                                    | Cloranfenicol       |        |                       |                       | LEITE | 0,30 *(III) | 60 |
|                                    | Carbaril            |        |                       |                       |       |             |    |
|                                    | Carbofuran          |        |                       |                       |       |             |    |
| Carbamatos                         | Metomil             | LEITE  | 20                    | 75                    |       |             |    |
|                                    | Propoxur            |        |                       |                       |       |             |    |
|                                    | Aldicarb            |        |                       |                       |       |             |    |
|                                    | Oxamil              |        |                       |                       |       |             |    |
|                                    | Metiocarb           |        |                       |                       |       |             |    |
|                                    | Clorpirifos Etil    |        |                       |                       | LEITE | 10          | 75 |
| Clorpirifos Metil                  |                     |        |                       |                       |       |             |    |
| Diazinon                           |                     |        |                       |                       |       |             |    |
| Metamidofós                        |                     |        |                       |                       |       |             |    |
| Mevinfós                           |                     |        |                       |                       |       |             |    |
| Acefato                            |                     |        |                       |                       |       |             |    |
| Pirimifós Metil                    |                     |        |                       |                       |       |             |    |
| Paration                           |                     |        |                       |                       |       |             |    |
| Pirimifós Etil                     |                     |        |                       |                       |       |             |    |
| Metidation                         |                     |        |                       |                       |       |             |    |
| Azinfós Metil                      |                     |        |                       |                       |       |             |    |
| Azinfós Etil                       |                     |        |                       |                       |       |             |    |

\*\* O resultado expresso no Certificado Oficial de Análise refere-se à concentração do analito no leite. Calcula-se a porcentagem de gordura na amostra e converte-se o resultado.

**QUADRO 2 – Recorte do Anexo I – Agrotóxicos Monitorados e Limites Máximos de Resíduos  
- PNCRC / Vegetal / 2010**

(<http://extranet.agricultura.gov.br/sislegis/loginAction.do?method=exibirTela>)

**SECRETARIA DE DEFESA AGROPECUÁRIA**

**INSTRUÇÃO NORMATIVA Nº 21, DE 2 DE SETEMBRO DE 2010**

O SECRETÁRIO SUBSTITUTO DE DEFESA AGROPECUÁRIA DO MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO, no uso das atribuições que lhe conferem os arts. 10 e 42 do Anexo I do Decreto nº 7.127, de 4 de março de 2010, tendo em vista o disposto na Portaria MAPA nº 45, de 22 de março de 2007, na Instrução Normativa SDA nº 42, de 31 de dezembro de 2008, e o que consta do Processo nº 21000.006279/2010-95, resolve:

Art. 1ª Aprovar os Programas Nacionais de Controle de Resíduos e Contaminantes para as culturas agrícolas de abacaxi, alface, alho, amendoim, arroz, banana, batata, café, castanha-do-brasil, feijão, laranja, limão, lima ácida, maçã, mamão, manga, melão, milho, morango, pimenta do reino, pimentão, soja, tomate, trigo e uva de que trata o Plano Nacional de Controle de Resíduos e Contaminantes em Produtos de Origem Vegetal - PNCRC/Vegetal para o ano-safra 2010/2011, na forma dos Anexos à presente Instrução Normativa.

Art. 2ª Os custos de envio das amostras integrantes do PNCRC/Vegetal ao laboratório oficial ou credenciado pertencente à Rede Nacional de Laboratórios Agropecuários do Sistema Unificado de Atenção à Sanidade Agropecuária correrão por conta do estabelecimento produtor.

Art. 3ª O PNCRC/Vegetal 2010/2011 inicia-se em julho de 2010 e tem vigência até junho de 2011.

Art. 4ª Esta Instrução Normativa entra em vigor na data de sua publicação.

JOSÉ GUILHERME TOLLSTADIUD LEAL

**ANEXO I**

**TABELA DE AGROTÓXICOS MONITORADOS E LIMITES MÁXIMOS DE RESÍDUOS**

| INGREDIENTE ATIVO | LIMITES MÁXIMOS DE RESÍDUOS (mg/kg) |          |          |          |          |          |          |
|-------------------|-------------------------------------|----------|----------|----------|----------|----------|----------|
|                   | Abacaxi                             | Alface   | Alho     | Amendoim | Arroz    | Banana   | Batata   |
| ACEFATO           | NPC                                 | NPC      | NPC      | 0,20     | NPC      | NPC      | 0,20     |
| ACETAMIPRIDO      | NPC                                 | NPC      | NPC      | NPC      | 0,05     | NPC      | 0,50     |
| ALACLORO          | NPC                                 | NPC      | NPC      | 0,05     | NPC      | NPC      | NPC      |
| ALDICARBE         | NPC                                 | NPC      | NPC      | NPC      | NPC      | NPC      | 1,00     |
| ALDICARBE SULFONA | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| ALDRIN            | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| ALETRIN           | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| ALFA-CIPERMETRINA | NPC                                 | NPC      | NPC      | NPC      | 0,05     | NPC      | 0,01     |
| AMETRINA          | 0,02                                | NPC      | NPC      | NPC      | NPC      | 0,07     | NPC      |
| AMICARBAZONA      | NPC                                 | NPC      | NPC      | NPC      | NPC      | NPC      | NPC      |
| AZINFÓS-ETILICO   | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| AZINFÓS-METÍLICO  | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| AZOXISTROBINA     | NPC                                 | 1,00     | 0,20     | 0,20     | 0,10     | 0,20     | 0,10     |
| BENFLURALIN       | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| BENOMIL           | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| BETA-CIPERMETRINA | NPC                                 | NPC      | NPC      | NPC      | 0,30     | NPC      | NPC      |
| BIFENTRINA        | NPC                                 | NPC      | NPC      | NPC      | 0,70     | 0,02     | 0,03     |
| BIOALETRINA       | PROIBIDO                            | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO | PROIBIDO |
| BITERTANOL        | NPC                                 | NPC      | NPC      | 0,20     | NPC      | NPC      | NPC      |
| BOSCALIDA         | NPC                                 | NPC      | 0,05     | NPC      | NPC      | NPC      | 0,05     |

### 3.5 Considerações Finais

O Programa Nacional de Controle de Resíduos e Contaminantes - PNCRC conduzido pelo Ministério da Agricultura, Pecuária e Abastecimento, vem despontando como um caso vitorioso

no contexto brasileiro de políticas públicas destinadas à maximização da segurança química de alimentos agropecuários. Entretanto, os conceitos aqui expostos e discutidos evidenciaram a necessidade de evolução da agenda do PNCRC rumo a uma arquitetura consistente de análise de risco, que passasse a considerar o triângulo virtuoso “dado-conhecimento-inteligência” por ela precipitado, de forma a expor e discutir as questões infra-estruturais, organizacionais/administrativas, políticas e legislativas associadas à referida evolução do programa.

Entende-se que, no contexto da necessidade de evolução da agenda do PNCRC rumo a uma arquitetura consistente de análise de risco, a abordagem Chem-Risk, formalmente apresentada na Parte III - Capítulo 7 desta dissertação, seja a resposta à esta solicitação promulgada pelos analistas de risco, peritos agropecuários e autoridades do PNCRC.

*Nota* – Os resultados deste trabalho de pesquisa foram formalmente apresentados e validados durante a Reunião Anual de Gestores Nacionais do PNCRC / Animal / Vegetal, em Novembro / 2010.

## **Parte II – Referencial Teórico**

# Capítulo 4

## Lógica Nebulosa ou *Fuzzy*

---

### 4.1 Considerações Iniciais

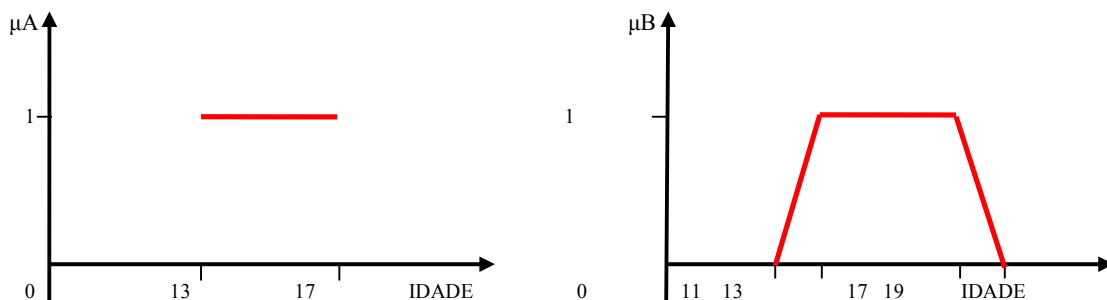
No capítulo 1 desta dissertação citamos que a produção agropecuária mundial está associada a um complexo de variáveis relacionadas às questões ambientais, sócio-político-econômicos, químicos, biológicos e físicos que são características dos sistemas vivos. Conjuntamente consideradas, essas variáveis formam um quadro de *incertezas* associadas ao processo de produção agrícola mundial. Segundo (SANDRI; CORREA, 1999) os dois principais aspectos de imperfeição da informação são a *imprecisão* e a *incerteza*. Os autores afirmam que as teorias mais conhecidas para tratamento da imprecisão e incerteza são, respectivamente, a teoria clássica dos conjuntos e a teoria das probabilidades, esclarecendo que, embora úteis, nem sempre conseguem captar a riqueza das informações fornecidas por seres humanos. Ainda segundo estes autores, a teoria dos conjuntos nebulosos (ZADEH, 1965) é capaz de tratar o aspecto vago da informação, ou seja, aspectos como imprecisão, ambigüidade e incerteza, que são características do pensamento humano. Juntamente com a teoria de possibilidades (ZADEH, 1978), tornam possível o tratamento destes aspectos para um conjunto de informações, quando em ambiente formal único. Utilizada em um contexto lógico, como o de sistemas baseados em conhecimento, a teoria dos conjuntos nebulosos é conhecida como lógica nebulosa, difusa ou *fuzzy*. A lógica *fuzzy* é considerada uma subárea da inteligência artificial e objetiva, em modo aproximado, modelar o raciocínio humano, procurando imitar a habilidade humana de tomar decisões racionais em um ambiente de incerteza e imprecisão. Por não ser exata, difere de outras mais conhecidas na literatura, como a teoria dos conjuntos clássica e a teoria das probabilidades, as quais baseiam-se na lógica binária [0, 1] ou booleana [verdadeiro, falso], tratando a informação com pouca flexibilidade. A lógica *fuzzy* estende a lógica booleana, pois ao invés de somente permitir dois valores (0 e 1), permite representar a informação imprecisa de forma mais adequada ao considerar uma gama infinita de valores neste intervalo, ou seja, [0; 1]. Potente e abrangente, a lógica *fuzzy* apresenta inúmeras definições, operações e operadores, os quais encontram-se além do escopo requerido por esta pesquisa.

## 4.2 Conceitos Básicos

Em continuidade às considerações iniciais, segundo (KLIR; YUAN, 1995), o mapeamento de um determinado elemento a um valor escalar entre 0 e 1 é realizado por meio de conjuntos *fuzzy*, os quais determinam o grau de pertinência deste elemento aos conjuntos. A função que realiza este mapeamento é denominada função de pertinência e o conjunto definido por ela denomina-se conjunto nebuloso, difuso ou *fuzzy*. Desta forma, torna-se possível tratar a *imprecisão* atribuindo-lhe determinado grau de veracidade ou falsidade. Formalmente, em (SANDRI; CORREA, 1999), um conjunto nebuloso A do universo  $\Omega$  é definido por uma função de pertinência  $\mu_A : \Omega \rightarrow [0; 1]$  que associa a cada elemento x de  $\Omega$  o grau  $\mu_A(x)$  com o qual x pertence a A:

- $\mu_A(x) = 1$  indica que x é completamente compatível com A;
- $\mu_A(x) = 0$  indica que x é completamente incompatível com A;
- $0 < \mu_A(x) < 1$  indica que x é parcialmente compatível com A, com grau  $\mu_A(x)$ .

A título de comparação, um conjunto A na teoria dos conjuntos clássica pode ser visto como um conjunto nebuloso específico, denominado usualmente de “*crisp*”, para o qual  $\mu_A : \Omega \rightarrow \{0,1\}$ , ou seja, a pertinência é do tipo “tudo ou nada”, “sim ou não”, e não gradual como nos conjuntos *fuzzy*. A figura 3 apresenta graficamente a diferença entre o conceito “adolescente” em relação à variável idade, por meio de um conjunto *crisp* (A) e de um conjunto *fuzzy* (B):



**Figura 3 – Funções de Pertinência Representativas do Conceito Adolescente por Conjunto *Crisp* ( $\mu_A$ ) Versus Conjunto *Fuzzy* ( $\mu_B$ ) na forma Trapezoidal (SANDRI; CORREA, 1999)**

Embora o exemplo apresente a forma trapezoidal para representação do conjunto *fuzzy* adolescente, existem outras formas de representação comumente utilizadas, tais como: triangular, exponencial, gaussiana, etc.

Observa-se, neste exemplo, que o conjunto *crisp* A não exprime completamente o conceito “adolescente”, pois uma pessoa com 12 anos e 11 meses seria considerada incompatível com este conceito, ocasionando classificação arbitrária. Por outro lado, o conjunto *fuzzy* B permite exprimir que qualquer pessoa com idade entre 13 e 17 anos é um adolescente e que abaixo de 11 ou acima de 19 não é considerado adolescente. Observa-se ainda que para os intervalos [11;13] e [17;19], considera-se uma pessoa tanto mais adolescente quanto mais próxima sua idade esteja de 13 ou 17 anos. Assim, o valor de  $\mu_B(x)$  representa o grau de pertinência de cada elemento  $x$ , representado pelo eixo  $x$ , ao conjunto *fuzzy* B. Quanto mais próximo de 1 for o valor de  $\mu_B(x)$ , maior será o grau de pertinência de  $x$  ao conjunto *fuzzy* B.

Visando facilitar a assimilação destes conceitos básicos relacionados à lógica difusa, ainda com relação à variável idade, em (KLIR; YUAN, 1995) tem-se uma definição de três conjuntos difusos para os conceitos (termos lingüísticos) jovem, adulto e idoso. A idade de um ser humano é considerada no intervalo real [0; 80], conforme apresentado na figura 4:

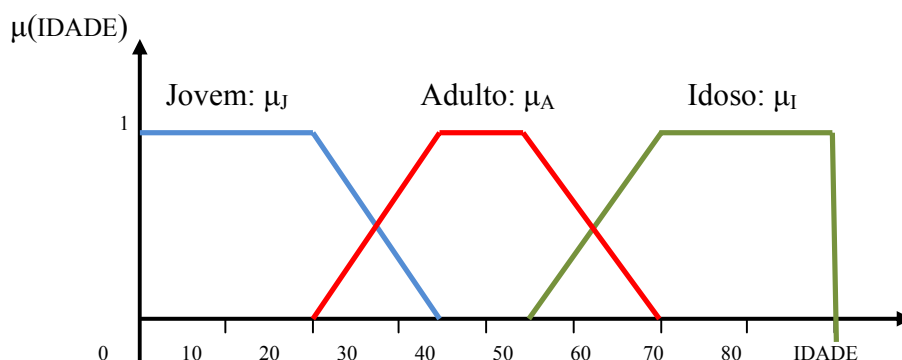


Figura 4 – Representação Gráfica dos Conceitos Jovem, Adulto e Idoso

Uma vez que esta forma de representação trapezoidal encontra-se diretamente aplicada no estudo de caso dessa pesquisa, é apresentada, na figura 5, sua representação gráfica:

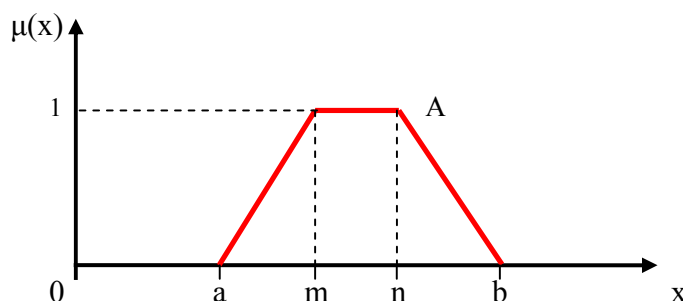


Figura 5 – Representação Gráfica de Função de Pertinência Trapezoidal



Analicamente, o conjunto *fuzzy* observado na figura 5 é traduzido por:

$$A(x) = \begin{cases} 0, & \text{se } x \leq a \\ (x - a) / (m - a), & \text{se } x \in (a, m) \\ 1, & \text{se } x \in [m, n] \\ (b - x) / (b - n), & \text{se } x \in (n, b) \\ 0, & \text{se } x \geq b \end{cases}$$

Logo, as funções de pertinência definidas para o conjunto Jovem, Adulto e Idoso, são:

- $\mu_J(x) = 1$  para  $x \leq 20$
- $\mu_J(x) = (35 - x) / 15$  para  $20 < x < 35$
- $\mu_J(x) = 0$  para  $x \geq 35$
  
- $\mu_A(x) = 1$  para  $35 \leq x \leq 45$
- $\mu_A(x) = (x - 20) / 15$  para  $20 < x < 35$
- $\mu_A(x) = (60 - x) / 15$  para  $45 < x < 60$
- $\mu_A(x) = 0$  para  $x \leq 20$  ou  $x \geq 60$
  
- $\mu_I(x) = 1$  para  $x \geq 60$
- $\mu_I(x) = (x - 45) / 15$  para  $45 < x < 60$
- $\mu_I(x) = 0$  para  $x \leq 45$

Uma variável linguística é definida como uma quádrupla  $(X, \Omega, T(X), M)$ , onde:

- $X$  é o nome da variável;
- $\Omega$  é o universo de  $X$ ;
- $T(X)$  é o conjunto de nomes para os valores de  $X$ ;
- $M$  é uma função que associa uma função de pertinência a cada elemento de  $T(X)$

Para o contexto deste trabalho de pesquisa, termos linguísticos referem-se aos elementos de  $T(X)$  ou às suas funções de pertinência.

### 4.3 Limitações Relevantes ao Contexto da Pesquisa

Embora apresente muitas vantagens, a lógica *fuzzy* também apresenta algumas limitações relevantes consideradas no contexto desta pesquisa, entre as quais foram relacionadas:

- Estabilidade: Sistemas *fuzzy* são estáticos não sendo capazes de se adaptarem a contextos extremamente dinâmicos;
- Incapacidade de aprendizagem: Por serem estáticos, os sistemas *fuzzy* não conseguem “aprender”.

#### **4.4 Considerações Finais:**

O modelo *fuzzy* traduz uma família de modelos matemáticos aplicados ao tratamento da *imprecisão* e *incerteza*, que coaduna perfeitamente com inúmeras situações vivenciadas em nosso cotidiano, para as quais exige-se análises cada vez mais complexas e inteligentes. Como veremos no capítulo 7, esse é o caso das análises de risco químico agropecuário desenvolvidas pelo Ministério da Agricultura, Pecuária e Abastecimento – MAPA, do Governo Federal Brasileiro. Por meio dos conceitos apresentados, é possível deduzir que a lógica *fuzzy*, especificamente considerando os números e conjuntos *fuzzy* apresentados, ao oferecer um tratamento para imprecisão e incerteza, contribui decisivamente para que a abordagem Chem-Risk alcance maior exatidão, precisão e, conseqüentemente, consiga expressar de forma mais fidedigna a realidade da segurança química dos produtos agropecuários brasileiros. Por meio desta técnica consegue-se reduzir parte da perda de informações valiosas que comprometem a qualidade da tomada de decisão realizada pelas autoridades do PNCRC / MAPA.

Por fim, acredita-se que associada às ontologias, as quais serão discutidas no próximo capítulo, e gerenciadas por uma plataforma computacional de base flexível apresentada no capítulo 7, as limitações da lógica *fuzzy* relacionadas no item 4.3, possam ser contornadas em nível satisfatório. Associadas à lógica *fuzzy*, as ontologias também contribuem para dirimir outra parte da perda de informações relacionadas ao processo de avaliação de riscos químicos dos produtos agropecuários brasileiros.

# Capítulo 5

## Ontologias

---

### 5.1 Considerações Iniciais

Para melhor compreender a contribuição das ontologias ao processo de organização, modelagem, compartilhamento e reuso da informação semântica, utilizados pela abordagem Chem-Risk, torna-se necessário, primeiramente, assimilar os principais conceitos e definições, conhecer os principais tipos e formas de representação, entender as principais motivações para sua criação e utilização, além de algumas limitações relevantes ao contexto desta pesquisa, que motivaram a adoção de técnicas colaborativas ao processo de engenharia e raciocínio integrado.

### 5.2 Definições e Conceitos Básicos

Na literatura encontram-se diversas definições e conceitos relacionados ao termo ontologia. Embora existam outras definições e considerações importantes, acredita-se que os autores abaixo, citados em ordem cronológica, ofereçam um bom entendimento sobre o termo:

Uma ontologia define um domínio de conhecimento, ou, mais formalmente, especifica uma conceitualização acerca dele (GRUBER, 1995).

No sentido filosófico, ontologias referem-se a um sistema particular de categorias que descreve uma determinada visão de mundo (GUARINO, 1998).

Uma ontologia é organizada em hierarquias de conceitos ou taxonomias e, pelo fato de não refletirem idealmente nenhum formalismo específico, pode-se considerar uma ontologia como a materialização do nível do conhecimento (CLARK, 1999).

Uma ontologia é criada com base no conhecimento de especialistas em um determinado domínio e define as regras que regulam a combinação entre os termos e relações de um domínio de conhecimento. O que se busca, em última instância, são melhorias nos processos de recuperação da informação (ALMEIDA; BAX, 2003).

Fornece uma representação codificada de um entendimento sobre determinado domínio compartilhado, expresso em termos e conceitos, em concordância por uma comunidade de pessoas (CULLOT et al, 2003).

Descrição formal de um domínio de discurso, objetivando compartilhá-lo em diferentes aplicações e expresso em uma linguagem que possa ser usada para raciocínio (NOY, 2004).

As ontologias oferecem um meio de se lidar com a representação de recursos de informação: o modelo de domínio descrito por uma ontologia pode ser usado como uma estrutura unificadora para dar semântica e uma representação comum à informação (FALBO; RUY; PEZZIN, 2004).

Amplamente citado em trabalhos relacionados à ciência da computação, Gruber (GRUBER, 2009) cita que: “No contexto da computação e ciências da informação, uma ontologia define um conjunto de primitivas de representação com o qual modela-se um domínio de conhecimento ou de discurso. As primitivas de representação são tipicamente classes (ou conjuntos), atributos (ou propriedades) e relacionamentos (ou relações entre os membros da classe). As definições das primitivas de representação incluem informações sobre o seu significado e as restrições à sua aplicação logicamente consistente. No contexto dos sistemas de banco de dados, a ontologia pode ser vista como um nível de abstração de modelos de dados, semelhante aos modelos hierárquicos e relacionais, mas destina-se à modelagem de conhecimento sobre os indivíduos, seus atributos, e suas relações com outros indivíduos. Ontologias são tipicamente especificadas em linguagens que permitem realizar uma abstração a partir de estruturas de dados e estratégias de implementação; na prática, as linguagens de ontologias estão mais próximas ao poder expressivo da lógica de primeira ordem que as linguagens utilizadas para modelar bancos de dados. Por esta razão, ontologias são ditas estarem ao "nível semântico", enquanto esquemas de bancos de dados são modelos de dados a “nível lógico ou físico”. Devido à sua independência dos modelos de dados de nível mais baixo, ontologias são utilizadas para integração de bancos de dados heterogêneos, habilitando a interoperabilidade entre sistemas diferentes, e especificando interfaces para serviços independentes baseados no conhecimento. Atualmente existem linguagens padrão e uma variedade de ferramentas comerciais e de código aberto para criar e trabalhar com ontologias.

### **5.3 Tipos e Representações**

Segundo (GOMES-PEREZ, 1999) existem diferentes tipos de ontologias considerando seu grau de genericidade:

- Ontologias de Representação: definem as primitivas de representação (frames, axiomas, atributos e outros) de forma declarativa;
- Ontologias Gerais: trazem definições abstratas necessárias para a compreensão de aspectos do mundo, como, tempo, processos, papéis, espaço, seres, coisas, etc.;
- Ontologias Centrais ou Genéricas: definem os ramos de estudos de uma área e/ou conceitos mais genéricos e abstratos desta área. Por exemplo, regras básicas do direito;
- Ontologias de Domínio: tratam de um domínio mais específico de uma área genérica de conhecimento, como direito tributário, microbiologia, etc;
- Ontologias de Aplicação: procuram solucionar um problema específico dentro de um domínio, como, identificar doenças do coração a partir de uma ontologia de domínio de cardiologia.

O autor afirma ainda existir outra classificação para ontologias, aplicável apenas às duas últimas acima citadas:

- Ontologias de tarefas: descrevem tarefas de um domínio – como, processos, planos, metas, escalonamentos, etc. – com uma visão mais funcional, embora declarativa, de um domínio;

Em outra abordagem, (GOMES-PÉREZ; FERNÁNDEZ-LOPES; CORCHO, 2003) afirmam que a comunidade de ontologias distingue entre ontologias que são principalmente uma taxonomia, denominadas ontologias *lightweight*, de ontologias que modelam o domínio de uma maneira mais profunda fornecendo mais restrições sobre a semântica do domínio, denominadas ontologias *heavyweight*. Ontologias *lightweight* incluem conceitos, taxonomias dos conceitos, relacionamentos entre conceitos e propriedades que descrevem os conceitos. Ontologias *heavyweight* adicionam axiomas e restrições às ontologias *lightweight*.

Complementando a abordagem anterior, (USCHOLD; GRÜNINGER, 2004) afirmam que a forma de especificar o significado dos termos distingue as diferentes abordagens de ontologias. Na figura 6 é apresentado um espectro evolutivo ilustrando os diferentes tipos de representações que geralmente são considerados como ontologias na literatura. À esquerda do espectro, existem representações simplificadas chamadas de “ontologias leves”, contendo apenas termos com pouca ou nenhuma especificação de significado. À medida que se avança para o lado direito da figura, aumenta-se o grau de formalização e a especificação do significado torna-se mais precisa,

reduzindo-se a ambigüidade. No extremo direito do espectro, encontram-se as teorias rigorosamente formalizadas que compõem as ontologias “autênticas”.

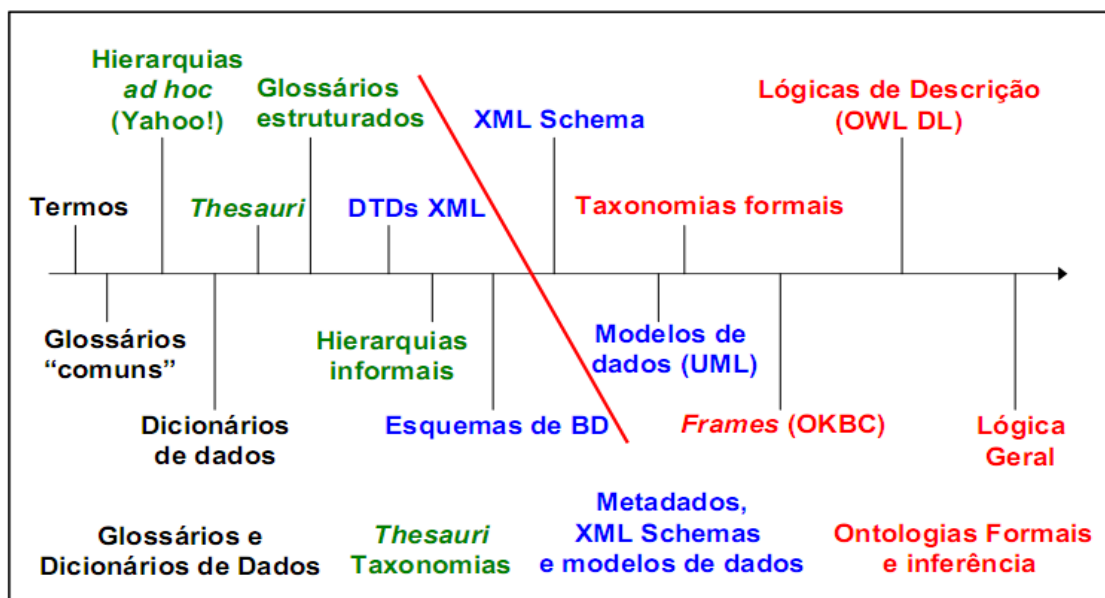


Figura 6 - Tipos de Representação de Ontologias (USCHOLD; GRÜNINGER, 2004)

Entre os tipos de representação de ontologia, as lógicas de descrição, apresentadas por (NARDI; BRACHMAN, 2003), são consideradas uma das famílias mais importantes para representação formal do conhecimento, e constituem a base para linguagens de representação como, por exemplo, a *Ontology Web Language (OWL)* (SMITH; WELTY; MCGUINNESS, 2004) a qual é oficialmente recomendada pelo *World Wide Web Consortium (W3C)* para a criação de ontologias na Web Semântica. A linguagem OWL subdivide-se em três sub-linguagens ou dialetos que diferem entre si apresentando aumento gradual de expressividade e complexidade:

- OWL Lite: provê os elementos básicos para representar conceitos, relacionamentos e restrições simples de propriedades, sendo o dialeto menos expressivo e complexo. Em função de suas construções mais simples, destina-se a usuários que necessitam obter uma classificação hierárquica com restrições simples. Por sua simplicidade, facilitam o desenvolvimento de ferramentas e máquinas de inferência para manipulação de ontologias, facilitando, por exemplo, migrações mais rápidas a partir de tesouros e outras taxonomias;

- OWL DL: por sua correspondência com as lógicas de descrição, este dialeto provê maior expressividade que a OWL Lite, aliada à garantia de que todas as inferências ou conclusões sejam computáveis e decidíveis, ou seja, podem ser resolvidas por sistemas computacionais e serão finalizadas em um tempo finito. Sendo uma das linguagens mais utilizadas para a criação de ontologias na Web Semântica, destina-se a usuários que desejam máxima expressividade;
- OWL Full: provê expressividade máxima sem garantias computacionais. Desta forma, ainda não foram desenvolvidas máquinas de inferência capazes de tratar todas as construções possíveis deste dialeto.

(YAGUINUMA, 2007) afirma que, de forma geral, as linguagens baseadas nas lógicas de descrição são adequadas para representar ontologias, de forma expressiva, além de oferecerem suporte ao raciocínio lógico através de máquinas de inferência, possibilitando a verificação de consistência, a classificação correta das instâncias e a obtenção de novas informações a partir das especificações presentes nas ontologias. A título de exemplo, a figura 7 apresenta uma ontologia simples, contendo conceitos representados por elipses em branco, propriedades por setas preenchidas e valores de atributos por retângulos. A instanciação é feita por setas tracejadas e as instâncias possuem tom acinzentado. O relacionamento “parte\_de” possui a restrição do axioma de transitividade, permitindo inferir que as instâncias da Cidade fazem parte das instâncias do País:

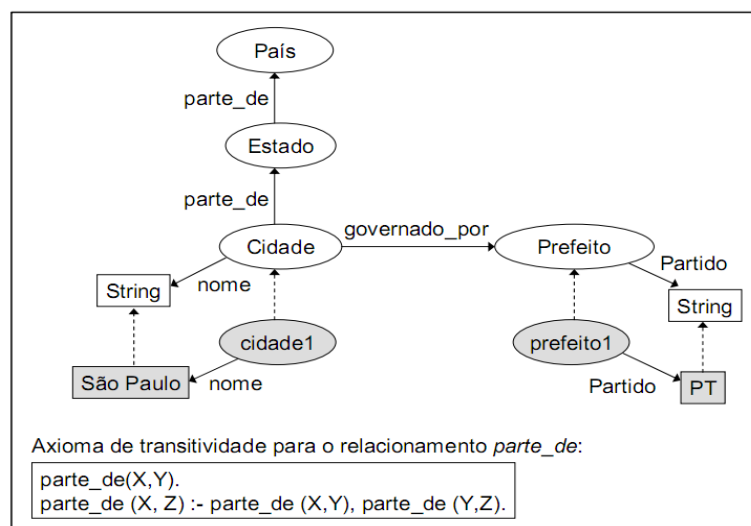


Figura 7 – Exemplo de uma Ontologia Simples (YAGUINUMA, 2007)

## 5.4 Engenharia de Ontologias

A engenharia de ontologias incorpora decisões sobre como representar uma ampla seleção de objetos e relações, em uma ordem lógica, levando a um modelo de nível ontológico. As abordagens para desenvolvimento de ontologias são normalmente específicas e limitadas. A ausência de atividades padronizadas, segundo (FERNANDEZ; GOMES-PEREZ; JURISTO, 1997), leva a desenvolvimentos conduzidos de forma artesanal e não como uma atividade científica. Os autores (NOY; McGUINNESS, 2001) também afirmam não existir um padrão definido para a criação de ontologias, mas sugerem sete passos a serem considerados durante sua criação:

1. Determinar o domínio e o escopo da ontologia;
2. Reutilizar ontologias existentes: considerar a existência de uma ontologia semelhante ou a possibilidade de refinar ou estender uma ontologia existente para uma tarefa específica. Muitas ontologias encontram-se disponíveis de forma eletrônica e podem ser importadas;
3. Levantar termos importantes a serem utilizados na ontologia: quais as propriedades dos termos? O que podemos dizer sobre estes termos?
4. Definir classes, sua hierarquia e especializações (subclasses);
5. Definir as propriedades (slots) das classes: uma classe de vinhos pode ter os seguintes slots: cor, sabor, quantidade de açúcar, ano da safra e produtor;
6. Restrições (facetadas) das propriedades: os slots podem possuir diferentes facetadas, como por exemplo, tipos de valores, valores permitidos, cardinalidade, etc. O nome de um vinho (ou *slot* nome) pode receber uma string de até vinte caracteres. Esta restrição é uma facetada do *slot* nome;
7. Criação de instâncias individuais para as classes na hierarquia: escolhe-se a classe, cria-se uma instância individual para a classe escolhida e preenche-se os valores dos slots;

Em outra abordagem para criação de ontologias, (USCHOLD; GRÜNINGER, 2004) afirmam que de forma geral, para modelar adequadamente uma conceitualização, ontologias são constituídas por conceitos, atributos, relacionamentos, axiomas formais e instâncias:

- Conceitos ou classes representam entidades que fazem parte de um domínio, tipicamente organizados em uma taxonomia de classes e subclasses;
- Atributos descrevem as características dos conceitos;



- Relacionamentos expressam como os conceitos estão interligados. Atributos e relacionamentos também podem ser chamados de propriedades dos conceitos;
- Axiomas são restrições sobre conceitos e propriedades, cujo principal propósito é expressar o significado de forma que as máquinas sejam capazes de interpretá-lo por meio de mecanismos de raciocínio automático.

#### **5.4.1 Método Unified Process for ONtology (UPON)**

Adicionalmente às metodologias citadas anteriormente, na literatura observa-se existir uma variedade de estratégias relacionadas ao desenvolvimento de ontologias, apresentando características diversas e direcionadas a diferentes propósitos e aplicações. Para (SILVA; SOUZA; ALMEIDA, 2008) a solução estaria centrada em uma proposta fundamentada em princípios teóricos e metodológicos que dessem sustentação científica ao processo de construção de ontologias. Em resposta a esta afirmativa e após constatar que entre os vários métodos para construção de ontologias, nenhum tenha se tornado uma clara referência, os autores (DE NICOLA; MISSIKOFF; NAVIGLI, 2009) apresentam o método Unified Process for ONtology (UPON), como uma nova abordagem para construção de ontologias, baseada em um processo bem estabelecido e amplamente utilizado em engenharia de software: o Processo Unificado de Desenvolvimento de Software (ou, simplesmente, Processo Unificado) aliado às técnicas de modelagem da Unified Modeling Language (UML). Trata-se de um método de cunho incremental para construção de ontologias que apresentem área de aplicação e objetivos bem definidos. O método engloba ciclos, fases, iterações e fluxos de trabalho. Cada ciclo consiste de quatro fases (concepção, elaboração, construção e transição) que eventualmente resultam em uma nova versão da ontologia. Cada fase é subdividida em iterações que, por sua vez, subdivide-se em cinco fluxos de trabalho: requerimentos, análise, projeto, implementação e testes. As fases iniciais estão mais concentradas no estabelecimento dos requerimentos da ontologia (identificação do domínio de interesse, escopo da ontologia, etc.) enquanto que as fases finais aprimoram a versão final da ontologia produzida no ciclo. Os principais objetivos do UPON são:

- Redução de tempo e custos na produção de ontologias em grande escala, fornecendo, também, orientações úteis para ontologias pequenas;
- Melhoria da qualidade da ontologia produzida, através da validação progressiva dos resultados intermediários;

- Criação de uma definição metodológica onde dois tipos de conhecimento, quais sejam, dos especialistas do domínio (DE) e dos engenheiros do conhecimento (KE), são explicitamente identificados e utilizados;
- Identificação clara das atividades, papéis e responsabilidades dos diferentes especialistas;
- Os resultados intermediários podem estar disponíveis para os usuários da ontologia por exemplo, para validação da busca semântica.

#### **5.4.2 Projeto Open Mind Common Sense (OCMS-Br)**

Devido à complexidade e abrangência demandadas pela ontologia Associações Perigosas de Analitos – APA, elaborada nesta pesquisa, o projeto Open Mind Common Sense - OMCS-Br (ANACLETO et al. - 2008) foi aplicado ao método UPON, viabilizando a existência de uma engenharia colaborativa e validativa que utiliza recursos de comunicação disponibilizados via web. Esta técnica inovativa de engenharia ontológica, adotada pela Abordagem Chem-Risk, incentivou fortemente a participação de diversos especialistas de domínio envolvidos no contexto da segurança alimentar química brasileira e mundial. A inclusão do Projeto OCMS-Br visou principalmente:

- Explorar e melhor conhecer a diversidade do vocabulário comumente utilizado por especialistas externos ao PNCRC / MAPA, particularmente, peritos agropecuários, consultores e analistas de risco;
- Validar e expandir o vocabulário representacional ou controlado utilizado no âmbito do PNCRC / MAPA;
- Validar e expandir o número de associações perigosas de analitos semanticamente modeladas na ontologia;
- Validar e aprimorar os novos processos de avaliação de riscos químicos sugeridos pela Abordagem Chem-Risk no âmbito das políticas públicas brasileiras que visam maximizar a segurança alimentar química, sob responsabilidade do MAPA.

Para facilitar a colaboração dos especialistas de domínio na coleta do senso comum, as seguintes medidas foram adotadas:

- O site OMCS-Br foi disponibilizado globalmente permitindo a participação de especialistas cadastrados;

- Os temas oferecidos auxiliaram a abordar e colher conhecimentos especializados que compõem o senso comum dos especialistas, os quais foram usados para aprimorar a ontologia.

Tecnicamente, a coleta do senso comum foi realizada por meio de modelos (templates) com frases e estruturas gramaticais simples associadas aos campos que devem ser preenchidos pelos especialistas. As informações já coletadas foram fornecidas aos especialistas, fazendo com que a base de conhecimento fosse constantemente retro-alimentada. A figura 8 ilustra, em alto nível de abstração, o esquema conceitual de operação do projeto OMCS-Br e sua interação com o domínio de conhecimento Associações Perigosas de Analitos – APA:

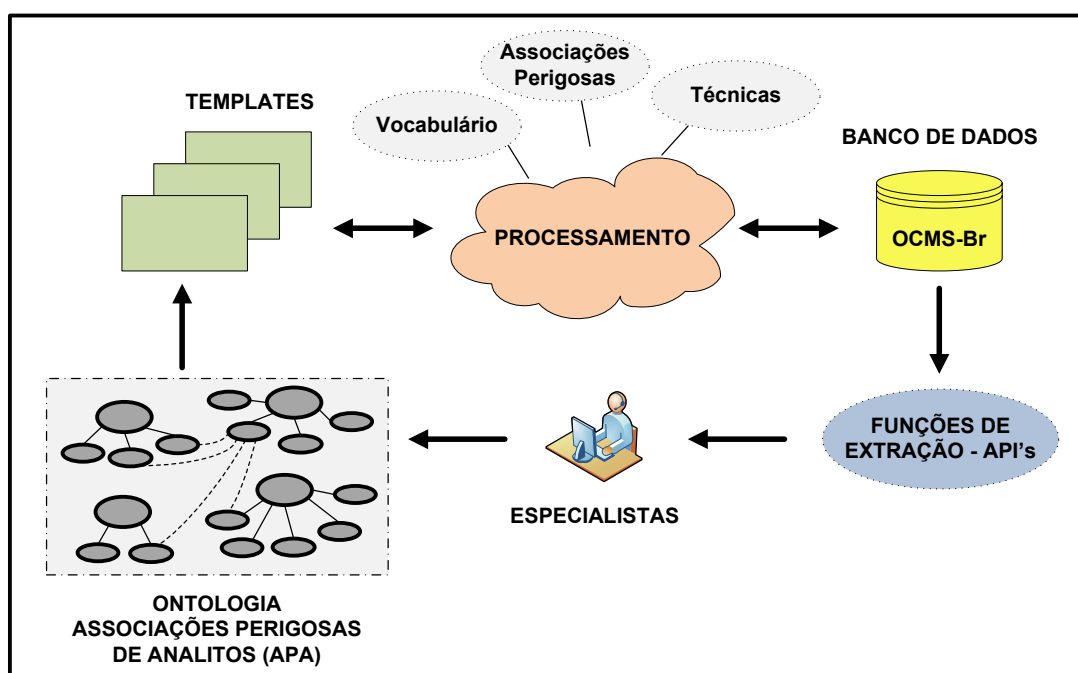


Figura 8: Esquema Conceitual de Operação do Projeto OMCS-Br e sua Interação com o Domínio de Conhecimento Associações Perigosas de Analitos - APA

## 5.5 Limitações Relevantes ao Contexto da Pesquisa

Entre as principais razões para o sucesso alcançado pela OWL, particularmente com relação ao dialeto OWL DL que permanece como uma das linguagens mais utilizadas para criação de ontologias na web semântica, cita-se o maior poder para expressividade dos conceitos,

objetos e relacionamentos de um domínio de conhecimento, aliado à garantia de que todas as inferências sobre este, sejam computáveis e decidíveis. Entre os principais desafios para aprimoramento da linguagem e ferramentas de apoio, encontra-se o aumento do poder de expressividade e desempenho, particularmente ao lidar com grandes ontologias e grandes massas de dados. Em atendimento a esta demanda, por todo o mundo pesquisadores buscam encontrar lógicas de descrição mais expressivas e sistemas de raciocínio (*reasoners*) otimizados, visando refinar e estender a capacidade da linguagem. Como resultado deste esforço conjunto, cita-se a habilidade para qualificar restrições de cardinalidade, qualificar propriedades como disjuntas, reflexivas e assimétricas, compor cadeias de propriedades, além de, entre outros, o suporte estendido para *datatypes* e *annotations*.

Entretanto, ao par destes inúmeros benefícios incorporados na forma de extensões que caracterizam a OWL2, (HORROCKS, 2008) afirma existirem situações em que mesmo as extensões da OWL2 não são adequadas ou mesmo suficientes para atender aos requisitos de algumas aplicações, conforme se verifica neste trabalho de pesquisa com relação à ontologia “Associações Perigosas de Analitos – APA”. As ferramentas de raciocínio OWL estão, sobretudo, relacionadas às classes e classificação, e, neste contexto, o raciocínio OWL é capaz de calcular todos os valores das propriedades que estejam implícitos às características das propriedades. Por esta razão, a título de exemplo, o raciocínio OWL não consegue afirmar que uma pessoa é o gerente de uma secretária, somente que a pessoa é um gerente. Segundo o autor, a expressividade da OWL pode ser aumentada por meio de implicações lógicas conhecidas como cláusulas de Horn, citando como exemplo:  $\text{parent}(x, y) \wedge \text{brother}(y, z) \Rightarrow \text{uncle}(x, z)$ . Por meio deste exemplo pode-se inferir que, se *y* é pai de *x* e *z* é irmão de *y*, conforme declarado na parte antecedente da regra, então, *z* é tio de *x*, conforme afirma a parte conseqüente da regra. Adotando esta linha de pesquisa (HORROCKS et al., 2004) apresentaram a Semantic Web Rule Language (SWRL), a qual propõe o uso de regras baseadas nas cláusulas de Horn, visando estender a expressividade da linguagem OWL de forma restrita e computável. Segundo os autores, a forma restrita de regras advém do fato de que a semântica das regras baseadas nas cláusulas de Horn aplica-se somente a instâncias ou indivíduos nomeados em uma ontologia. Esta é a razão pela qual a forma restrita de regras, conhecida como “*DL-Safe*”, não prejudica a decidibilidade do dialeto OWL-DL.

## 5.6 O Ambiente da Semantic Web Rule Language (SWRL)

A partir do esforço desenvolvido por pesquisadores ao redor do mundo em busca de uma padronização para regras de inferência no contexto da World Wide Web, a linguagem de marcação RuleML surge como resposta à publicação e compartilhamento de regras neste ambiente. Entre as características da linguagem RuleML destaca-se a possibilidade de construção de sub-linguagens de regras sobre as linguagens XML, RDF e OWL, gerando, assim, a SWRL. Em suma, a linguagem de regras para web semântica consiste em uma proposta para combinar ontologias e regras de inferência (OWL-DL + RuleML = SWRL) que incorpora uma sintaxe abstrata de alto-nível para regras baseadas nas cláusulas de Horn (Horn-like), formando um ambiente que restringe a expressão de todas as regras em termos dos conceitos OWL, quais sejam: classes, propriedades e indivíduos.

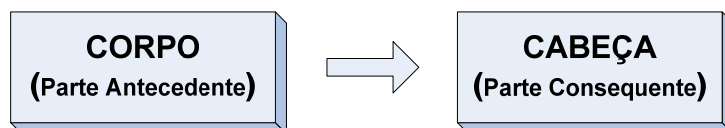
Entre as principais motivações para adoção de uma linguagem de regras ao nível da ontologia empregada no contexto desta pesquisa, cita-se:

- Aumento do poder de expressividade da linguagem OWL;
- Integração entre os raciocínios SWRL e OWL;
- Inúmeras ferramentas de suporte, tais como, Bossam, R2ML, Hoolet, Pellet, KAON2, RacerPro e SWRLTab (Protege);
- As regras são consideradas instâncias de dados e armazenadas como parte da ontologia;
- Reusabilidade do conjunto de regras;
- Maior facilidade para leitura e escrita de regras utilizando uma linguagem de regras verdadeira;
- Utilizar o recurso de que uma regra pode ser verdadeira em algumas situações, mas não em todos os casos.

Com relação à capacidade de trabalhar com *reasoners*, citada no segundo item das principais motivações, ressalta-se que a capacidade de inferência da linguagem OWL é preponderantemente baseada nas características das propriedades das classes, tais como, inversão, simetria e transitividade, ao passo que a capacidade de inferência da linguagem SWRL baseia-se no conjunto de regras SWRL. Diferentemente do raciocínio OWL, o raciocínio SWRL ocorre ao nível das propriedades e instâncias (objetos), forçando a existência de uma interação funcional

entre os raciocínios. Com relação ao terceiro item, deve ser ressaltado que a SWRLTab fornece um ambiente completo para criação, edição e eliminação de regras SWRL, que atua de forma complementar ao ambiente do Protegé-OWL.

O formato das regras utilizado na linguagem SWRL é ilustrado na figura 9:



**Figura 9 – Formato das Regras Utilizado na Linguagem SWRL**

Regras SWRL possuem a forma de implicação entre a parte antecedente e a consequente. Ambas consistem de zero ou múltiplos átomos. Múltiplos átomos são tratados como conjunções, não sendo permitida a existência de disjunções entre os mesmos. O significado deste formato pode ser entendido como: se a parte antecedente acontece (é verdadeira) então a parte consequente também irá acontecer.

Entre as principais limitações da linguagem SWRL, destacam-se:

- Suporta apenas inferência monótona, ou seja, as regras SWRL não podem ser usadas para modificar informações existentes em uma ontologia. Caso as regras SWRL permitissem a não monotonicidade as modificações poderiam acontecer, inclusive possibilitando a eliminação de informações na ontologia;
- A capacidade de ligar-se a indivíduos que não são conhecidos (nomeados) em uma ontologia é conhecida causa de indecidibilidade. Por esta razão, conforme anteriormente citado, as regras SWRL são do tipo DL-Safe, ou seja, em virtude da desejável propriedade de decidibilidade existe uma restrição para que as regras SWRL associem-se somente com indivíduos nomeados em uma ontologia, formando, assim, um subconjunto restrito de regras que garante a decidibilidade.

Adicionalmente, o ambiente de regras SWRL permite a incorporação da linguagem de consulta SQWRL-Semantic Query-Enhanced Web Rule Language (O'CONNOR; DAS, 2009), a qual, baseada na linguagem SWRL, destina-se a promover consultas às ontologias OWL por meio de operadores baseados na linguagem SQL. Por meio deste mecanismo, as consultas SQWRL podem operar em conjunção com a linguagem SWRL e, assim, serem usadas para

recuperar o conhecimento oferecido pelas ontologias, inferido, neste caso particular, por meio de regras SWRL. Para ser capaz de executar as consultas às ontologias, utilizando regras SWRL, a linguagem SQWRL requer uma máquina de regras, como por exemplo, a máquina de regras Jess utilizada no contexto desta pesquisa. A partir deste ponto, os resultados das consultas podem ser acessados internamente no ambiente do Protege-OWL, por meio do mecanismo SQWRLTab, ou a partir de um ambiente externo de programação por meio da SQWRLQueryAPI, que fornece uma interface baseada em JDBC para execução e recuperação de consultas a partir de aplicações Java – A preparação, execução e acesso aos resultados das consultas SQWRL utilizando a API Java, foi o método de trabalho escolhido para recuperar o conhecimento oferecido pela ontologia Associação Perigosa de Analitos – APA. Em ontologias que apresentam comportamento mais dinâmico, permitindo constantes alterações em sua estrutura, a linguagem SQWRL é indicada por ser genérica e apropriada à execução de consultas no contexto da web semântica expressa em OWL, semelhantemente à linguagem SQL e os bancos de dados relacionais. Em muitos casos, a partir de modificações na estrutura da ontologia, faz-se uma única modificação na consulta SQWRL.

## **5.7 Considerações Finais**

Yaguinuma (YAGUINUMA, 2007) afirma que na ciência da computação, ontologias podem ser aplicadas nas mais diversas formas. Em geral, são utilizadas em contextos de organização, reuso e compartilhamento de informação semântica, facilitando a compreensão e a comunicação do conhecimento de um domínio entre humanos e sistemas computacionais. É neste contexto de organização, compartilhamento e reuso da informação semântica, que as ontologias desempenham um papel fundamental em diversas áreas de atuação, seja na esfera privada, pública ou na governamental, entre as quais, situa-se o Governo Federal Brasileiro.

O Capítulo 7 desta dissertação apresenta o uso de ontologias para capturar o conhecimento inerente ao domínio “Associações Perigosas de Analitos - APA”, modelando-o semanticamente por meio da descrição dos conceitos e relacionamentos mantidos entre estes. Conforme sugerido nesta pesquisa, este domínio de conhecimento representa um fator de risco adicional que incide ao nível das amostras das matrizes agropecuárias monitoradas pelo PNCRC. Por esta razão, deve ser incorporado aos procedimentos de análise de riscos químicos associados ao processo de produção agropecuária brasileiro, gerenciado pelo Ministério da Agricultura, Pecuária e Abastecimento –

MAPA, podendo, em maior nível de abstração, ser incorporado ao processo maior de produção agropecuária mundial, gerenciado pela Comissão do *Codex Alimentarius* – OMS/ONU.

O próximo capítulo discute o Processo de Descoberta de Conhecimento em Bases de Dados – DCBD, utilizado como linha mestra para a criação da plataforma base flexível utilizada na abordagem Chem-Risk.



# Capítulo 6

## Descoberta de Conhecimento em Bases de Dados - DCBD

---

### 6.1 Considerações Iniciais

O problema da explosão de dados, atualmente vivenciado, tem sido largamente citado por inúmeros autores na ciência da computação (CHAKRABARTI et al, 2009.), (TAN; STEINBACH; KUMAR, 2009), (WITTEN; FRANK, 2005). Acredita-se que esse seja decorrente de um processo maior de informatização generalizada aliada aos rápidos avanços tecnológicos. É possível deduzir que somos ricos em dados, mas muito pobres em informação (HAN; KAMBER, 2006) e, conseqüentemente, no conhecimento superior que pode ser proporcionado por esta. Foi com base neste contexto, largamente observado nos dias de hoje, que, segundo (FAYYAD, 1996), teorias e ferramentas computacionais surgiram, visando auxiliar a extração de conhecimento a partir de grandes massas de dados, dando origem a área denominada Descoberta de Conhecimento em Bases de Dados - DCBD. Neste projeto de pesquisa, considera-se o dado como algo bruto ou a matéria prima da qual pode-se extrair informação; informação como sendo o dado processado apresentando significado e contexto bem definido, e, conhecimento, o uso inteligente da informação.

### 6.2 O processo DCBD

A Descoberta de Conhecimento em Bases de Dados (DCBD) é compreendida como a integração entre as áreas de estatística, inteligência artificial (redes neurais e sistemas simbólicos) e banco de dados (PRADO, 2001). A computação encontra-se diretamente aplicada em todas as etapas do processo DCBD, participando não somente na execução de algoritmos para identificação de padrões e tendências, mas auxiliando o desenvolvimento do conjunto de métodos, técnicas e ferramentas resultantes. Encontram-se no âmbito da DCBD as atividades de coleta, limpeza, redução de dados e seleção de parâmetros para a execução de algoritmos, visando a identificação de padrões ou tendências, além da representação e processamento do conhecimento. Mineração de Dados é a principal etapa do processo de DCBD, consistindo na aplicação de algoritmos para a

geração de padrões sobre um conjunto de dados. As figuras 10 e 11 ilustram o processo DCBD em diferentes níveis de abstração:

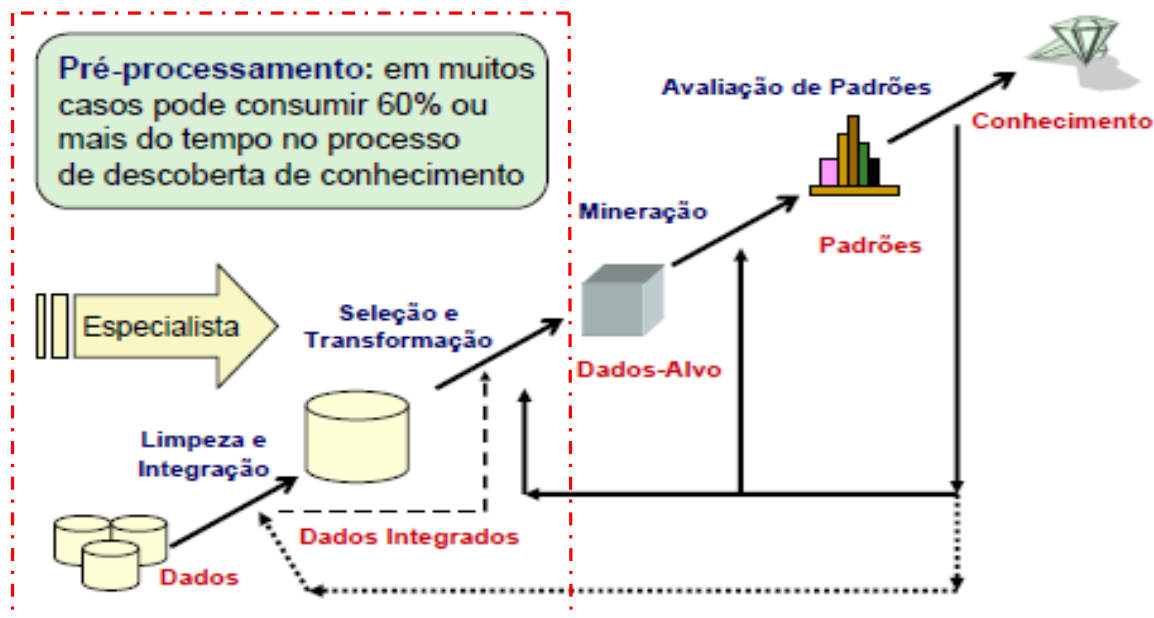


Figura 10 - Atividades do Processo de Descoberta de Conhecimento em Bases de Dados

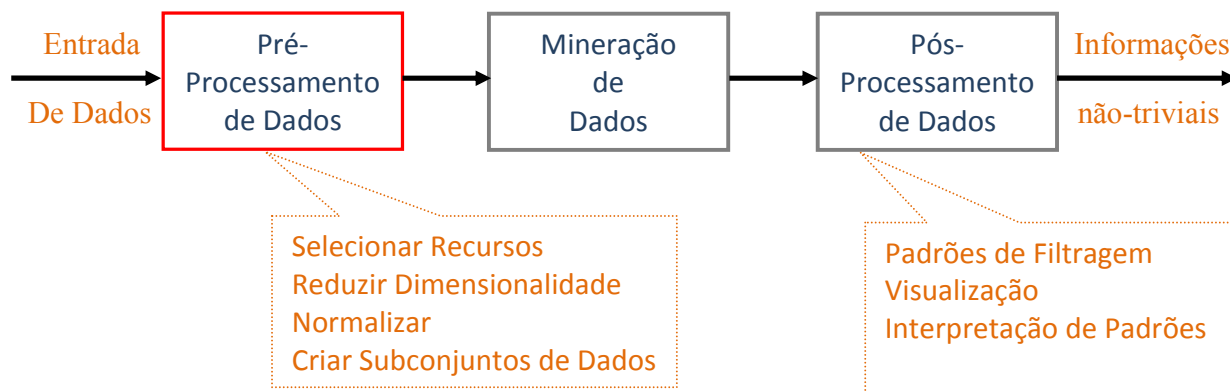


Figura 11 - O Processo de Descoberta de Conhecimento em Bases de Dados baseado em (TAN; STEINBACH; KUMAR, 2009)

O presente trabalho abrange as atividades de preparação ou pré-processamento de dados, mineração e visualização de dados em um processo DCBD.

### **6.3 Pré-Processamento de Dados**

A preparação dos dados tem merecido a atenção de inúmeros autores na literatura da ciência da computação, tais como, (CHAKRABARTI et al., 2009.), (TAN; STEINBACH; KUMAR, 2009), (HAN; KAMBER, 2006), (WITTEN; FRANK, 2005), (BRAGA, 2005), entre outros. (HAN; KAMBER, 2006) afirmam que “uma baixa qualidade dos dados irá levar a resultados de mineração de baixa qualidade”, relacionando uma série de técnicas de pré-processamento capazes de melhorar a eficácia do processo de mineração, em termos de qualidade dos padrões extraídos ou com relação ao tempo necessário para desenvolvê-la, quais sejam:

- Limpeza dos dados: Destina-se a remoção de ruídos e inconsistências;
- Integração de dados: Combinação de diferentes fontes ou repositórios de dados;
- Seleção de dados: Selecionar um conjunto de dados relevantes às tarefas de mineração;
- Transformação de dados: Transformar, consolidar ou normalizar os dados em formas apropriadas às tarefas de mineração, por meio de sumarização, agregação, etc.;
- Redução de dados: Obter uma representação resumida em volume, mas que produza resultados analíticos similares.

Parafraseando Jiawei Han (HAN; KAMBER, 2006), a abordagem Chem-Risk, ao adicionar qualidade aos dados, considerados nesta pesquisa como resultados das avaliações de risco químico agropecuário, adiciona qualidade e precisão aos resultados obtidos na atividade de mineração de dados do processo DCBD.

### **6.4 Mineração de Dados**

A mineração de dados é considerada a principal atividade do processo DCBD, possibilitando realizar uma extração de conhecimento, na forma de padrões não triviais, implícitos e desconhecidos, a partir de grandes massas de dados (WITTEN; FRANK, 2005). Trata-se de uma atividade de natureza exploratória que pode ser realizada de maneira descritiva, ou seja, descrevendo o conjunto de dados analisado de forma resumida, informativa e discriminante, ou de maneira preditiva, por meio de modelos utilizados para previsão de tendências e propriedades de dados desconhecidos. Estes modelos podem ser do tipo supervisionado, os quais exigem a intervenção de um especialista para funcionamento adequado, ou do tipo não supervisionado, os quais prescindem desta intervenção. Entretanto, segundo

(TAN; STEINBACH; KUMAR, 2009), embora a mineração de dados traduza um processo de descoberta de informações úteis em grandes depósitos de dados, nem todas as tarefas de descoberta de informação são consideradas mineração de dados. Por exemplo, tarefas baseadas nas técnicas tradicionais da ciência da computação para organizar e recuperar a informação por meio de estruturas de índices em um sistema gerenciador de banco de dados ou encontrar páginas web consultando um mecanismo de busca na internet, são tarefas relacionadas à área de recuperação de dados. Ainda de acordo com estes autores, as tarefas de mineração de dados são divididas em duas categorias principais:

- **Tarefas de Previsão:** Objetivam prever o valor de determinado atributo baseado nos valores de outros atributos. O atributo a ser previsto é comumente conhecido como variável dependente ou alvo, enquanto que os atributos usados para fazer a previsão, são conhecidos como variáveis independentes ou explicativas. Existem dois tipos de tarefas de modelagem de previsão: **classificação e regressão**;
- **Tarefas Descritivas:** Objetivam derivar padrões de conhecimento, quais sejam, correlações, tendências, agrupamentos, trajetórias ou anomalias, que resumam os relacionamentos subjacentes dos dados. Em mineração de dados são consideradas de natureza exploratória, exigindo frequentemente técnicas de pós-processamento para validar e explicar os resultados. Existem dois tipos de tarefas descritivas: **análise de associação**, usada para descobrir padrões que descrevam características altamente associadas dentro do conjunto de dados, e **análise de grupo**, usada para encontrar grupos de observações intimamente relacionadas, de modo que as observações pertencentes a um mesmo grupo sejam semelhantes entre si, comparativamente às pertencentes a outros grupos;

A figura 12 ilustra as principais tarefas que podem ser executadas durante a atividade de mineração de dados do processo DCBD, em busca de padrões não triviais de conhecimento a serem revelados para o conjunto de dados analisado:

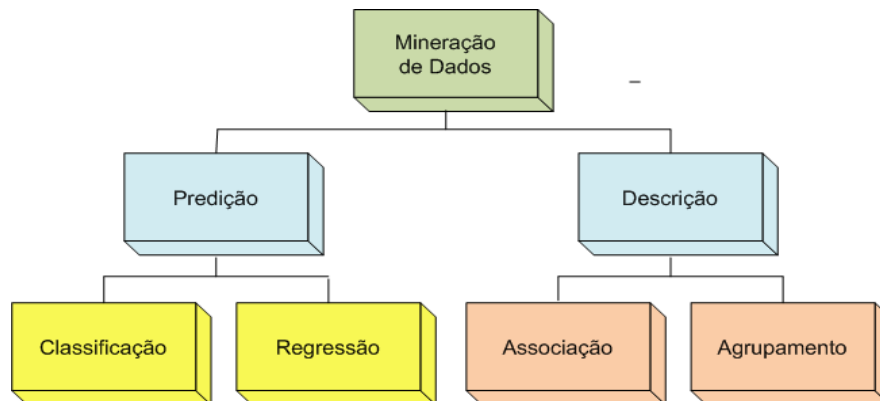


Figura 12 - Principais Tarefas de Mineração de Dados

## 6.5 Pós-Processamento de Dados

Segundo (TAN; STEINBACH; KUMAR, 2009), “fechar o laço” é uma expressão frequentemente usada para se referir ao processo de integrar os resultados da atividade de mineração de dados com os sistemas de apoio a decisão. Este processo de integração pode requerer uma etapa de pós-processamento que assegure, por exemplo:

- A incorporação somente de resultados válidos e úteis ao sistema de apoio a decisão. Medições estatísticas e métodos para testes de hipóteses, por exemplo, podem ser aplicados durante o pós-processamento para eliminar resultados não legítimos encontrados durante a atividade de mineração de dados;
- Uma representação adequada do conhecimento descoberto, tornando-o mais facilmente acessível à compreensão dos usuários, que neste estudo de caso, representam os analistas de risco, peritos agropecuários e autoridades do Ministério da Agricultura, Pecuária e Abastecimento - MAPA. Técnicas de visualização podem ser empregadas permitindo que os analistas possam explorar os resultados da mineração, a partir de uma diversidade de pontos de vista. Ao exibir a informação na forma gráfica ou tabular, a visualização objetiva contribuir para que uma pessoa possa interpretar visualmente a informação, facilitando a assimilação do conhecimento ao formar um “modelo mental” das informações produzidas (TAN; STEINBACH; KUMAR, 2009). O uso das técnicas de visualização é conhecido como “mineração visual dos dados” e normalmente desempenham um papel chave, facilitando a análise das informações produzidas e a assimilação do conhecimento descoberto.

## **6.6 Limitações Relevantes ao Contexto da Pesquisa**

Embora seja de suma importância para a geração de conhecimento por meio de tecnologias avançadas, uma limitação considerável do processo completo de DCBD refere-se ao elevado grau de subjetividade que normalmente caracteriza um processo não totalmente automatizado. Pode-se afirmar que, neste processo, a atividade que apresenta maior nível de automatização é a de mineração de dados, embora também existam ferramentas de suporte para as demais atividades. Acredita-se que a principal razão, esteja em função da diversidade e da complexidade das técnicas que podem ser empregadas ao longo das fases de pré-processamento, mineração de dados e pós-processamento.

## **6.7 Considerações Finais**

Como veremos, a partir do capítulo 7 desta dissertação, as técnicas de pré-processamento, mineração e pós-processamento de dados que compõem o processo de descoberta de conhecimento, desempenharam um papel fundamental para que a abordagem Chem-Risk pudesse alcançar o objetivo de fornecer maiores níveis de conhecimento não-trivial, relacionado ao domínio maior Análise de Risco Químico Agropecuário, para analistas de risco, consultores agropecuários e autoridades do governo que comandam o Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC. As técnicas de mineração de dados, em particular, permitiram refinar e estender o conhecimento superior oferecido na atividade de pré-processamento de dados por meio de ontologias e lógica *fuzzy*, detectando a existência de novas associações perigosas de analitos, sugerindo novas faixas de risco na forma de agrupamentos para analitos e amostras, e permitindo realizar uma comprovação formal dos resultados superiores alcançados pela abordagem Chem-Risk, comparativamente às abordagens Tradicional e Vetor de Risco Ponderado, empregadas no contexto das políticas públicas brasileiras que visam maximizar a segurança alimentar química.

Os resultados alcançados na etapa de mineração de dados da Abordagem Chem-Risk complementam a realidade da segurança química dos alimentos agropecuários brasileiros e expandem o nível de conhecimento especializado em direção à tomada de decisões não triviais.

Os procedimentos para avaliação de riscos químicos empregados por estas abordagens serão analisados ao longo do Capítulo 7.

## **Parte III – Trabalho de Pesquisa**



# Capítulo 7

## Abordagem Chem-Risk

---

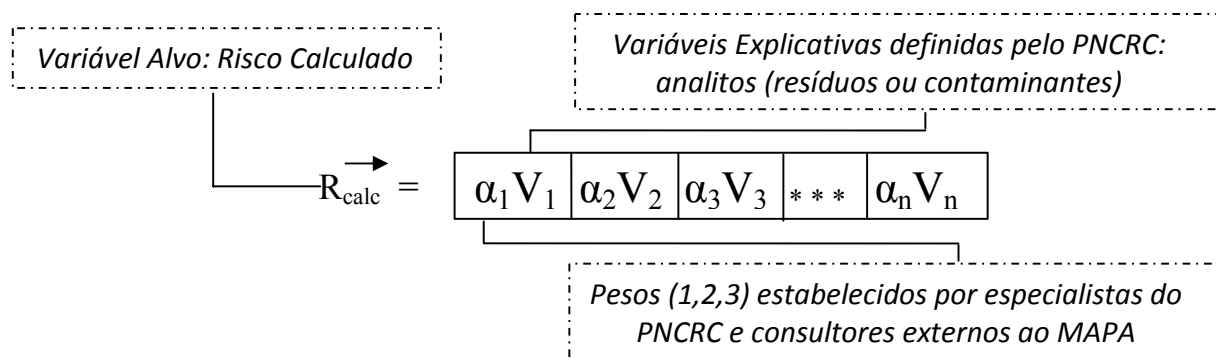
### 7.1 Considerações Iniciais

Conforme apresentado no Capítulo 3, diferentes tipos e técnicas de análises laboratoriais são empregados por especialistas do Ministério da Agricultura para avaliação dos riscos inerentes a aproximadamente 430 analitos que são passíveis de serem encontrados em amostras de matrizes de produtos agropecuários. Tais amostras são sistematicamente colhidas ao longo do país e analisadas nos laboratórios credenciados pelo Plano Nacional de Controle de Resíduos e Contaminantes- PNCRC. Os resultados dessas análises laboratoriais são publicados em diferentes formatos de dados, de acordo com o tipo da substância analisada. Entre os dados considerados nessas análises, cita-se além da concentração do analito, a frequência recomendada para a análise laboratorial (anual, semestral, trimestral, mensal, semanal, outra), a localidade da amostra (região, estado, município, outra) e o proprietário da amostra (cooperativa, instituição, laticínio, outro). Na abordagem que tem sido utilizada para avaliação de riscos químicos no contexto das políticas públicas brasileiras, os especialistas analisam esses resultados por meio de diferentes métodos e técnicas para aferição dos níveis de riscos químicos. Normalmente a análise é efetuada considerando um conjunto de amostras. A abordagem atualmente praticada no âmbito do PNCRC / MAPA, no contexto do presente trabalho, será referenciada como Abordagem Tradicional.

Entre as inúmeras dificuldades enfrentadas pelo PNCRC para gerenciamento de processo de avaliação de riscos químicos nacional, destacam-se o alto custo das avaliações de risco e a rotatividade de profissionais envolvidos, particularmente, com relação aos analistas de risco, estatísticos, peritos agropecuários e consultores externos. O processo maior das avaliações de risco incorrem em alto custo em função de fatores como a alta tecnologia empregada para análise das concentrações dos analitos, o tempo requerido por estas análises, os salários, viagens dos especialistas e consultores, o envio de amostras e a execução das políticas de intervenção e ações regulatórias ao processo produtivo dos alimentos agropecuários brasileiros.

A ação de avaliação e controle de riscos atualmente desenvolvidos pelo PNCRC em nível nacional, levaram o País ao reconhecimento e respeito internacional em fóruns que definem as diretrizes para o comércio de alimentos agropecuários, observando as questões de segurança alimentar. No entanto, é possível observar que essa ação pode ser bastante melhorada se houver diminuição da subjetividade e erros decorrentes da avaliação humana, além do estabelecimento de uma padronização efetiva de procedimentos.

Uma abordagem alternativa para avaliação de riscos, sugerida por assessores do MAPA, enfoca um modelo estatístico aplicado à determinação do risco calculado ( $R_{calc}$ ) associado a uma amostra de matriz agropecuária. Conforme ilustrado na figura 13, trata-se de um vetor de risco destinado a promover uma classificação ponderada de amostras, em função dos diversos analitos (resíduos e contaminantes) detectados, que apresenta a capacidade de ser gradualmente estendido até ao nível do processo de produção nacional do alimento analisado. No contexto deste trabalho, essa abordagem será referida como abordagem Vetor de Risco Ponderado.



**Figura 13: Abordagem Vetor de Risco Ponderado Sugerida ao MAPA**

Apesar do esquema de ponderação estar aplicado ao nível dos analitos, o nível de risco é aferido para uma única amostra de matriz agropecuária. Dessa forma, a amostra constitui a menor granularidade de informação trabalhada por este modelo, o que denota um avanço técnico em comparação à abordagem Tradicional de análise de risco adotada pelo MAPA, cujas granularidades mínimas são normalmente os grupos de amostras.

Entretanto, embora os benefícios proporcionados por este modelo sejam consideráveis, sua estrutura de cálculo também acarreta perda significativa de valiosas informações, em virtude da produção agropecuária mundial estar associada a um complexo de variáveis relacionadas às

questões ambientais, sócio-político-econômicas, químicas, biológicas e físicas (BONNET; SILVA, 2007).

Por esta razão, torna-se imprescindível a existência de um modelo que seja capaz de agregar novos fatores de risco e técnicas de análise de risco a fim de que os níveis de risco inferidos apresentem maior abrangência e precisão em seu cálculo, e, assim, consigam expressar eficazmente, em maior ou menor granularidade, a real qualidade dos produtos agropecuários consumidos ao longo das regiões do país ou exportados. Na próxima seção são explicitadas as macro-etapas que compõem a abordagem Chem-Risk.

## 7.2 Visão Geral da Abordagem Chem-Risk

A visão geral da Abordagem Chem-Risk é apresentada na figura 14. A abordagem foi dividida em Macro-Etapas que correspondem aos principais módulos ou procedimentos computacionais realizados para estimativa de riscos químicos associados ao processo de produção dos produtos agropecuários brasileiros.

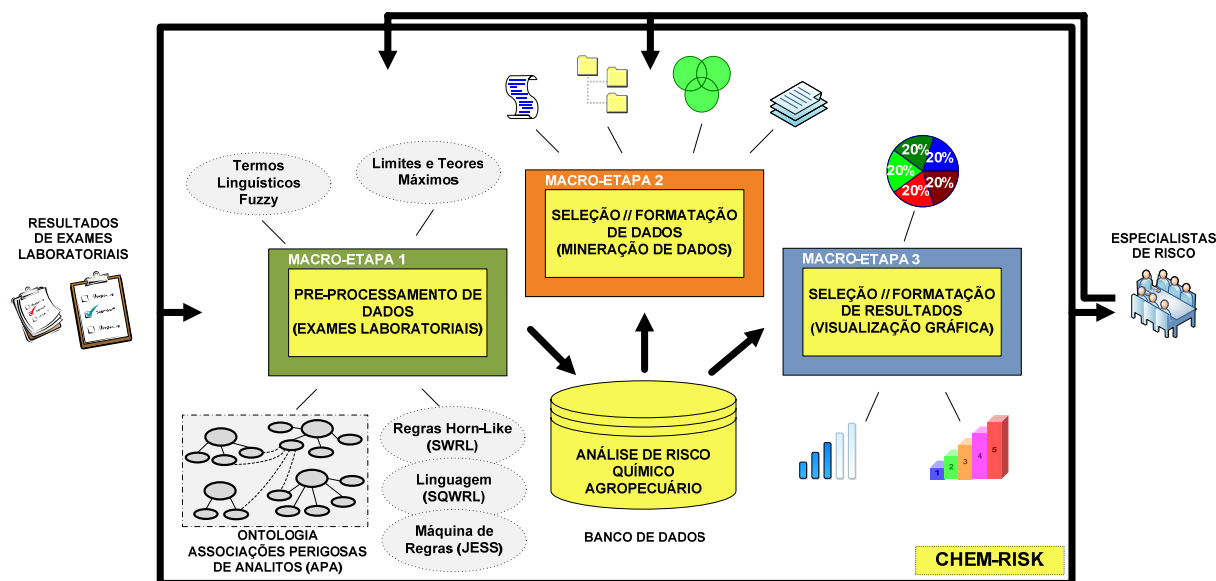


Figura 14: Visão Geral da Abordagem Chem-Risk

Conforme ilustrado na figura 14, na Macro-Etapa 1 a Abordagem Chem-Risk realiza o pré-processamento dos resultados dos exames laboratoriais realizados nas matrizes dos produtos

agropecuários. Em um primeiro momento, utiliza os conceitos da lógica *fuzzy* para estimativa dos níveis de risco associados aos analitos detectados na amostra, de acordo com os termos linguísticos pré-estabelecidos conjuntamente com os especialistas do PNCRC, e com os limites e teores máximos pré-estabelecidos mundialmente pelo *Codex Alimentarius* para as concentrações dos analitos. Em um segundo momento, os níveis de risco da amostra são estimados, levando-se em consideração os termos linguísticos pré-estabelecidos para amostras e a ontologia Associações Perigosas de Analitos - APA. Todos os resultados aferidos na Macro-Etapa 1 são armazenados no banco de dados Análise de Risco Químico Agropecuário – ARQA, de forma que na Macro-Etapa 2, estes resultados possam ser formatados adequadamente para submissão aos processos de mineração de dados. Na macro-Etapa 3, a fim de que todos estes resultados gerados nas macro-etapas anteriores possam ser melhor compreendidos e interpretados pelos especialistas do PNCRC, a abordagem Chem-Risk permite que sejam visualizados graficamente. As próximas subseções detalham as macro-etapas da Abordagem Chem-Risk:

### **7.2.1 Macro-Etapa 1 - Pré-processamento de Dados**

Conforme mencionado anteriormente, na primeira macro-etapa, a Abordagem Chem-Risk realiza o pré-processamento dos dados gerados a partir da massa de exames laboratoriais realizadas em amostras de matrizes de produtos agropecuários pelos laboratórios credenciados do PNCRC / MAPA. Os resultados dos exames laboratoriais são publicados individualmente na forma de Concentração de Analito (CA) para cada analito detectado na amostra analisada, respeitando o conjunto de analitos pré-estabelecido mundialmente pela Comissão do *Codex Alimentarius* de acordo com a matriz agropecuária.

#### **Plataforma Base Flexível de Pré-processamento**

A Plataforma Base Flexível gerencia todos os procedimentos computacionais realizados na Macro-Etapa 1 da abordagem Chem-Risk para pré-processamento das avaliações de riscos químicos ao nível dos analitos, amostras e lotes de amostras agropecuárias. Adicionalmente ao fator de risco Concentração do Analito (CA), igualmente considerado nas Abordagens Tradicional e Vetor de Risco Ponderado, a Abordagem Chem-Risk cria uma Plataforma Base Flexível para permitir a incorporação de diversos outros fatores de risco ao cálculo dos níveis de

risco associados aos analitos, amostras e lotes de amostras. Os fatores de risco adicionais à Concentração do Analito (CA) que incidem ao nível dos analitos são em número de três:

- Tendência da Concentração do Analito (TDCA);
- Prazo de Ajustamento da Concentração do Analito (PZCA) aos índices mundiais de segurança química alimentar;
- Custo de Ajustamento da Concentração do Analito (CTCA) aos índices mundiais.

Ao nível das amostras e lotes de amostras, os fatores de risco adicionais individualmente considerados, são em número de seis:

- Local de Colheita da Amostra (LCA);
- Nível de Toxicidade dos Analitos não conformes (NTA);
- Associações Perigosas de Analitos (APA) agregando os fatores individuais:
  - Efeito Toxicológico Aditivo – ETA;
  - Efeito Toxicológico Sinérgico – ETS;
  - Efeito Toxicológico de Potenciação – ETP;
  - Efeito Toxicológico Redutor – ETR.

Tecnicamente, à luz das ciências administrativas, a Plataforma Base Flexível realiza integralmente os procedimentos recomendados pelo Processo de Gerência de Riscos (ALENCAR; SCHMITZ, 2006), citado no Capítulo 2, sendo considerada uma variante da Técnica de Kepner - Tregoe (KEPNER; TREGOE, 1991) aplicada para identificação, análise, organização e solução de problemas. À luz da ciência da computação, inspirada na lógica do processo de Descoberta de Conhecimento em Bases de Dados (DCBD), a plataforma incorpora técnicas computacionais inerentes à lógica *fuzzy*, visando pré-processar adequadamente os dados (resultados das avaliações de risco em analitos, amostras e lotes de amostras) a serem submetidos aos processos de mineração de dados, para que estejam mais exatos, precisos e fidedignos à realidade da segurança química alimentar mundial. Adicionalmente, incorpora técnicas computacionais inerentes às ontologias, criando a ontologia Associações Perigosas de Analitos – APA para viabilizar a existência de análises de risco complexas envolvendo os fatores de risco adicionais, ao nível dos analitos e amostras, tornando o dado ainda mais completo, fidedigno e tecnicamente confiável. A incorporação dessas técnicas computacionais à Plataforma Base Flexível será separadamente discutida nos itens 8.3 e 8.4, respectivamente. A figura 15 apresenta um diagrama enfocando as principais etapas que compõem a Macro-Etapa 1:

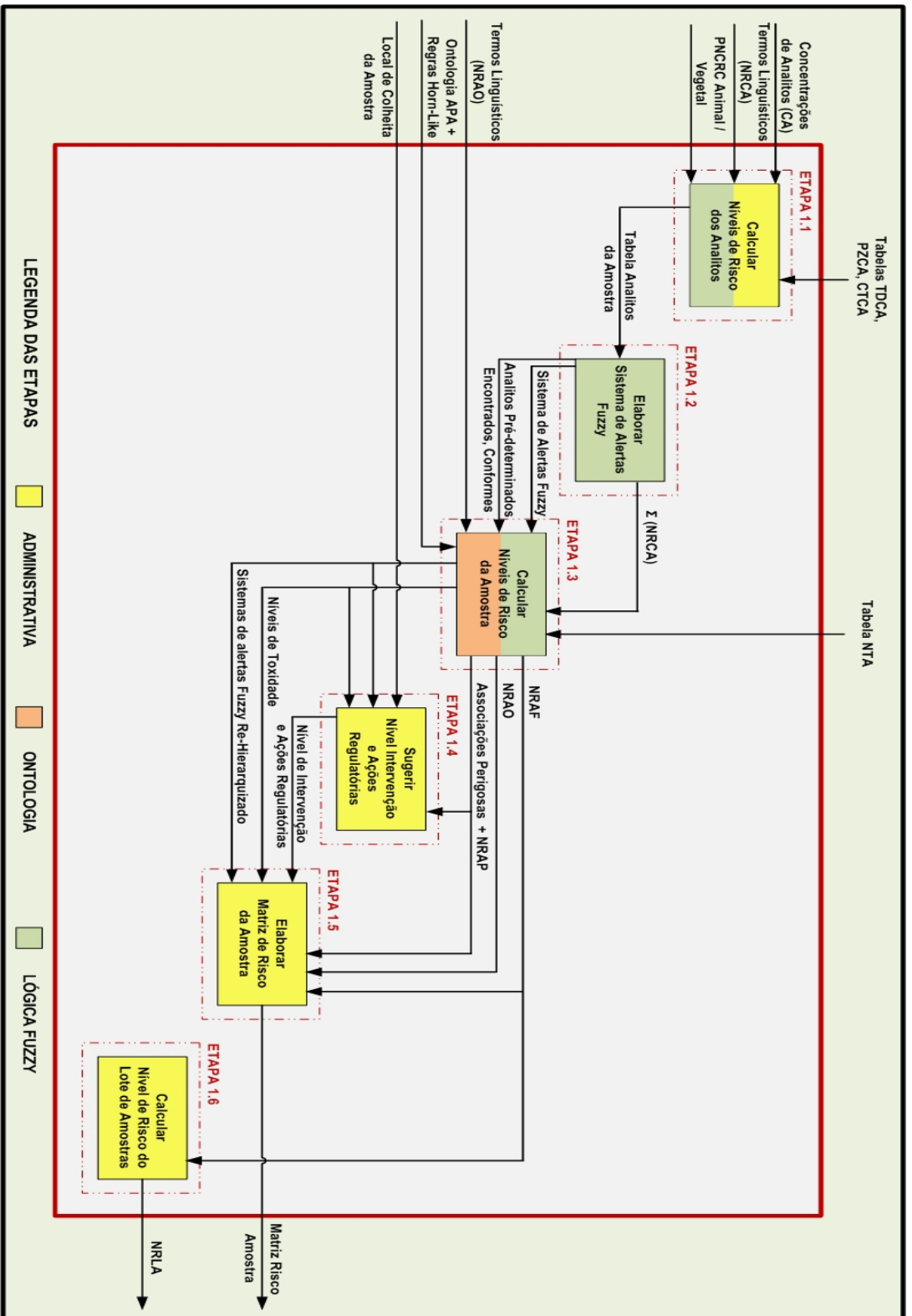


Figura 15 – Principais Etapas que Compõem a Plataforma Base Flexível - Macro-Etapa I da Abordagem Chem-Risk

Detalha-se, a seguir, as principais etapas que compõem a Plataforma Base Flexível ilustrada na figura 15, tomando como referência os 43 analitos pré-definidos como resíduos e contaminantes pelo PNCRC / Codex Alimentarius para análise de riscos químicos em uma amostra da matriz agropecuária leite. No exemplo apresentado a seguir, 31 analitos foram detectados na amostra de leite analisada, entre os 43 obrigatoriamente pesquisados.

**Etapa 1.1:** Nesta etapa são aferidos os níveis de risco associados aos analitos detectados na amostra de leite, particularmente com relação às concentrações que apresentam. A figura 16 ilustra a representação gráfica da Etapa 1.1 que engloba as Sub-Etapas 1.1.1 a 1.1.7. As cores amarelo e verde destacadas no processo, indicam, respectivamente, a existência de procedimentos administrativos e computacionais inerentes à lógica *fuzzy*. Visando facilitar a compreensão destes procedimentos e cálculos associados, serão considerados os resultados aferidos para 6 analitos entre os 31 detectados na amostra.

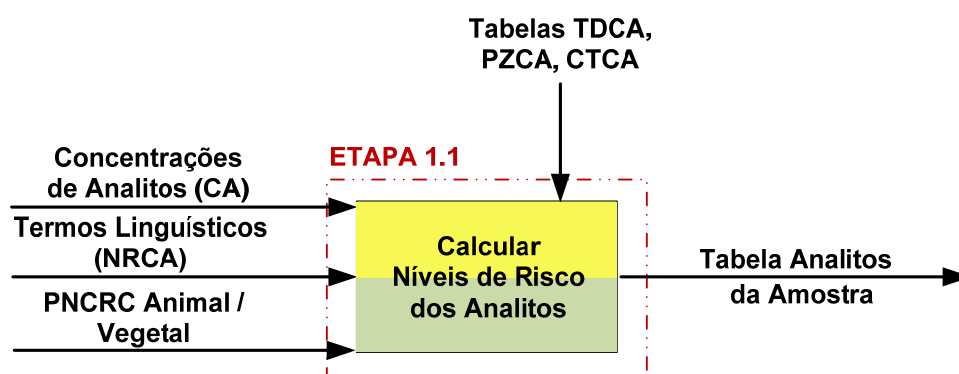


Figura 16 - Representação Gráfica da Etapa 1.1 da Plataforma Base Flexível – Macro-Etapa 1

**Sub-Etapa 1.1.1:** A partir das análises laboratoriais desenvolvidas nos laboratórios credenciados pelo PNCRC, obtém-se a Concentração do Analito (CA) ou intensidade de cada substância encontrada na amostra analisada. Ressalta-se que, em conformidade com a substância encontrada, a concentração pode ser publicada na forma de dado discreto, contínuo, nominal ou intervalar, para as diferentes matrizes agropecuárias analisadas pelo PNCRC, formando a Tabela Concentrações de Analitos – CA. A tabela 2 ilustra um fragmento desta tabela para a amostra de leite analisada:

TABELA 2 – Fragmento da Tabela Concentrações de Analitos – CA

| <b>Matriz:</b> Leite <b>Amostra</b>   <b>Lote:</b> 001   023 <b>Data:</b> 05/09/2010 |  |                                      |
|--|--|--------------------------------------|
| <b>Local da colheita:</b>  | Após ordenha                           |                                      |
| <b>Proprietário:</b>   | Cooperativa Central dos Produtores ... |                                      |
| <b>Município / UF:</b>   | Alto Jequitibá / MG                    |                                      |
| <b>Mesorregião:</b>  | Zona da Mata / MG                      |                                      |
| <b>Laboratório:</b>  | LANAGRO / MG                           |                                      |
| Ítem   | Analito                                | Concentração do Analito (CA em µg/L) |
| 1  | Metidation                             | 7                                    |
| 2  | Aflatoxina M1                          | 0.43                                 |
| 3  | Metamidofós                            | 10                                   |
| 4  | Ivermectina                            | ND                                   |
| 5  | Heptacloro                             | 4.92                                 |
| ....   | ....                                   | ....                                 |
| 43   | PCBs                                   | 2.56                                 |

ND = Não Detectado na Amostra

**Sub-Etapa 1.1.2:** Confronta-se a Concentração do Analito (CA) com seu Limite de Referência (LR), ou seja, com o Limite Máximo de Resíduo (LMR), caso a substância esteja classificada como resíduo, ou com o Teor Máximo de Contaminante (TMC), caso esteja classificada como contaminante. Conforme apresentado no Capítulo 3, os limites e teores máximos encontram-se publicados no PNCRC / Animal e PNCRC / Vegetal elaborados por especialistas do PNCRC baseados nas recomendações da Comissão do *Codex Alimentarius* – FAO/OMS/ONU. Uma parte dos Limites de Referência (LR) para as Concentrações dos Analitos (CA) detectados em uma amostra de leite está ilustrada na tabela 3.

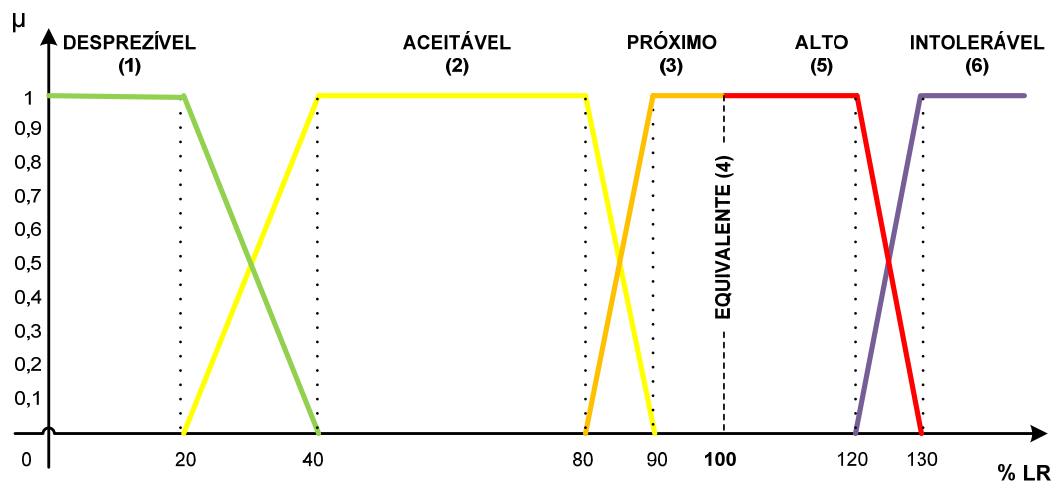
TABELA 3 – Limites de Referência para as Concentrações de Analitos Detectados em Amostra de Leite

| Analito       | LMR (em µg/L) | TMC  |
|---------------|---------------|------|
| Metidation    | 20            | ---- |
| Aflatoxina M1 | ----          | 0.5  |
| Metamidofós   | 10            | ---- |
| Ivermectina   | 10            | ---- |



|            |       |       |
|------------|-------|-------|
| Heptacloro | 4     | ----  |
| .....      | ..... | ..... |
| PCBs       | 2     | ----  |

**Sub-Etapa 1.1.3:** De acordo com o Limite Máximo de Resíduo (LMR) ou Teor Máximo de Contaminante (TMC) pré-determinado para cada analito, classifica-se nominalmente o fator de risco “Concentração do Analito (CA)” de acordo com os termos lingüísticos inerentes às faixas de risco associadas aos conjuntos *fuzzy*, definidos conjuntamente com os especialistas do PNCRC (Ver figura 17). Determina-se, assim, o Nível de Risco não agregado para cada Concentração de Analito (NRCA) detectada na amostra da matriz agropecuária analisada.



**Figura 17 – Conjuntos *fuzzy* Associados às Faixas de Risco para as Concentrações de Analitos Detectadas na Amostra de Matriz Agropecuária Analisada**

Analiticamente, as funções de pertinência ( $\mu$ ) associadas à forma trapezoidal dos conjuntos *fuzzy*, são:

$$\mu(\text{Desprezível}) = \begin{cases} 1, & \text{para } x \leq 20 \\ (40 - x) / (40 - 20), & \text{para } x \in (20,40) \text{ ou } 20 < x < 40 \\ 0, & \text{para } x \geq 40 \end{cases}$$

$$\mu(\text{Aceitável}) = \begin{cases} 0, & \text{para } x \leq 20 \\ (x - 20) / (40 - 20), & \text{para } x \in (20,40) \text{ ou } 20 < x < 40 \\ 1, & \text{para } x \in [40,80] \text{ ou } 40 \leq x \leq 80 \\ (90 - x) / (90 - 80) & \text{para } x \in (80,90) \text{ ou } 80 < x < 90 \\ 0, & \text{para } x \geq 90 \end{cases}$$

$$\mu(\text{Próximo}) = \begin{cases} 0, & \text{para } x \leq 80 \\ (x - 80) / (90 - 80), & \text{para } x \in (80,90) \text{ ou } 80 < x < 90 \\ 1, & \text{para } x \in [90,100) \text{ ou } 90 \leq x < 100 \\ 0, & \text{para } x \geq 100 \end{cases}$$

$$\mu(\text{Equivalente}) = \begin{cases} 1, & \text{para } x = 100 \end{cases}$$

$$\mu(\text{Alto}) = \begin{cases} 0, & \text{para } x \leq 100 \\ 1, & \text{para } x \in (100,120] \text{ ou } 100 < x \leq 120 \\ (130 - x) / (130 - 120), & \text{para } x \in (120,130) \text{ ou } 120 < x < 130 \\ 0, & \text{para } x \geq 130 \end{cases}$$

$$\mu(\text{Intolerável}) = \begin{cases} 0, & \text{para } x \leq 130 \\ (x - 120) / (130 - 120), & \text{para } x \in (120,130) \text{ ou } 120 < x < 130 \\ 1, & \text{para } x \geq 130 \end{cases}$$

Nesse novo processo para encontrar o NRCA, as funções de pertinência aos conjuntos *fuzzy* mapeiam o Resultado Percentual (RP), originado a partir do confronto entre a Concentração do Analito (CA) e seu respectivo Limite de Referência (LMR ou TMC), a um valor escalar entre 0 e 1 que mede o grau de pertinência do RP aos conjuntos *fuzzy*. Os números *fuzzy* inferidos a partir deste mapeamento traduzem os graus de pertinência do RP aos conjuntos *fuzzy*, e participam do cálculo associado ao NRCA como pequenos sobrepesos associados às faixas de risco.

Conforme observado na tabela 2 da Sub-Etapa 1.1.1, a análise laboratorial revelou uma concentração de 7 µg/L para o analito Metidation. Conforme informado na tabela 3 da Sub-Etapa 1.1.2 existe um LMR pré-determinado de 20 µg/L para este analito. Desta forma, a Concentração do Analito (CA) apresenta um Resultado Percentual (RP) equivalente a 35% do seu Limite de Referência (LR), no caso representado pelo seu LMR respectivo. Aplicado às funções de pertinência, o Resultado Percentual (RP) apresenta os seguintes graus de pertinência aos conjuntos *fuzzy*:

- $\mu(\text{Desprezível}) = (40 - 35) / (40 - 20) = 0.25$
- $\mu(\text{Aceitável}) = (35 - 20) / (40 - 20) = 0.75$
- $\mu(\text{Próximo}) = 0$
- $\mu(\text{Equivalente}) = 0$
- $\mu(\text{Alto}) = 0$
- $\mu(\text{Intolerável}) = 0$

Logo,

o NRCA para o Metidation apresentando concentração de 7 µg/kg será estabelecido em:

$$(0.25 \times 1) + (0.75 \times 2) + 0 + 0 + 0 + 0 = 1.75$$

Graficamente, o mapeamento do Resultado Percentual (RP = 35%) aos conjuntos *fuzzy* é ilustrado pela figura 18:

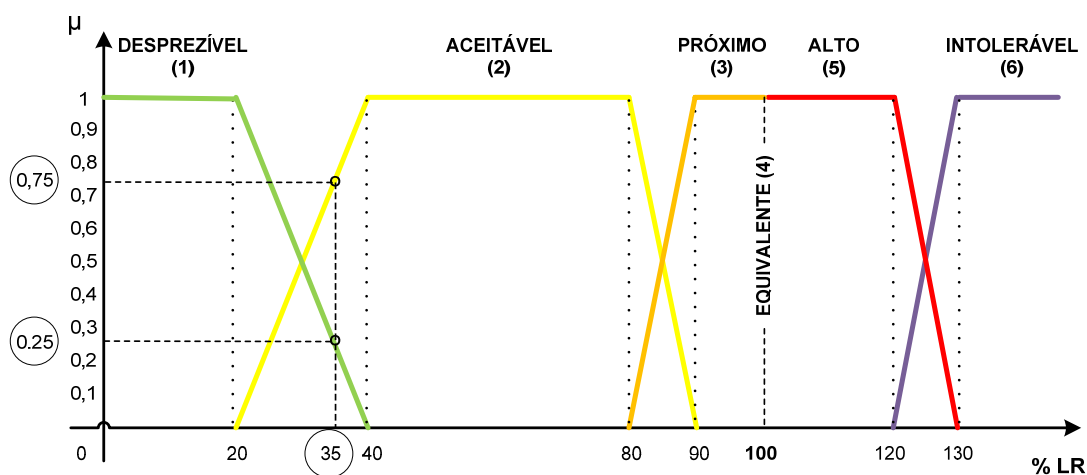


Figura 18 – Mapeamento Gráfico do Resultado Percentual (RP=35%) aos Conjuntos *Fuzzy*

A tabela 4 exibe um fragmento da Tabela Analitos da Amostra relacionando todos os NRCAs inferidos para os analitos da amostra de leite analisada.

**TABELA 4 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCAs Inferidos a partir dos Conjuntos *Fuzzy***

| Analito       | Concentração do Analito (CA em µg/L) | NRCA  |
|---------------|--------------------------------------|-------|
| Metidation    | 7                                    | 1.75  |
| Aflatoxina M1 | 0.43                                 | 2.70  |
| Metamidofós   | 10                                   | 4     |
| Ivermectina   | ND                                   | ND    |
| Heptacloro    | 4.92                                 | 5.30  |
| .....         | .....                                | ..... |
| PCBs          | 2.56                                 | 5.80  |

**Sub-Etapa 1.1.4:** Nesta sub-etapa os especialistas associam um valor representativo para o fator de risco “Tendência da Concentração do Analito” (ver tabela 6) para cada concentração detectada na amostra, de acordo com as características químicas, biológicas e patogênicas da substância, utilizando para esta finalidade, a tabela Tendências da Concentração do Analito – TDCA (ver tabela 5):

**TABELA 5 – Tendências da Concentração do Analito – TDCA**

| Fatores de Escala para Tendência da Concentração do Analito | Tendência da CA (TDCA) |
|---|------------------------|
| Regredir / Estabilizar-se                                   | 0                      |
| Agravar-se vagorosamente ou a longo prazo                   | 1                      |
| Agravar-se moderadamente ou a médio prazo                   | 2                      |
| Agravar-se rapidamente ou a curto prazo                     | 3                      |

**TABELA 6 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCA e TDCA**

| Analito       | Concentração do Analito (CA em µg/L) | NRCA | TDCA |
|---------------|--------------------------------------|------|------|
| Metidation    | 7                                    | 1.75 | 0    |
| Aflatoxina M1 | 0.43                                 | 2.70 | 2    |
| Metamidofós   | 10                                   | 4    | 3    |
| Ivermectina   | ND                                   | ND   | ND   |
| Heptaclo      | 4.92                                 | 5.30 | 0    |
| ....          | ....                                 | .... | .... |
| PCBs          | 2.56                                 | 5.80 | 0    |

**Sub-Etapa 1.1.5:** Nesta sub-etapa os especialistas consideram o fator de risco “Prazo de ajustamento aos índices mundiais de segurança química (Limites de Referência)” ou, de forma equivalente, aos Sistemas de Autocontroles promovidos pelo MAPA visando o *continuum* da produção e comercialização dos produtos agropecuários brasileiros (ver tabela 8). A tabela PZCA – Prazos Estimados de Ajustamento da Concentração do Analito (tabela 7) é utilizada em atendimento a esta finalidade. Atualmente sete autocontroles encontram-se definidos pela norma ISO 22000/2006, quais sejam: Boas Práticas Pecuárias (BPP); Boas Práticas Agrícolas (BPA); Boas Práticas de Fabricação (BPF); Boas Práticas de Higiene (BPH); Boas Práticas de

Manipulação (BPM); Boas Práticas de Distribuição (BPD) e Boas Práticas de Comercialização (BPC).

**TABELA 7 – Prazos Estimados de Ajustamento da Concentração do Analito (CA) aos Autocontroles do MAPA – PZCA**

| Fatores de Escala para Prazos Estimados de Ajustamento aos Autocontroles | Prazo Estimado de Ajustamento (PZCA) |
|--|--------------------------------------|
| Desnecessário  | 0                                    |
| Curto  | 1                                    |
| Médio  | 2                                    |
| Longo  | 3                                    |

Deve ser observado que os prazos “1”, “2” e “3” não se coadunam aos níveis de risco da concentração “1”, “2”, “3” e “4” ilustrados pelos conjuntos *fuzzy* na figura 18. Da mesma forma, o prazo “0” não se coaduna aos níveis de risco “5” e “6”. O prazo “0” também não se coaduna às tendências de agravamento “1”, “2” e “3”.

**TABELA 8 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCA, TDCA e PZCA**

| Analito       | Concentração do Analito (CA em µg/L) | NRCA | TDCA | PZCA |
|---------------|--------------------------------------|------|------|------|
| Metidation    | 7                                    | 1.75 | 0    | 0    |
| Aflatoxina M1 | 0.43                                 | 2.70 | 2    | 1    |
| Metamidofós   | 10                                   | 4    | 3    | 1    |
| Ivermectina   | ND                                   | ND   | ND   | ND   |
| Heptacloro    | 4.92                                 | 5.30 | 0    | 2    |
| ....          | ....                                 | .... | .... | .... |
| PCBs          | 2.56                                 | 5.80 | 0    | 3    |

**Sub-Etapa 1.1.6:** Semelhantemente à Sub-Etapa 1.1.5, os especialistas nesta sub-etapa necessitam associar o fator de risco “Custo de ajustamento aos índices de segurança química mundiais” ou, de forma equivalente, aos Sistemas de Autocontroles promovidos pelo MAPA visando o *continuum* da produção e comercialização dos produtos agropecuários brasileiros (ver tabela 10). A tabela CTCA – Custos Estimados de Ajustamento da Concentração do Analito, apresentada na tabela 9 é utilizada em atendimento a esta finalidade.

**TABELA 9 – Tabela de Custos Estimados de Ajustamento da Concentração do Analito (CA) aos Autocontroles do MAPA – CTCA**

| Fatores de Escala para Custos Estimados de Ajustamento aos Autocontroles | Custo Estimado de Ajustamento (CTCA) |
|--|--------------------------------------|
| Desnecessário  | 0                                    |
| Baixo  | 1                                    |
| Médio  | 2                                    |
| Alto   | 3                                    |

Deve ser observado que os custos “1”, “2” e “3” não se coadunam aos níveis de risco da concentração “1”, “2”, “3” e “4” ilustrados pelos conjuntos *fuzzy* na figura 18. Da mesma forma, o custo “0” não se coaduna aos níveis de risco “5” e “6”. O custo “0” também não se coaduna às tendências de agravamento “1”, “2” e “3”.

**TABELA 10 – Fragmento da Tabela de Analitos da Amostra Exibindo NRCA, TDCA, PZCA e CTCA**

| Analito       | Concentração do Analito (CA em µg/L) | NRCA  | TDCA  | PZCA  | CTCA  |
|---------------|--------------------------------------|-------|-------|-------|-------|
| Metidation    | 7                                    | 1.75  | 0     | 0     | 0     |
| Aflatoxina M1 | 0.43                                 | 2.70  | 2     | 1     | 2     |
| Metamidofós   | 10                                   | 4     | 3     | 1     | 1     |
| Ivermectina   | ND                                   | ND    | ND    | ND    | ND    |
| Heptacloro    | 4.92                                 | 5.30  | 0     | 2     | 1     |
| .....         | .....                                | ..... | ..... | ..... | ..... |
| PCBs          | 2.56                                 | 5.80  | 0     | 3     | 2     |

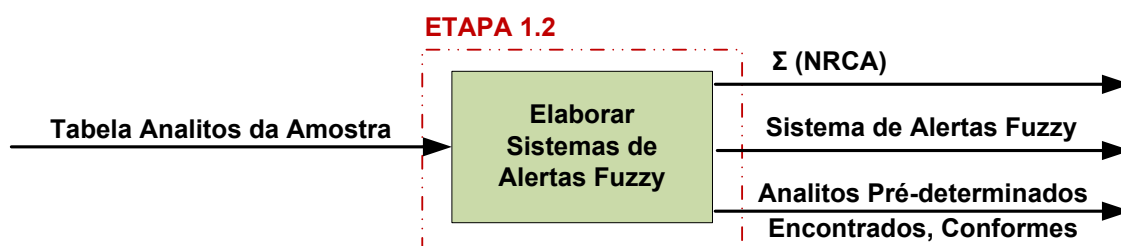
**Sub-Etapa 1.1.7:** A partir das informações obtidas nas sub-etapas anteriores, os especialistas podem calcular, para cada analito encontrado na amostra da matriz analisada, o Nível de Risco Agregado do Analito (NRAA) de acordo com o intervalo  $1 \leq \text{NRAA} \leq 15$ . O valor 1 (um) reflete o Limite mínimo de Conformidade do NRAA (LmC-NRAA) e o valor 15 reflete o Limite Máximo de Não Conformidade do NRAA (LMNC-NRAA). A tabela Analitos da Amostra (ver tabela 11) exhibe o NRAA inferido para cada analito detectado, assim como, o somatório dos níveis de risco NRCA e NRAA inferidos ao final deste processo.

*Nota.* Deve ser observado que o intervalo acima engloba tanto os analitos “conformes” quanto os analitos “não conformes”, porventura encontrados na amostra analisada. Esta situação é refletida pelos somatórios ( $\Sigma$ ) do NRCA e NRAA.

**TABELA 11 – Fragmento da Tabela Analitos da Amostra Apresentando o Somatório ( $\Sigma$ ) dos Níveis de Risco NRCA e NRAA**

| Analito                                      | Concentração do Analito (CA em $\mu\text{g/L}$ ) | NRCA (1)     | TDCA (2) | PZCA (3) | CTCA (4) | NRAA (1+2+3+4) |
|--|--|--------------|----------|----------|----------|----------------|
| Metidation                                   | 7  | 1.75         | 0        | 0        | 0        | 1.75           |
| Aflatoxina M1                                | 0.43   | 2.70         | 2        | 1        | 2        | 7.70           |
| Metamidofós                                  | 10   | 4            | 3        | 1        | 1        | 9              |
| Ivermectina                                  | ND   | ND           | ND       | ND       | ND       | ND             |
| Heptacloro                                   | 4.92   | 5.30         | 0        | 2        | 1        | 8.30           |
| .....  | .....  | .....        | .....    | .....    | .....    | .....          |
| PCBs   | 2.56   | 5.80         | 0        | 3        | 2        | 10.80          |
| <b>(<math>\Sigma</math>) Níveis de Risco</b> |  | <b>53.37</b> |          |          |          | <b>117.30</b>  |

**Etapa 1.2:** Nesta etapa o Sistema de Alertas *Fuzzy* é elaborado a partir dos resultados alcançados nas sub-etapas que compõem a Etapa 1.1. A figura 19 ilustra graficamente a Etapa 1.2. A cor verde indica a existência de procedimentos computacionais inerentes à lógica *fuzzy*:



**Figura 19 - Representação Gráfica da Etapa 1.2 da Plataforma base Flexível – Macro-Etapa 1**

Com base na tabela 11, um sistema de alertas *fuzzy* sugere aos analistas de risco e autoridades do PNCRC quais problemas devem ser prioritariamente considerados. Os analitos que porventura apresentaram NRCA > 4 serão destacados na cor vermelha e agrupados nas primeiras posições da tabela Sistema de Alertas (ver tabela 12), sugerindo intervenção máxima ao sistema produtivo. Os analitos que porventura apresentaram tendência de agravamento da concentração, ou seja, TDCA > 0, serão destacados na cor laranja e agrupados na segunda posição da tabela, sugerindo a implantação de ações regulatórias preventivas por parte das autoridades do PNCRC. Os analitos que porventura apresentaram conformidade aos seus limites de referência serão destacados pela

cor azul claro e agrupados na terceira posição da tabela, indicando não existir necessidade de intervenções e ações regulatórias ao sistema produtivo. Todos os grupos de analitos destacados em vermelho, laranja ou azul claro serão hierarquizados individualmente em ordem decrescente, segundo o maior valor do NRAA inferido no grupo.

*Notas.*

1) A hierarquização dos NRAA's fundamenta-se na experiência de que a tarefa de adequação simultânea de todos os analitos que apresentaram não conformidade aos respectivos índices mundiais de segurança química ou apresentaram tendência de agravamento da concentração, pode traduzir uma ação confusa, contraproducente ou mesmo ineficiente;

2) O NRCA revela quais analitos apresentaram conformidade ou não conformidade aos índices mundiais de segurança química, enquanto que o NRAA realiza uma previsão para cada concentração detectada, incorporando a esta, de acordo com a experiência de especialistas, uma tendência estimada, uma estimativa de prazo e de custo de ajustamento do processo produtivo aos índices de segurança mundiais ou, em última palavra, aos sistemas de autocontrole do processo produtivo agropecuário promovidos pelo PNCRC / MAPA. Ao revelar as concentrações de analitos com tendência de superação a estes índices, o NRAA possibilita a existência de tratamento preventivo por parte do PNCRC, facilitando a aplicação antecipada de ações regulatórias ou intervenções ao processo produtivo agropecuário, evitando ações de emergência, desperdício de tempo e capital.

**TABELA 12 - Fragmento Tabela Analitos da Amostra Apresentando o Sistema de Alertas Fuzzy Baseado no Nível de Risco Agregado do Analito - NRAA**

| <b>Analito</b>             | <b>Concentração do Analito (CA em µg/L)</b> | <b>NRCA (1)</b> | <b>TDCA (2)</b> | <b>PZCA (3)</b> | <b>CTCA (4)</b> | <b>NRAA (1+2+3+4)</b> |
|----------------------------|---|-----------------|-----------------|-----------------|-----------------|-----------------------|
| PCBs                       | 2.56  | 5.80            | 0               | 3               | 2               | 10.80                 |
| Heptacloro                 | 4.92  | 5.30            | 0               | 2               | 1               | 8.30                  |
| Metamidofós                | 10  | 4               | 3               | 1               | 1               | 9                     |
| Aflatoxina M1              | 0.43  | 2.70            | 2               | 1               | 2               | 7.70                  |
| .....                      | .....                                       | .....           | .....           | .....           | .....           | .....                 |
| Metidation                 | 7   | 1.75            | 0               | 0               | 0               | 1.75                  |
| Ivermectina                | ND  | ND              | ND              | ND              | ND              | ND                    |
| <b>(Σ) Níveis de Risco</b> |   | <b>53.37</b>    |                 |                 |                 | <b>117.30</b>         |



**Etapa 1.3:** Nesta etapa são aferidos os níveis de risco associados à amostra analisada, avaliados os níveis de toxicidade dos analitos não conformes porventura detectados, e analisada a existência de associações perigosas de analitos na amostra, sendo informados os níveis de risco das associações perigosas detectadas. A figura 20 ilustra graficamente a Etapa 1.3 que engloba as Sub-Etapas 1.3.1 a 1.3.4. As cores verde e salmão, destacadas no processo, indicam, respectivamente, a existência de procedimentos computacionais inerentes à lógica *fuzzy* e ontologias.

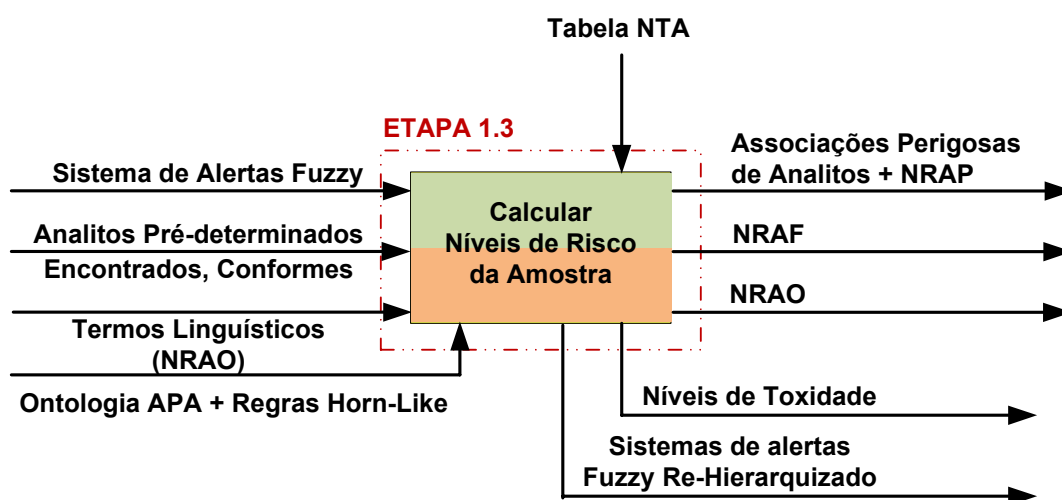


Figura 20 - Representação Gráfica da Etapa 1.3 da Plataforma base Flexível – Macro-Etapa 1

**Sub-Etapa 1.3.1:** Nesta sub-etapa é calculado o Nível de Risco da Amostra Original (NRAO), considerando tanto a existência conjunta de analitos conformes e não conformes, quanto a possibilidade da existência somente de analitos conformes ou não conformes em uma amostra. Deve ser ressaltado que basta a existência de um único analito “não conforme” para que a amostra inteira seja classificada como “não conforme” ou “não apropriada” ao consumo humano ou animal. Por esta razão, o Nível de Risco da Amostra (NRAO) necessita ser inferido de forma distinta para amostras conformes e não conformes. Amostras conformes apresentam todos os analitos detectados em conformidade com seus LR’s respectivos, traduzidos formalmente pelo intervalo de conformidade  $1 \leq NRCA \leq 4$ . Amostras não conformes apresentam um ou mais analitos em não conformidade com seus LR’s respectivos, traduzidos formalmente por  $NRCA > 4$ . Os limites mínimo e máximo do intervalo de conformidade, correspondem aos índices Limite

mínimo de Conformidade do NRCA (LmC-NRCA) e Limite Máximo de Conformidade do NRCA (LMC-NRCA).

Para encontrar o NRAO respectivo a uma amostra “conforme” ou “apropriada” ao consumo humano, os seguintes passos devem ser considerados:

- Estabelecer o Intervalo de Conformidade da Amostra (ICA) considerando os limites mínimo (LmC-NRAO) e máximo (LMC-NRAO) de conformidade para o Nível de Risco da Amostra Original (NRAO), em função do número de analitos detectados, de acordo com as seguintes fórmulas:

$$\circ \text{LmC-NRAO} = (\text{Nr}^{\circ} \text{ de Analitos Detectados} * \text{LmC-NRCA})$$

$$\circ \text{LMC-NRAO} = (\text{Nr}^{\circ} \text{ de Analitos Detectados} * \text{LMC-NRCA})$$

- Encontrar o Percentual de Conformidade (PC) da amostra analisada, de acordo com a seguinte fórmula:

$$\text{PC} = (((\Sigma\text{-NRCA} - \text{LmC-NRAO}) / (\text{LMC\_NRAO} - \text{LmC-NRAO})) * 100)$$

- Confrontar o PC com os conjuntos *fuzzy*, estabelecidos conjuntamente com os especialistas do PNCRC, para as faixas de risco associadas às amostras conformes (ver figura 21), visando encontrar o Nível de Risco da Amostra Original (NRAO):

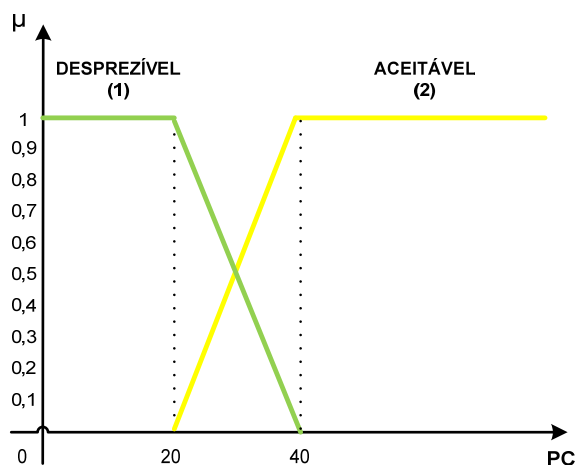


Figura 21 – Conjuntos *Fuzzy* Pré-estabelecidos para Faixas de Risco de Amostras Conformes

Analicamente, as funções de pertinência associadas à forma trapezoidal dos conjuntos *fuzzy*, pré-estabelecidos para faixas de risco de amostras conformes, são:

$$\mu(\text{Desprezível}) = \begin{cases} 1, & \text{para } x \leq 20 \\ (40 - x) / (40 - 20), & \text{para } x \in (20,40) \text{ ou } 20 < x < 40 \\ 0, & \text{para } x \geq 40 \end{cases}$$

$$\mu(\text{Aceitável}) = \begin{cases} 0, & \text{para } x \leq 20 \\ (x - 20) / (40 - 20), & \text{para } x \in (20,40) \text{ ou } 20 < x < 40 \\ 1, & \text{para } x \geq 40 \end{cases}$$

Neste exemplo estamos considerando que foram detectados 31 analitos, entre os 43 pré-determinados para uma amostra de leite. Vamos supor que todos os analitos detectados apresentaram conformidade em relação aos seus respectivos LR's, traduzindo um  $\Sigma$ -NRCA de 53.37.

O Intervalo de Conformidade da Amostra (ICA) seria estabelecido da seguinte forma:

- LmC-NRAO = 31 \* 1 = 31
- LMC-NRAO = 31 \* 4 = 124

Logo,

o ICA = [31; 124] e o Percentual de Conformidade (PC) é:

$$PC = (((53.37 - 31) / (124 - 31)) * 100) = 24.05\%$$

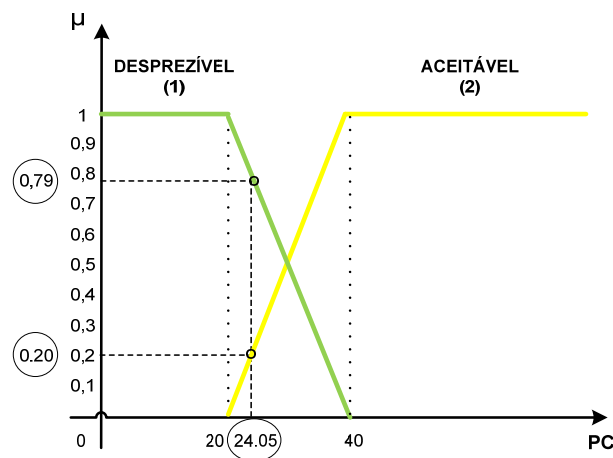
Aplicado às funções de pertinência, o Percentual de Conformidade (PC) apresenta os seguintes graus de pertinência aos conjuntos *fuzzy*:

- $\mu(\text{Desprezível}) = (40 - 24.05) / (40 - 20) \approx 0.79$
- $\mu(\text{Aceitável}) = (24.05 - 20) / (40 - 20) \approx 0.20$

Logo, o NRAO para a amostra conforme deverá ser estabelecido em:

$$(0.79 \times 1) + (0.20 \times 2) = 1.19$$

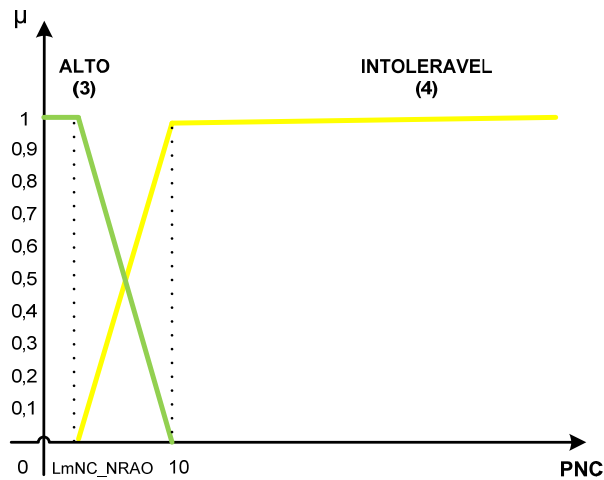
Graficamente, o mapeamento do Percentual de Conformidade (PC = 24.05%) aos conjuntos *fuzzy* é ilustrado pela figura 22:



**Figura 22 – Mapeamento Gráfico do Percentual do Conformidade (PC=24.05%) aos Conjuntos *Fuzzy* Pré-estabelecidos para Faixas de Risco de Amostras Conformes**

De outra forma, para encontrar o NRAO respectivo a uma amostra “não conforme” ou “não apropriada” ao consumo humano e animal, que apresenta um ou mais analitos em não conformidade com seu(s) respectivo(s) LR’s, traduzido(s) por NRCA > 4, os seguintes passos deverão ser considerados:

- Observar o Intervalo de Não Conformidade da Amostra (INCA) considerando os limites mínimo (LmNC-NRAO), correspondente a encontrar um único analito “não conforme” e máximo (LMNC-NRAO), correspondente a encontrar todos os analitos em não conformidade, de acordo com as seguintes fórmulas:
  - $LmNC-NRAO = (1 / Nr^o \text{ de Analitos Pesquisados}) * 100 = (1 / 43) * 100 = 2.32 \%$
  - $LMNC-NRAO = 100\%$
  
- Encontrar o Percentual de Não Conformidade (PNC) da amostra analisada, de acordo com a seguinte fórmula:
 
$$PNC = (Nr^o \text{ de Analitos Não Conformes} / Nr^o \text{ de Analitos Pesquisados}) * 100$$
  
- Confrontar o PNC com os conjuntos *fuzzy*, estabelecidos conjuntamente com os especialistas do PNCRC, para as faixas de risco associadas às amostras não conformes (ver figura 23), visando encontrar o Nível de Risco da Amostra Original (NRAO):



**Figura 23 – Conjuntos Fuzzy Pré-estabelecidos para Faixas de Risco de Amostras Não Conformes**

Analicamente, as funções de pertinência associadas à forma trapezoidal dos conjuntos *fuzzy*, pré-estabelecidos para faixas de risco de amostras não conformes, são:

$$\mu(\text{Alto}) = \begin{cases} 1, & \text{para } x \leq \text{LmNC\_NRAO} \\ (10 - x) / (10 - \text{LmNC\_NRAO}), & \text{para } \text{LmNC\_NRAO} < x < 10 \\ 0, & \text{para } x \geq 10 \end{cases}$$

$$\mu(\text{Intolerável}) = \begin{cases} 0, & \text{para } x \leq \text{LmNC\_NRAO} \\ (x - \text{LmNC\_NRAO}) / (10 - \text{LmNC\_NRAO}), & \text{para } \text{LmNC\_NRAO} < x < 10 \\ 1, & \text{para } x \geq 10 \end{cases}$$

Neste exemplo vimos que ao final da Etapa 1.2, o Sistema de Alertas *Fuzzy* destacou a existência de 2 analitos apresentando não conformidade aos seus respectivos LR's, para um total de 43 analitos pesquisados na amostra de leite.

O Intervalo de Não Conformidade da Amostra (INCA) seria estabelecido da seguinte forma:

- LmNC-NRAO = 2.32 % (1 analito apresentando não conformidade)
- LMNC-NRAO = 100% (Todos os 43 analitos apresentando não conformidade)

Logo,

o INCA = [2.32; 100]. O Percentual de Não Conformidade (PNC) é:

$$\text{PNC} = 2 / 43 * 100 = 4.65\%$$

Aplicado às funções de pertinência, o Percentual de Não Conformidade (PNC) apresenta os seguintes graus de pertinência aos conjuntos *fuzzy*:

- $\mu(\text{Alto}) = (10 - 4.65) / (10 - 2.32) = 0.69$
- $\mu(\text{Intolerável}) = (4.65 - 2.32) / (10 - 2.32) = 0.31$

Logo,

o NRAO para a amostra não conforme deverá ser estabelecido em:

$$(0.69 \times 3) + (0.31 \times 4) = 3.31$$

Graficamente, o mapeamento do Percentual de Não Conformidade (PNC = 4.65%) aos conjuntos *fuzzy* é ilustrado pela figura 24:

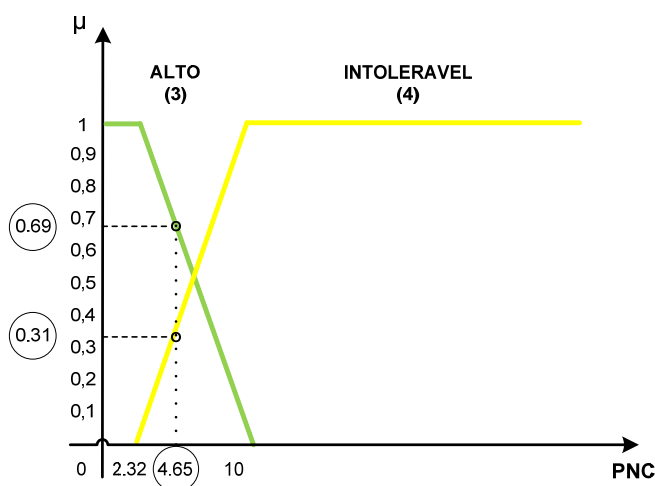


Figura 24 – Mapeamento Gráfico do Percentual de Não Conformidade (PNC=4.65%) aos Conjuntos *Fuzzy* Pré-estabelecidos para Faixas de Risco de Amostras Não Conformes

**Sub-Etapa 1.3.2:** De forma complementar ao Sistema de Alertas *Fuzzy* que denuncia a existência de analitos não conformes na amostra analisada, os especialistas necessitam considerar o fator de risco Nível de Toxicidade dos Analitos Não Conformes, visando promover uma rehierarquização dos analitos não conformes, porventura encontrados, de acordo com a gravidade da enfermidade que poderá causar ao organismo humano ou animal. Desta forma, um analito não conforme que apresente NRCA inferior aos outros analitos não conformes detectados na amostra analisada pode ocupar a primeira posição na tabela Analitos da Amostra, promovendo uma reclassificação dos analitos não conformes devido à superioridade do seu potencial maligno. Na tabela 14, este é o caso do analito Heptacloro que apresenta potencial maligno superior aos PCBs

detectados na amostra de leite analisada. A tabela 13 ilustra um fragmento da tabela Níveis de Toxicidade dos Analitos Não Conformes utilizada em atendimento a esta finalidade, cujo Nível de Toxicidade do Analito (NTA) apresenta uma escala limitada aos valores “Baixo”, “Médio” e “Alto”:

**TABELA 13 – Níveis de Toxicidade de Analitos Não Conformes - NTA**

| <b>Grupo</b>     | <b>Analito</b> | <b>Nível de Toxicidade do Analito (NTA)</b> |
|------------------|----------------|---|
| Organofosforados | Metidation     | Alto  |
|                  | Metamidofós    | Alto  |
| PCBs             | PCBs           | Médio                                       |
| Organoclorados   | Aldrin         | Alto  |
|                  | Heptacoloro    | Alto  |
|                  | Alfa HCH       | Alto  |
| Micotoxinas      | Aflatoxina M1  | Alto  |
| .....            | .....          | .....                                       |

**TABELA 14 – Fragmento da Tabela Analitos da Amostra Apresentando o Sistema de Alertas *Fuzzy* Re-hierarquizado para os Analitos Não Conformes Detectados na Amostra Analisada**

| <b>Analito</b>             | <b>Concentração do Analito (CA em µg/L)</b> | <b>NRCA (1)</b> | <b>TDCA (2)</b> | <b>PZCA (3)</b> | <b>CTCA (4)</b> | <b>NRAA (1+2+3+4)</b> |
|----------------------------|---|-----------------|-----------------|-----------------|-----------------|-----------------------|
| Heptacoloro                | 4.92  | 5.30            | 0               | 2               | 1               | 8.12                  |
| PCBs                       | 2.56  | 5.80            | 0               | 3               | 2               | 10.27                 |
| Metamidofós                | 10  | 4               | 3               | 1               | 1               | 9                     |
| Aflatoxina M1              | 0.43  | 2.70            | 2               | 1               | 2               | 7.70                  |
| .....                      | .....                                       | .....           | .....           | .....           | .....           | .....                 |
| Metidation                 | 7   | 1.75            | 0               | 0               | 0               | 1.75                  |
| Ivermectina                | ND  | ND              | ND              | ND              | ND              | ND                    |
| <b>(Σ) Níveis de Risco</b> |   | <b>53.37</b>    |                 |                 |                 | <b>117.30</b>         |

**Sub-Etapa 1.3.3:** Adicionalmente à maior acurácia, precisão e escopo proporcionados pelo NRAO por aplicar os conceitos da lógica *fuzzy* e trabalhar com o menor grão de informação ao nível de risco dos analitos, os especialistas necessitam considerar a existência do fator de risco agregado “Associações Perigosas de Analitos - APA”, evitando, assim, acumular nova perda de informações importantes no cálculo do nível de risco das amostras analisadas, às perdas inicialmente detectadas ao nível dos analitos e amostras detectadas pela aplicação da lógica *fuzzy*.

Este fator de risco agregado engloba a existência de fatores de risco individuais que relatam a interação entre dois ou mais agentes intoxicantes (analitos) que causam danos ao sistema biológico humano e animal, seja alterando uma função vital ou mesmo levando à morte, sob certas circunstâncias. Os efeitos tóxicos e os antídotos são geralmente proporcionais à Concentração do Analito (CA) detectado, podendo, quando combinados, serem maiores ou menores que a soma dos efeitos tóxicos ou antídotos individuais. Os fatores de risco individuais que compõem as Associações Perigosas de Analitos - APA são:

- Efeito Toxicológico Aditivo – ETA: O efeito tóxico final da associação é equivalente à soma dos efeitos tóxicos de cada um dos analitos envolvidos ( $1+1=2$ ) ( $1+1+1=3$ ), etc.;
- Efeito Toxicológico Sinérgico – ETS: O efeito tóxico final da associação é superior à soma dos efeitos tóxicos de cada um dos analitos envolvidos ( $1+1>2$ ), ( $1+1+1>3$ ), etc.;
- Efeito Toxicológico de Potenciação – ETP: O efeito tóxico final da associação é ocasionado pelo aumento da concentração de um analito particular (CA) quando em combinação com outro analito ( $0+1>1$ );
- Efeito Toxicológico Redutor – ETR: O efeito tóxico final de um analito particular é reduzido em função da presença de outro.

A presença dos fatores de risco individuais em uma amostra, em geral, agrava o Nível de Risco da Amostra Original (NRAO) inferido na Sub-Etapa 1.3.1. Deve ser esclarecido, contudo, que os efeitos redutores proporcionados pelos antídotos mascaram o verdadeiro nível de concentração (CA) que seria normalmente detectado para determinados analitos, ocultando, assim, a provável existência de práticas ilegais durante o processo produtivo da matriz analisada que contrariam as recomendações dos sistemas de autocontroles. Por esta razão, os efeitos redutores também necessitam ser detectados e corrigidos a tempo.

A tabela 15 apresenta um fragmento da tabela Associações Perigosas de Analitos – APA, cujo NRAP (Nível de Risco da Associação Perigosa) caracteriza-se por apresentar uma escala limitada aos valores “Alto” e “Intolerável”, devido ao contexto desta pesquisa focar nos efeitos toxicológicos prejudiciais das associações.



**TABELA 15 - Fragmento da Tabela Associações Perigosas de Analitos – APA**

| APA  | Descrição Técnica da Associação Perigosa   | Nível de Risco (NRAP) |
|------|--|-----------------------|
| 001  | Descrição Técnica da Associação Perigosa Nrº 1 ...   | 3-Alto                |
| 002  | Descrição Técnica da Associação Perigosa Nrº 2 ...   | 4-Intolerável         |
| 003  | Analitos pertencentes ao grupo organofosforados encontrados conjuntamente em uma amostra de matriz agropecuária qualquer, apresentam efeito aditivo, ocasionando hepato-enfermidades graves.   | 4-Intolerável         |
| 004  | Aflatoxina B1 detectada em qualquer nível de concentração encontrada conjuntamente com Aflatoxina G1 detectada em qualquer nível de concentração, ambas em amostra de matriz de feijão, apresentam efeito sinérgico com alto potencial cancerígeno | 4-Intolerável         |
| 005  | Descrição Técnica da Associação Perigosa Nrº 5 ...   | 3-Alto                |
| .... | ....   | ....                  |

**Sub-Etapa 1.3.4:** Nesta sub-etapa o Nível de Risco da Amostra Final (NRAF) é inferido a partir do NRAO e do maior Nível de Risco das Associações Perigosas de Analitos detectadas na amostra da matriz de leite analisada. Dando continuidade ao exemplo, observamos na Sub-Etapa 1.3.1 que devido à presença de 2 analitos não conformes, o NRAO estimado para a amostra de leite foi de 3.31, correspondendo a aproximadamente 0.69 do nível de risco “3-Alto” e 0.31 do nível “4-Intolerável”, para um Percentual de Não Conformidade (PNC) de 4.65%. Entretanto, a existência da associação perigosa APA-003, revelada pela presença dos organofosforados Metidation e Metamidofós na tabela Analitos da Amostra (Sub-Etapa 1.3.2, tabela 14), sugere um NRAP = “4-Intolerável” para a amostra analisada. Caso existissem outras associações perigosas na amostra analisada, o maior NRAP sugerido seria adotado. Uma vez que o NRAP sugerido é superior ao NRAO inferido, ou seja,  $NRAP > NRAO$ , a plataforma base flexível recomenda como diagnóstico final para a amostra de leite analisada, um Nível de Risco da Amostra Final (NRAF) de “4 - Intolerável”.

*Nota.* Observe que, para os casos em que  $NRAO = NRAP$  ou  $NRAO > NRAP$ , o NRAO irá prevalecer como Nível de Risco da Amostra Final (NRAF).

**Etapa 1.4:** Com base nos resultados das análises de risco associadas aos analitos e amostras de matrizes agropecuárias, divulgadas pelos laboratórios credenciados, a plataforma base flexível sugere aos especialistas e autoridades do PNCRC um Nível de Intervenção e Ações Regulatórias (NIAR) ao processo de produção e comercialização dos produtos agropecuários desenvolvido pelo proprietário da amostra em determinada região do país. A figura 25 ilustra graficamente a Etapa 1.4 que apresenta a cor amarela destacando a existência de um processo administrativo:

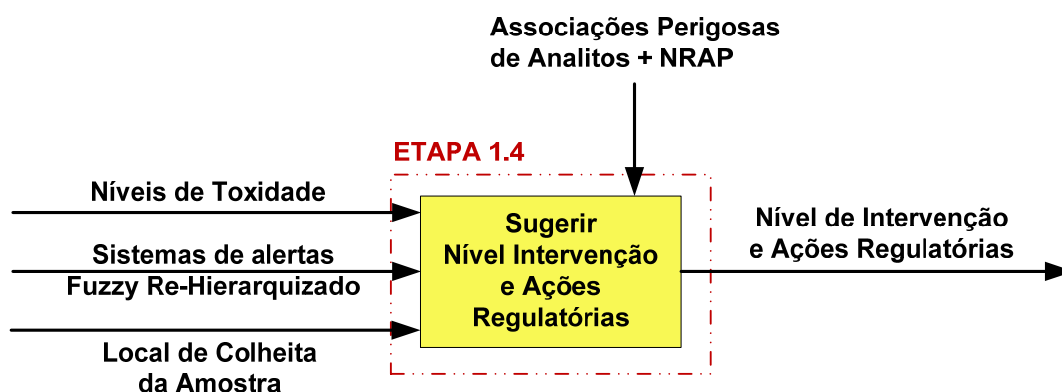


Figura 25 - Representação Gráfica da Etapa 1.4 da Plataforma base Flexível – Macro-Etapa 1

Entre outros fatores, a sugestão do NIAR engloba o fator de risco “Local de Colheita da Amostra” considerando as alternativas “Anteriormente à embalagem” e “Posteriormente à embalagem”. O NIAR representa um conjunto de ações (medidas) de cunho fiscal-administrativo que visa impedir a existência de práticas ilegais durante o processo produtivo, auxiliando o produtor a realizar um balizamento da sua produção, especificamente em relação ao conhecimento e observação dos autocontroles no *continuum* da sua produção, visando que os produtos agropecuários brasileiros respeitem os índices de qualidade e segurança alimentar mundialmente recomendados. Tecnicamente, perante o contexto maior do processo de gerência de riscos (ALENCAR; SCHMITZ, 2006), citado ao longo da seção 2.2.4, trata-se de um conjunto de ações pós-análise de risco, ou seja, estas medidas não mais encontram-se relacionadas ao sub-processo análise de riscos, e, sim, ao sub-processo controle de riscos. Este conjunto de medidas pós-análise de riscos visa maximizar os impactos positivos e minimizar os impactos negativos dos fatores de risco que incidem ao nível dos analitos e amostras de matrizes agropecuárias, mantendo-os em uma faixa de variabilidade e incerteza aceitáveis. Adicionalmente à severidade

do nível de intervenção ao processo produtivo, as ações regulatórias inerentes à intervenção encontram-se descritas na tabela de Níveis de Intervenções e Ações Regulatórias – NIAR, ilustrada na tabela 16:

**TABELA 16 – Tabela Níveis de Intervenção e Ações Regulatórias – NIAR**

| <b>Nível de Intervenção (NIAR)</b> | <b>Ações Regulatórias Aplicadas ao Sistema Produtivo</b>   |
|------------------------------------|--|
| 1 – Desnecessário                  | * Observa-se bom nível de adequação do processo produtivo aos autocontroles recomendados.  |
| 2 - Investigatório                 | * Analisar as causas de proximidade (em 10%) ou equivalência ao LR;<br>* Emitir orientações técnicas visando maior adequação aos autocontroles recomendados;<br>* Provável período “imediato – curto” de ajustamento aos autocontroles a “custo baixo”.  |
| 3 - Moderado                       | * Analisar em curto prazo as causas de superação moderada (em 10%) ao LR para amostras colhidas anteriormente à embalagem;<br>* Suspensão imediata do processo de comercialização com rastreamento das causas de não conformidade;<br>* Exigir adequação imediata aos autocontroles;<br>* Período estimado de ajustamento aos autocontroles: Curto – Médio;<br>* Custo estimado de ajustamento aos autocontroles: Médio.   |
| 4 - Máximo                         | * Analisar em curtíssimo prazo a(s) causa(s) de:<br>- Superação ao LR para amostras colhidas pós-embalagem;<br>- Superação moderada (em 10%) ao LR com existência de Níveis de Toxicidade de Analitos Não-Conformes com propensão a causar enfermidades graves;<br>- Forte superação ao LR (acima de 10%);<br>- Existência de Associações Perigosas de Analitos;<br>- Existência de Analitos Proibidos ou Não Permitidos para a Cultura (NPC);<br>* Suspensão imediata do processo de produção e/ou comercialização com rastreamento das causas de não conformidade;<br>* Abertura de processo de investigação e monitoramento;<br>* Exigir adequação imediata aos autocontroles;<br>* Período estimado de ajustamento aos autocontroles: Médio-Longo;<br>* Custo estimado de ajustamento aos autocontroles: Alto. |

Supondo que a amostra de leite, considerada neste exemplo, tenha sido colhida pós-embalagem e, com base nas informações de que o analito PCBs apresenta concentração superior a 10% do seu LR e existe a associação perigosa de analitos número 003 na amostra analisada, a

Plataforma Base Flexível sugere aos especialistas do PNCRC um NIAR = 4-Máximo para a amostra analisada.

**Etapa 1.5:** A partir das informações produzidas pelas etapas anteriores, os especialistas do Ministério podem obter uma Matriz de Riscos para a amostra de leite analisada, apresentando as seguintes informações ilustradas na figura 26:

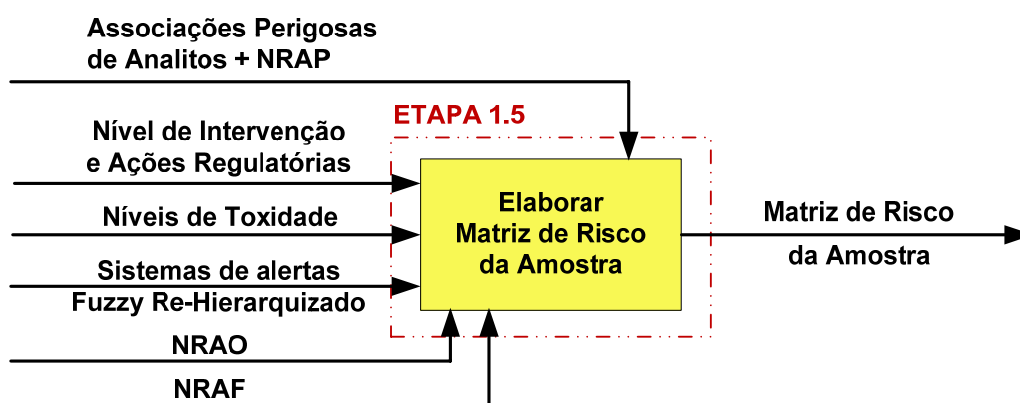


Figura 26 - Representação Gráfica da Etapa 1.5 da Plataforma base Flexível – Macro-Etapa 1

Detalhadamente estas informações compreendem

- Sistema de Alertas *Fuzzy* Re-hierarquizado:
  - Relação dos analitos pré-determinados para a matriz de leite analisada;
  - Concentrações dos Analitos detectados (CA);
  - Níveis de Risco das Concentrações dos Analitos (NRCA);
  - Tendências das Concentrações dos Analitos (TDCA);
  - Prazos Estimados de Ajustamento do sistema produtivo aos índices mundiais de segurança alimentar de acordo com as orientações dos Sistemas Autocontroles do MAPA (PZCA);
  - Custo de Ajustamento a estes índices (CTCA);
  - Níveis de Risco Agregados dos Analitos (NRAA);
- Nível de Risco da Amostra Original - NRAO;
- Níveis de Toxicidade dos Analitos Não Conformes;
- Relação das Associações Perigosas de Analitos – APA, apresentando:

- Efeito Toxicológico Aditivo – ETA;
- Efeito Toxicológico Sinérgico – ETS;
- Efeito Toxicológico de Potenciação – ETP;
- Efeito Toxicológico Redutor – ETR.
- Nível de Risco da Amostra Final – NRAF;
- Nível de Intervenção e Ações Regulatórias – NIAR.

Por meio da Matriz de Riscos, os gestores do MAPA podem avaliar o nível atual de exposição da população aos resíduos, contaminantes e suas associações perigosas, facilitando priorizar intervenções reguladoras ao processo produtivo agrícola ou retirar produtos de circulação, conforme o caso. Valendo-se das séries históricas armazenadas no banco de dados Análise de Riscos Químicos Agropecuários – ARQA, a Abordagem Chem-Risk permite que os especialistas de risco possam estender a Matriz de Riscos ao nível dos medicamentos veterinários comumente utilizados no agronegócio brasileiro.

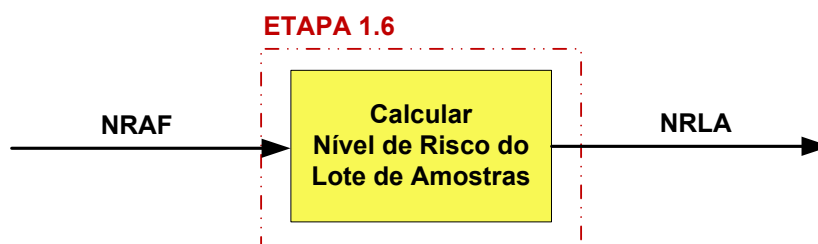
**Etapa 1.6:** De acordo com a semântica das avaliações de risco desenvolvidas pelo PNCRC, toda amostra pertence a um determinado lote de amostras. Por esta razão, após o processo de aferição das avaliações de risco ao nível dos analitos e amostras, a Plataforma Base Flexível permite aferir o Nível de Risco do Lote de Amostras (NRLA) analisado, de acordo com o seguinte método:

- 1) Caso existam amostras não conformes no lote:
  - 1.1) Calcular o total de amostras não conforme altas, ou seja,  $3 \leq \text{NRAF} < 4$ ;
  - 1.2) Calcular o total de amostras não conforme intoleráveis, ou seja,  $\text{NRAF} = 4$ ;
  - 1.3) Calcular o total de amostras não conformes (1.1 + 1.2);
  - 1.4) Calcular o NRLA de acordo com a seguinte fórmula:  

$$\text{NRLA} = ((\text{total de amostras não conforme altas} * 3) + (\text{total de amostras não conforme intoleráveis} * 4) / \text{total de amostras não conformes});$$
 Nível de intervenção = “3 - Máxima”
- 2) Caso todas as amostras do lote estejam conformes:
  - 2.1) Encontrar a média dos riscos (NRAF) das amostras conformes;
  - 2.2) Se média dos riscos  $\geq 2.7$ , ou seja, 10% próximo ao nível de risco 3 - Alto;

- Nível de Intervenção = “2 - Investigatória”;
- 2.3) Se média dos riscos < 2.7  
Nível de Intervenção = “1 - Desnecessária”.

A figura 27 ilustra graficamente a Etapa 1.6:



**Figura 27 - Representação Gráfica da Etapa 1.6 da Plataforma base Flexível – Macro-Etapa 1**

O modelo completo para estimativa dos níveis de risco ao nível dos analitos, amostras e lotes de amostras de matrizes agropecuárias, gerenciado pela Plataforma Base Flexível de Pré-processamento na Macro-Etapa 1 da Abordagem Chem-Risk é apresentado na figura 28.

Nas próximas seções, as macro-etapas seguintes à Macro-Etapa 1 que complementam os procedimentos computacionais realizados pela Abordagem Chem-Risk, serão discutidos

### **7.2.2 Macro-Etapa 2 – Seleção // Formatação de Dados**

Na segunda macro-etapa, a Plataforma Base Flexível converte o formato original dos dados armazenados no banco de dados Análise de Risco Químico Agropecuário - ARQA, adaptando-os aos formatos particulares recomendados para mineração de dados utilizando as técnicas de associação, classificação e agrupamento. Estas técnicas permitem que os especialistas do PNCRC possam refinar e estender, ainda mais, o conhecimento superior incorporado na etapa de pré-processamento por meio de ontologias e lógica *fuzzy*. Cita-se, por exemplo:

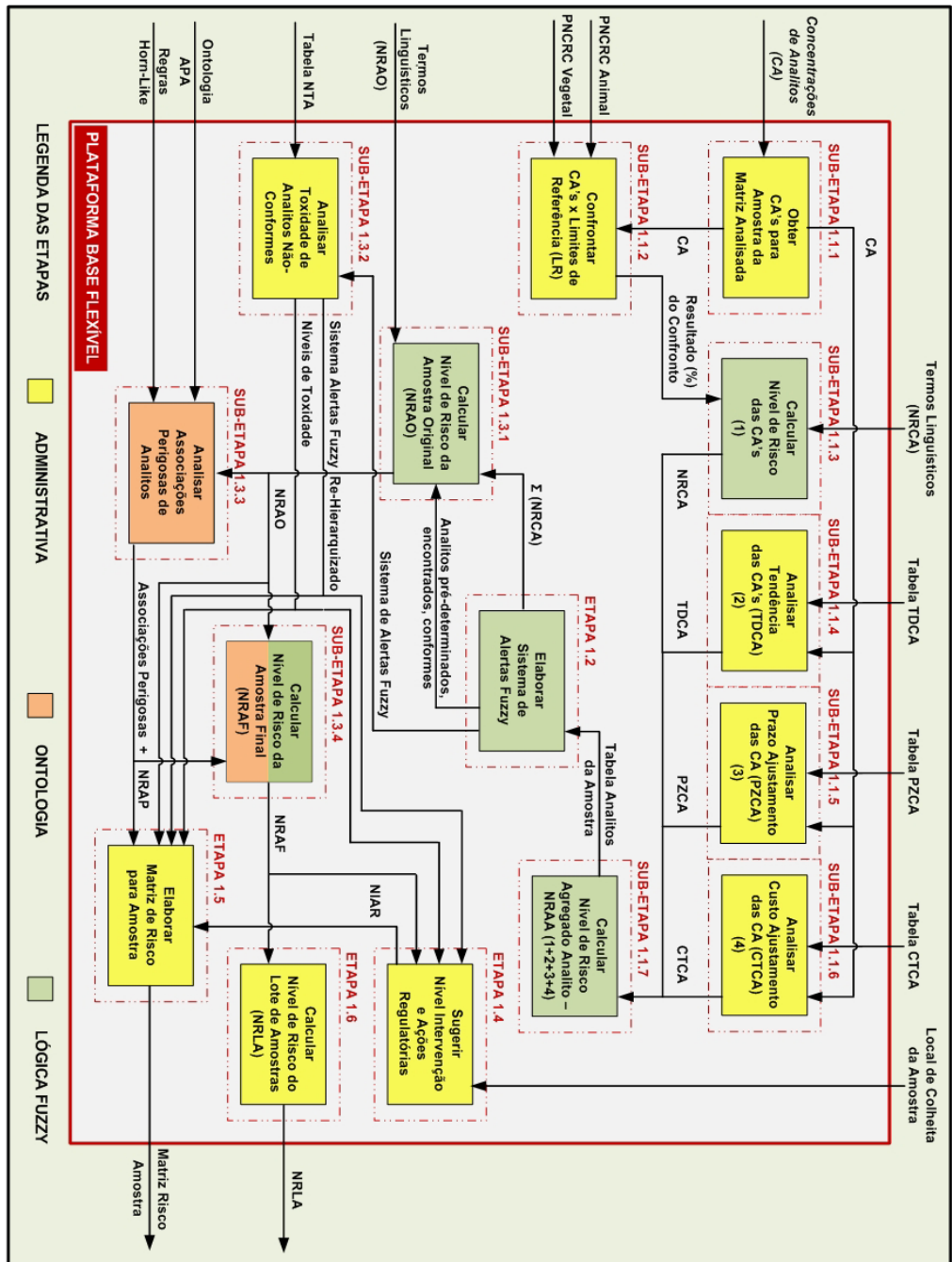


Figura 28 – Sequência de Etapas, Sub-Etapas e Fluxo das Informações Trabalhadas pela Plataforma Base Flexível para Avaliação de Riscos Químicos Agropecuários – Macro-Etapa I

- Por meio das técnicas de associação, particularmente com os algoritmos *Apriori* e *Predictive Apriori*, detectar a existência de novas associações perigosas de analitos adicionalmente àquelas modeladas por ontologias, em conjunto com os especialistas do PNCRC;

- Por meio das técnicas de agrupamento, particularmente com o algoritmo *Simple k-Means*, sugerir novas faixas de risco para analitos e amostras;
- Por meio das técnicas de classificação, particularmente com o algoritmo *J-48*, realizar uma comprovação formal dos resultados superiores alcançados pela abordagem Chem-Risk, comparativamente à abordagem tradicional comumente utilizada pelo PNCRC.

A partir da geração de dados mais exatos, precisos e fidedignos à realidade da segurança química dos alimentos brasileiros, obtidos durante a Macro-Etapa 1 de pré-processamento por meio das técnicas computacionais, obtêm-se resultados mais exatos, precisos, fidedignos à realidade e, portanto, tecnicamente mais confiáveis, durante a etapa de mineração de dados.

Os resultados aprimorados tecnicamente pela lógica *fuzzy*, aliados à viabilização de análises complexas por meio de ontologias, auxiliam aos analistas de risco e autoridades do PNCRC a alcançarem o objetivo de verdadeiramente expressarem a realidade da segurança química dos alimentos agropecuários brasileiros, em parte mascarada, devido à abordagem tradicional não considerar inúmeros fatores de risco importantes durante o processo de avaliação dos riscos químicos agropecuários.

Os resultados da etapa de mineração, em parte surpreendentes, expandem o nível de conhecimento especializado do PNCRC em direção a uma tomada de decisões não triviais e inéditas ao nível do agronegócio brasileiro.

### **7.2.3 Macro-Etapa 3 - Seleção // Formatação de Resultados**

Na terceira Macro-Etapa, a abordagem seleciona e formata os resultados armazenados no banco de dados Análise de Risco Químico Agropecuário - ARQ que foram gerados a partir da aplicação da lógica *fuzzy* e ontologias ao processo de pré-processamento realizado na Macro-Etapa 1 pela Plataforma Base Flexível. Os resultados selecionados e formatados são submetidos a um processo de visualização gráfica que facilita a interpretação dos mesmos por parte dos especialistas envolvidos com os procedimentos de avaliação dos riscos químicos agropecuários, no caso, os analistas de risco, peritos agropecuários e autoridades do PNCRC.



### **7.3 Considerações Finais**

Apesar dos benefícios proporcionados e ser largamente utilizada no contexto das políticas públicas brasileiras, a Abordagem Tradicional acarreta significativa perda de informações que podem ser contornadas por meio da aplicação das técnicas de cunho administrativo-computacional empregadas pela Abordagem Chem-Risk. Por outro lado, apesar do avanço técnico proporcionado, a Abordagem Vetor de Risco Ponderado também não se mostra capaz de reduzir eficazmente a perda de informações revelada nas seções anteriores deste capítulo. Pode-se concluir que através do ganho de informação produzido pela Abordagem Chem-Risk, adiciona-se qualidade aos dados e resultados aferidos nas avaliações de riscos químicos, afetando positivamente o processo de tomada de decisão por parte das autoridades do PNCRC / MAPA, com relação ao processo de produção e comercialização, interna e externa, dos alimentos agrícolas brasileiros.

Para melhor evidenciar o ganho de informação produzido pela Abordagem Chem-Risk, o próximo capítulo apresenta as principais contribuições e os resultados do estudo de viabilidade, obtidos a partir da aplicação da Abordagem Chem-Risk sobre a massa de exames laboratoriais em amostras de matrizes agropecuárias, animais e vegetais monitoradas pelo PNCRC / MAPA.

# Capítulo 8

## Resultados do Estudo de Viabilidade

---

### 8.1 Considerações Iniciais

Neste capítulo abordaremos em maiores detalhes as principais contribuições e os resultados do estudo de viabilidade realizados a partir da aplicação da Abordagem Chem-Risk ao Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC / Animal e Vegetal, do Ministério da Agricultura, Pecuária e Abastecimento – MAPA. As contribuições e os resultados obtidos demonstram a superioridade da Abordagem Chem-Risk, que apresenta sensíveis aprimoramentos à estrutura das avaliações dos riscos químicos dos produtos agropecuários brasileiros, considerando os limites de segurança química pré-estabelecidos mundialmente pela *Comissão do Codex Alimentarius* – FAO/OMS/ONU. Ao longo das seções, detalha-se em maior profundidade a importância da aplicação das técnicas computacionais, baseadas na lógica difusa e ontologias, à estrutura de cálculo dos riscos químicos associados à produção agropecuária brasileira, aprimorando a qualidade dos dados durante a etapa de pré-processamento. Em sequência, ilustra-se como os resultados obtidos durante a etapa de mineração de dados complementam e aprimoram os resultados qualitativamente superiores, alcançados por meio da aplicação das técnicas computacionais à etapa de pré-processamento de dados da Abordagem Chem-Risk.

### 8.2 Principais Contribuições

Em comparação com a Abordagem Tradicional, analisada ao longo do capítulo 3, e com o Vetor de Risco Ponderado, analisado no item 7.1, resume-se nesta seção as principais contribuições da Abordagem Chem-Risk em termos de medidas qualitativas e quantitativas de desempenho:

## 8.2.1 Medidas Qualitativas de Desempenho

As principais medidas de desempenho de cunho qualitativo utilizadas para avaliar o desempenho da Abordagem Chem-Risk, foram comprovadas a partir da implantação dos procedimentos computacionais realizados pela Plataforma Base Flexível (Macro-Etapa 1 - Item 7.2.1) sobre a massa de exames laboratoriais gerada pelos laboratórios credenciados do PNCRC, quais sejam:

1) Expressar, em termos de maior ou menor granularidade de informação, a atual qualidade dos produtos agropecuários consumidos internamente no país ou exportados. A abordagem Chem-Risk permite trabalhar com a menor parte da informação (ou grão) ao nível de risco dos analitos, fornecendo maior exatidão e precisão para a estimativa dos riscos associados à segurança química dos alimentos agropecuários. O grão permite que a abordagem ofereça maior escopo para as avaliações de risco, possibilitando inferir níveis de risco químico incrementais para propriedades particulares, bacias, zonas geográficas (mesorregiões), estados, regiões e país, associados com um analito particular, grupo ou classe de analitos, amostra de matriz agropecuária e também para grupo ou classe especial de amostras. Como exemplo, cita-se:

a) Analito particular detectado na amostra da matriz analisada, por exemplo: Qual o nível de risco associado ao analito Tetraciclina na carne bovina produzida no Rio Grande do Sul? E ao nível da carne bovina brasileira? E ao nível dos produtos agropecuários em geral comercializados no país?;

b) Grupo ou classe de analitos, por exemplo: Qual o nível de risco associado ao grupo dos Antimicrobianos no mel brasileiro? Qual o nível de risco associado ao grupo Carbamatos no leite produzido na Zona da Mata em Minas Gerais?;

c) Amostra de matriz analisada englobando todos os analitos detectados, por exemplo: Qual o nível de risco inferido para a amostra número xxxxxx/2008 ?;

d) Grupo ou classe especial de amostras, por exemplo: plantas submetidas a tratamento específico, tratamento homeopático ministrado em animais, animais submetidos a alimentação diferenciada, manejo diferenciado, etc.;

e) Propriedades específicas, por exemplo: Qual o nível de risco químico associado ao leite produzido na propriedade xxxx?;

f) Associações, cooperativas, empresas, etc., por exemplo: Qual o nível de risco químico associado ao leite produzido pela Embrapa Gado de Leite? E pela cooperativa Completa? Qual o nível de risco químico associado ao camarão produzido pela empresa Fish?;

g) Cidades, estados, regiões, zonas geográficas e país, por exemplo: Qual o risco químico associado à carne bovina produzida na cidade do Rio de Janeiro? E à carne bovina produzida na região sul? E à carne bovina brasileira? Qual o risco químico associado aos vegetais produzidos no país? Qual o risco químico, expresso em termos de resíduos e contaminantes, associado aos produtos agropecuários brasileiros?

2) Detectar o nível de proximidade das concentrações dos analitos (CA) aos respectivos limites de referências pré-estabelecidos por legislação nacional, convenção internacional ou, em sua ausência, na melhor evidência científica disponível. Caso determinada concentração esteja acima do seu limite de referência respectivo, o analito é classificado como não conforme ao consumo humano e animal;

3) Aprimorar a capacidade inferencial dos programas públicos, permitindo a sistemática agregação de novos fatores de risco, evitando, assim, significativa perda de informações durante a estimativa dos riscos, por não considerá-los ao nível dos analitos (a-d) e amostras (e-f). Na abordagem Chem-Risk, estes novos fatores de risco são considerados riscos químicos potenciais que individualmente afetam os resultados das avaliações de risco. No contexto desta pesquisa, são considerados os seguintes fatores de risco adicionais:

a) Nível de risco não agregado para cada concentração de analito detectado na amostra da matriz agrícola (Nível de Risco da Concentração do Analito – NRCA);

b) Tendência da concentração do analito de acordo com a experiência dos analistas de risco, peritos em produtos agropecuários e dados históricos;

c) Período de ajustamento aos índices de segurança química (Limites de Referência) recomendados mundialmente;

d) Custo de ajustamento a estes índices;

e) Nível de Toxicidade dos Analitos não conformes (NTA);

f) Nível de Risco de Associações Perigosas de Analitos (NRAP).

4) Propor níveis de risco agregados (NRAA) para cada analito encontrado na amostra analisada, a partir da incorporação dos fatores Concentração do Analito, Tendência da Concentração, Prazo de Ajustamento e Custo de Ajustamento;

5) Propor um Sistema de Alertas *Fuzzy* baseado no NRAA com capacidade para diferenciar e agrupar de forma hierárquica tanto os analitos não conformes, como os analitos que apresentem tendência de agravamento da concentração detectada;

6) Propor Nível de Risco Original da Amostra (NRAO) com capacidade para diferenciar amostras conformes e não conformes ao consumo humano e animal;

7) Propor Nível de Risco Final da Amostra (NRAF) de acordo com o maior nível de risco detectado entre NRAO e NRAP;

8) Fornecer informações para a priorização de intervenções e ações regulatórias ao sistema produtivo agropecuário ou mesmo para a retirada de produtos de circulação, considerando o local de colheita (origem) da amostra como fator de risco ao nível das amostras;

9) Estabelecer uma matriz de risco para os medicamentos veterinários detectados na amostra da matriz analisada, considerando os níveis de risco incrementais, os fatores de risco adicionais, assim como, as demais informações geradas nos itens anteriores.

## **8.2.2 Medidas Quantitativas de Desempenho**

Em adição às medidas qualitativas apresentadas no item 7.3.1, as medidas quantitativas usadas para avaliar o desempenho da Abordagem Chem-Risk foram:

- 1) Aumento do número de intervenções governamentais ao sistema brasileiro de produção agropecuária animal:
  - Intervenções máximas (12.7%);
  - Intervenções moderadas (4.6%).
- 2) Aumento do número de intervenções governamentais ao sistema brasileiro de produção agropecuária vegetal:
  - Intervenções máximas: 8.43%;
  - Intervenções moderadas: 18.1%.
- 3) Aumento dos tipos de avaliações de risco (níveis de risco incrementais);
- 4) Aumento do número de fatores de risco considerados no cálculo dos níveis de risco ao nível dos analitos, ou seja, NRCA e NRAA;

- 5) Aumento do número de fatores de risco considerados no cálculo dos níveis de risco ao nível das amostras, ou seja, NRAO, NTA e NRAP;
- 6) Redução dos custos anuais das avaliações de risco realizadas pelo PNCRC/MAPA:
  - o Salário de consultores externos: 42,2%;
  - o Passagens e despesas diárias: 73,9%.

### 8.3 Análise da Incorporação da Lógica *Fuzzy*

No Capítulo 4 – Lógica Nebulosa, na seção 4.4 – Considerações Finais, afirmamos que: “... é possível deduzir que a lógica *fuzzy*, especificamente considerando os números e conjuntos *fuzzy* apresentados, ao oferecer um tratamento para imprecisão e incerteza, contribui, decisivamente, para que a abordagem Chem-Risk alcance maior exatidão, precisão e, conseqüentemente, consiga expressar de forma mais fidedigna a realidade da segurança química dos produtos agropecuários brasileiros, reduzindo, assim, a perda de informações valiosas que comprometem a qualidade da tomada de decisão realizada pelas autoridades do PNCRC/MAPA”. Por meio de exemplos e considerações técnicas, a seguir, demonstra-se os benefícios alcançados, a curto e longo prazo, pela incorporação da lógica *fuzzy* à Macro-Etapa de Pré-processamento de Dados, gerenciada pela Plataforma Base Flexível.

#### 8.3.1 Estimativa do Nível de Risco da Concentração do Analito - NRCA

A título de comparação, vamos utilizar o esquema de avaliação do NRCA na Sub-Etapa 1.1.3 da Plataforma Base Flexível e observar os métodos utilizados na Abordagem Tradicional e Vetor de Risco Ponderado. A figura 29 ilustra individualmente a Sub-Etapa 1.1.3:

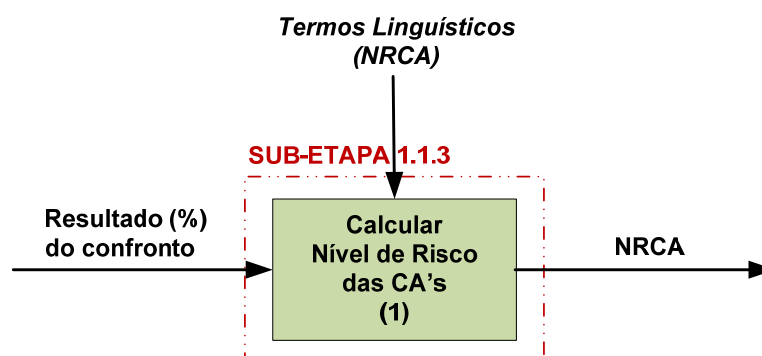
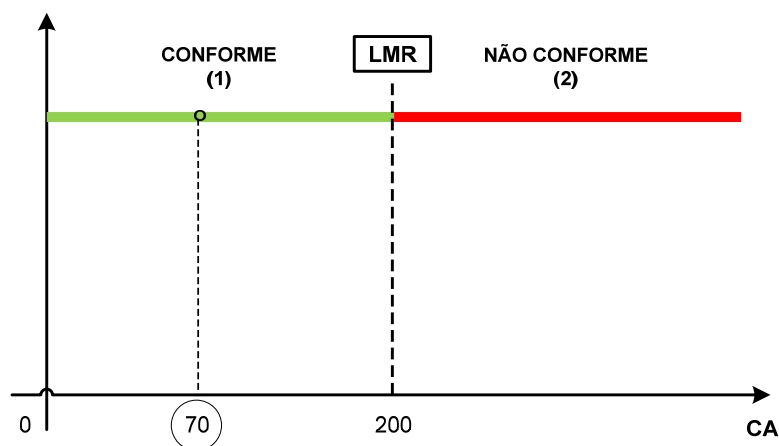


Figura 29 - Representação Gráfica da Sub-Etapa 1.1.3 da Abordagem Chem-Risk

Suponha que a análise laboratorial tenha sido realizada em uma amostra da matriz “Carne Bovina – Músculo” e tenha revelado uma concentração (CA) de 70 µg/kg do analito Clortetraciclina, para um Limite de Referência (LR) pré-estabelecido mundialmente em 200 µg/kg e assinalado no PNCRC / Animal / 2010 - Uma vez que este analito é um tipo de resíduo, o Limite de Referência (LR) representa o seu Limite Máximo de Resíduo – LMR. Esta concentração, juntamente com as outras concentrações detectadas para os demais analitos encontrados na amostra, irá compor a tabela de Concentrações de Analitos (Sub-Etapa 1.1.1 – Tabela 2) da Plataforma Base Flexível, e deverá ser confrontada com o limite máximo admitido para este resíduo (LMR) na Sub-Etapa 1.1.2.

A partir deste ponto, as abordagens diferem significativamente:

- Utilizando um único limite rígido (*crisp*), ou seja, o Limite de Referência (LR), a Abordagem Tradicional avalia diretamente a concentração detectada classificando-a como “Conforme” ou “Não Conforme”, caso a concentração encontre-se dentro do limite pré-estabelecido ou supere este limite. A figura 30 ilustra graficamente a Avaliação da Concentração do Analito (ACA) de 70 µg/kg, sugerida pela Abordagem Tradicional como “1- Conforme”:



**Figura 30 – Avaliação da Concentração do Analito (ACA) Clortetraciclina - Abordagem Tradicional**

- O Vetor de Risco Ponderado também sugere limites rígidos (*crisp*) para as faixas de risco da concentração “1-Baixo”, “2-Médio” e “3-Alto” que estabelece. O Resultado

Percentual (RP) expressa o resultado do confronto CA x LMR que, neste exemplo, situa-se em 35% do Limite de Referência (LR). Com base no Resultado Percentual inferido, a abordagem classifica o Nível de Risco da Concentração (NRC) do analito Clortetraciclina como “2- Médio”. A figura 31 ilustra graficamente o método utilizado:

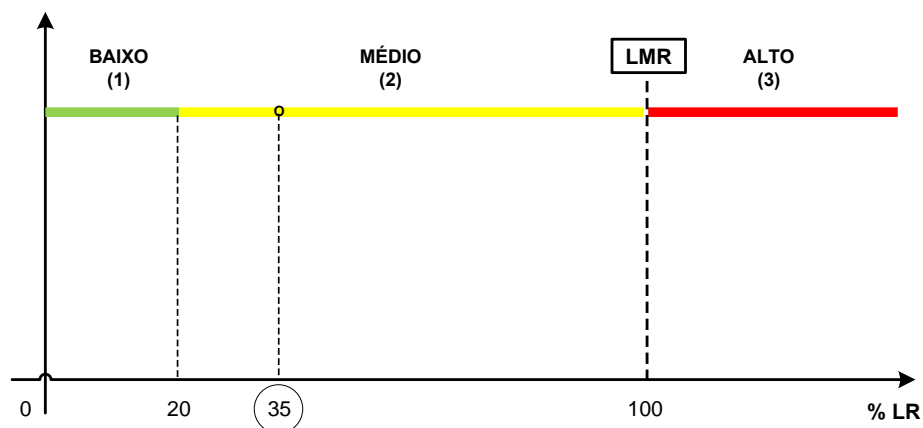


Figura 31 – Associação do Resultado Percentual (RP) aos Conjuntos *crisp* para Inferência do Nível de Risco da Concentração (NRC) - Abordagem Vetor de Risco Ponderado

### Deficiências Ocultas dos Cálculos Baseados na Teoria Clássica dos Conjuntos

Perante o contexto em que se insere, a aparência correta dos cálculos das avaliações de risco realizadas pelas abordagens Tradicional e Vetor de Risco Ponderado, baseadas na teoria clássica de conjuntos (*crisp*), oculta deficiências importantes que comprometem a semântica das avaliações de risco. Por exemplo, considerando a abordagem Vetor de Risco Ponderado, observa-se que:

- De acordo com a Faixa de Risco “2 - Médio”, todas as CA’s situadas no intervalo  $10 \mu\text{g}/\text{kg} \leq \text{CA} \leq 200 \mu\text{g}/\text{kg}$  apresentam o mesmo nível de risco químico. Apesar de ser uma situação inteligível que representa uma faixa das concentrações permitidas para o analito Clortetraciclina em uma amostra de “Músculo de Carne Bovina”, a razão nos diz que “uma concentração de  $14 \mu\text{g}/\text{kg}$  apresenta um risco médio menor que uma concentração de  $199 \mu\text{g}/\text{kg}$  (um risco quase alto)”. Por outro lado, pode-se afirmar que “uma concentração de  $199 \mu\text{g}/\text{kg}$  apresenta um pouco do risco médio e um pouco do risco alto”



- Estas abordagens perdem informações importantes por utilizarem limites rígidos em suas avaliações.
- Observe ainda que, considerando somente a presença do analito “Clortetraciclina” nas amostras de músculo de carne bovina analisadas, o nível de risco “2 - Médio” atribuído a uma amostra apresentando uma concentração de 10 µg/kg seria o mesmo para outra amostra apresentando uma concentração de 200 µg/kg, ou seja, com 2.000% a mais de Clortetraciclina.

É exatamente neste ponto que a teoria de conjuntos nebulosos, apresentando um tratamento para a imprecisão e incerteza, contribui, decisivamente, para que a abordagem Chem-Risk alcance uma exatidão e precisão superiores às abordagens Tradicional e Vetor de Risco Ponderado. A exatidão e precisão superior dos cálculos das avaliações de risco aliadas às faixas de risco adicionais, representadas por conjuntos *fuzzy* elaborados conjuntamente com especialistas, expressam de forma mais fidedigna a realidade das avaliações químicas dos produtos agropecuários brasileiros, consumidos internamente ou exportados.

Vimos na Sub-Etapa 1.1.3 da Plataforma Base Flexível da Abordagem Chem-Risk que o Resultado Percentual (RP) aplicado às funções de pertinência apresenta os seguintes graus de pertinência aos conjuntos *fuzzy* que representam as faixas de risco:

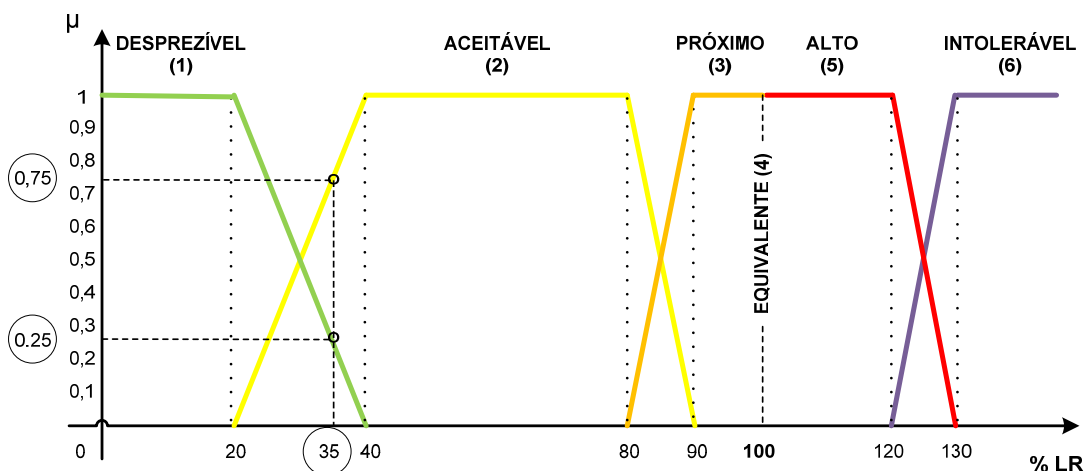
- $\mu(\text{Desprezível}) = (40 - 35) / (40 - 20) = 0.25$
- $\mu(\text{Aceitável}) = (35 - 20) / (40 - 20) = 0.75$
- $\mu(\text{Próximo}) = \mu(\text{Equivalente}) = \mu(\text{Alto}) = \mu(\text{Intolerável}) = 0$

Logo,

o NRCA inferido para o analito Clortetraciclina apresentando concentração de 70 µg/kg em uma amostra de músculo de carne bovina, foi de:

$$(0.25 \times 1) + (0.75 \times 2) + 0 + 0 + 0 + 0 = 1.75$$

Visando manter a sequência de raciocínio utilizada para comparação das abordagens, reapresenta-se a figura 18 da Sub-Etapa 1.1.3 ilustrando o mapeamento do Resultado Percentual (RP) aos conjuntos *fuzzy*:



**Figura 18 – Mapeamento Gráfico do Resultado Percentual (RP=35%) aos Conjuntos Fuzzy**

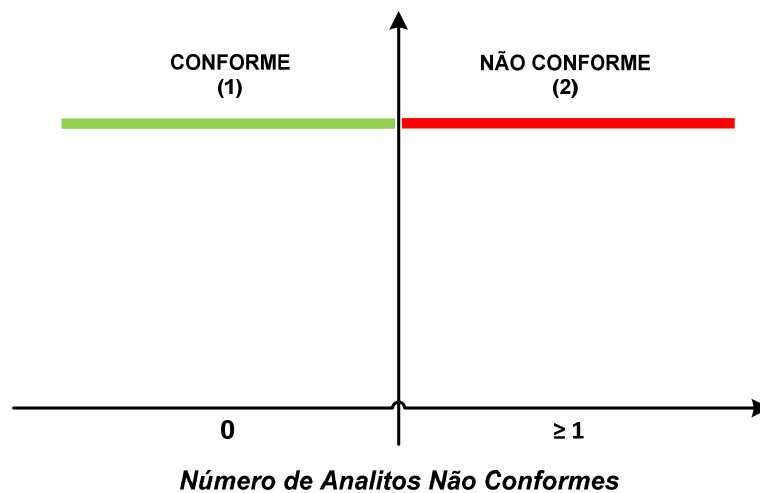
Deve ser ressaltado que o NRCA, ao integrar a composição do Nível de Risco Agregado do Analito (NRAA), contribui para que este último também alcance exatidão e precisão superior.

Observe que as “deficiências ocultas” dos cálculos realizados pelas abordagens Tradicional e Vetor de Risco Ponderado, citados anteriormente nesta seção, encontram-se devidamente tratadas pela Abordagem Chem-Risk.

### 8.3.2 Estimativa do Nível de Risco da Amostra Original - NRAO

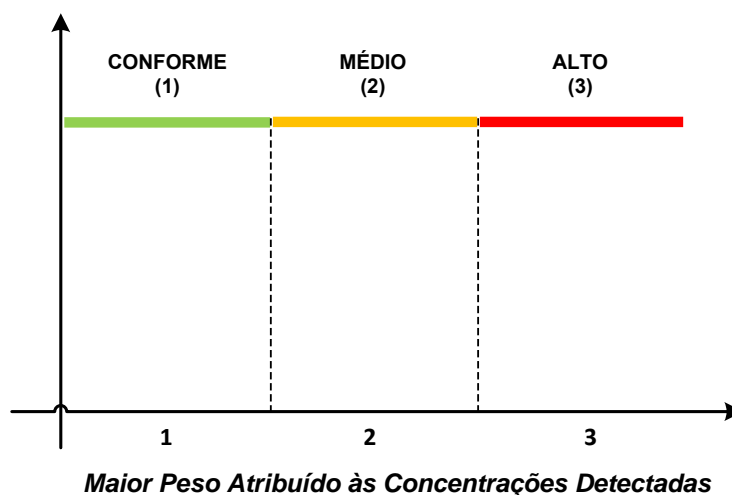
De forma semelhante aos cálculos do NRCA, nesta seção vamos comparar os métodos utilizados para cálculo do nível de risco da amostra nas abordagens Tradicional, Vetor de Risco Ponderado e Chem-Risk. Deve ser ressaltado que, de acordo com a semântica das avaliações de risco realizadas ao nível das amostras, basta a existência de um único analito não conforme, detectado entre os analitos pré-determinados para a matriz, para que a amostra inteira seja classificada como “não conforme” ao consumo humano e animal. Este comportamento é observado em todos os métodos analisados a seguir:

Na Abordagem Tradicional os especialistas de risco classificam a Avaliação da Amostra (AA) utilizando os limites rígidos “1 - Conforme” ou “2 - Não Conforme” em função da existência ou não de analitos não conformes detectados. A figura 32 ilustra o método utilizado:



**Figura 32 – Método para Avaliação de Amostra (AA) - Abordagem Tradicional**

Na Abordagem Vetor de Risco Ponderado os especialistas de risco classificam o Nível de Risco da Amostra (NRA) utilizando limites rígidos para as seguintes faixas de risco: “1 - Conforme”, caso não sejam detectados analitos não conformes, “2 – Médio” ou “3 – Alto” em função do maior peso atribuído ao potencial maligno dos analitos não conformes detectados na amostra. A figura 33 ilustra o método para avaliação do NRA:

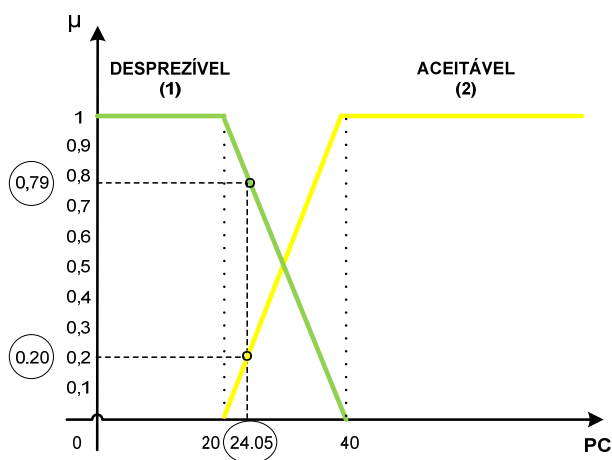


**Figura 33 – Método para Avaliação do NRA - Abordagem Vetor de Risco Ponderado**

O raciocínio utilizado pela Abordagem Chem-Risk para estimativa do NRCA pode ser estendido para estimativa do primeiro nível de risco inferido para as amostras analisadas, ou seja, o NRAO, visando que este nível de risco também esteja mais exato, preciso e fidedigno.

Para uma amostra conforme, a qual apresenta todos os analitos encontrados em conformidade com seus limites de referência respectivos:

- A Abordagem Tradicional classifica quimicamente a AA como “1 – Conforme”;
- A Abordagem Vetor de Risco Ponderado classifica o NRA como “1 – Conforme”;
- Mantendo a sequência de raciocínio utilizada para comparação das abordagens, na Sub-Etapa 1.3.1 da Plataforma Base Flexível da Abordagem Chem-Risk, supomos que foram detectados 31 analitos conformes entre 43 pesquisados em uma amostra de leite, traduzindo um  $\Sigma$ -NRCA de 53.37. Logo, o Percentual de Conformidade (PC) da amostra de 24.05%, aplicado às funções de pertinência aos conjuntos *fuzzy*, determinou um NRAO de 1.19, ilustrado na rerepresentação da figura 22:

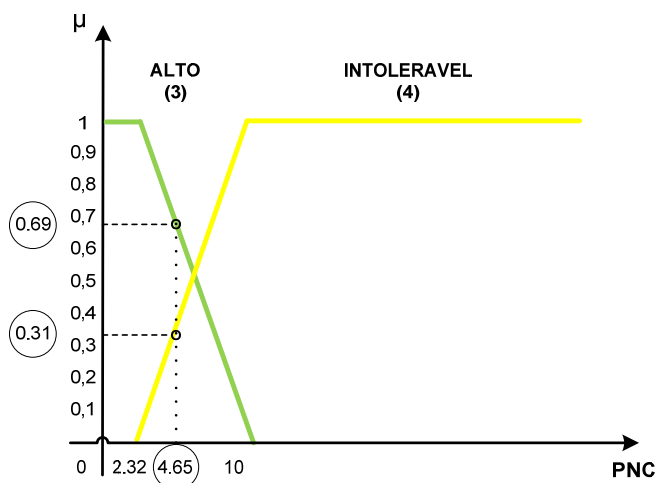


**Figura 22 – Mapeamento Gráfico do Percentual de Conformidade (PC=24.05%) aos Conjuntos *Fuzzy* Pré-estabelecidos para Faixas de Risco de Amostras Conformes**

Para uma amostra “não conforme”, ou seja, apresentando um ou mais analitos em não conformidade aos seus limites de referência respectivos:

- A Abordagem Tradicional classifica a AA como “2 – Não Conforme”;

- A Abordagem Vetor de Risco Ponderado classifica o NRA como “2 – Médio” ou “3 – Alto” de acordo com o maior potencial maligno atribuído aos analitos não conformes detectados na amostra analisada;
- Mantendo a sequência de raciocínio utilizada para comparação das abordagens, na Sub-Etapa 1.3.1 da Plataforma Base Flexível da Abordagem Chem-Risk, supomos que o Sistema de Alertas *Fuzzy* destacou a existência de 2 analitos não conformes, entre 43 analitos pesquisados na amostra de leite. Logo, o Percentual de Não Conformidade (PNC) estabelecido foi de 4.65%, cujo mapeamento às funções de pertinência dos conjuntos *fuzzy*, revelou um NRAO de 3.31, conforme ilustrado na rerepresentação da figura 24:



**Figura 24 – Mapeamento Gráfico do Percentual de Não Conformidade (PNC=4.65%) aos Conjuntos *Fuzzy* Pré-estabelecidos para Faixas de Risco de Amostras Não Conformes**

### 8.3.3 Conclusões Acerca da Incorporação da Lógica *Fuzzy*

Por meio das análises comparativas, acima demonstradas, conclui-se que os limites rígidos inerentes a Teoria Clássica dos Conjuntos (ou *crisp*) não se mostram adequados para representar corretamente a semântica das avaliações de riscos químicos agropecuários, tanto ao nível dos analitos como também das amostras analisadas, permitindo que concentrações 2.000% superiores apresentem o mesmo nível de risco químico. Além disto, não se mostram capazes de expressar termos comumente utilizados na linguagem dos especialistas de risco, tais como, “próximo a”, “parte de”, “quase”, etc., que são inerentes à semântica do domínio avaliações de riscos químicos agropecuários. Em consequência, os limites rígidos também acarretam a perda de importantes

informações que afetam negativamente a qualidade dos dados trabalhados e armazenados, a qualidade dos resultados da etapa de mineração e, por consequência, a qualidade dos resultados do processo de descoberta de conhecimento.

## **8.4 Análise da Incorporação da Ontologia APA**

Na Sub-Etapa 1.3.3 da Plataforma base Flexível da Abordagem Chem-Risk afirmamos que “...os especialistas necessitam considerar a existência do fator de risco agregado “Associações Perigosas de Analitos - APA”, evitando, assim, acumular nova perda de informações importantes no cálculo do nível de risco das amostras analisadas, às perdas inicialmente detectadas ao nível dos analitos e amostras pela aplicação da lógica *fuzzy*. Este fator de risco “agregado” engloba a existência de fatores de risco individuais que relatam a interação entre dois ou mais agentes intoxicantes (analitos) que causam danos ao sistema biológico humano e animal, seja alterando uma função vital ou mesmo levando à morte, sob certas circunstâncias...”

Ressalta-se, também, a existência de elevados custos inerentes ao processo de avaliação dos riscos químicos agropecuários, desenvolvido pelo PNCRC em nível nacional, relacionados, principalmente com as altas tecnologias empregadas, o tempo necessário para detecção das concentrações, as coletas diversificadas do plano de amostragem, os salários e viagens constantes de autoridades, técnicos e consultores. Devido aos elevados custos, existe um certo grau de rotatividade entre analistas de risco, estatísticos, técnicos agrícolas e consultores, durante a execução dos métodos de avaliação utilizados na Abordagem Tradicional, refletindo a ocorrência de subjetividade e erros causados por avaliação humana.

Contudo, a incorporação dos novos fatores de risco ao nível das amostras, os quais individualmente compõem o nível de risco agregado Associações Perigosas de Analitos – APA, exige novos tipos de análises complexas e dispendiosas no âmbito do PNCRC.

De acordo com o acima exposto, os principais objetivos que nortearam a aplicação de ontologias pela Abordagem Chem-Risk foram:

- “Aprender” com exatidão o conceito “Associações Perigosas de Analitos”
- Viabilizar a análise das Associações Perigosas de Analitos no âmbito do PNCRC;
- Modelar individualmente a semântica dos fatores de risco Efeito Toxicológico Aditivo (ETA), Efeito Toxicológico Sinérgico (ETS), Efeito Toxicológico de Potenciação

(ETP) e Efeito Toxicológico Redutor (ETR), que podem ser encontrados durante a análise da amostra;

- Como pré-condição para pesquisa e recuperação semântica de dados, responder corretamente às consultas da Abordagem Chem-Risk sobre a existência ou não de uma ou mais associações perigosas de analitos nas amostras analisadas, consideradas como fatores de risco individuais de acordo com a tabela 15 citada na Sub-Etapa 1.3.3, relacionando as associações detectadas e o Nível de Risco da Associação Perigosa (NRAP);
- Reduzir a forte dependência dos profissionais envolvidos;
- Evitar nova perda de informações no cálculo do nível de risco das amostras, de forma acumulativa às perdas detectadas pela aplicação da lógica *fuzzy* ao nível dos analitos e amostras.

#### **8.4.1 Engenharia da Ontologia APA: a Abordagem ACATO**

Entre as diversas abordagens para desenvolvimento de ontologias, citadas na seção 5.4, a abordagem ACATO foi desenvolvida para a engenharia da Ontologia APA. A abordagem é considerada uma variante do método UPON por incorporar as técnicas do projeto Open Mind Common Sense (OMCR-Br) e os resultados da etapa de mineração de dados da abordagem Chem-Risk, em particular, com relação às técnicas de associação, citadas na seção 8.6. O acrônimo ACATO refere-se à *Abordagem Colaborativa para Avaliação de Termos Ontológicos*.

Embora todo o processo da modelagem semântica tenha sido inicialmente definido em conformidade com a experiência e orientação dos especialistas do PNCRC e da Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA, devido a abrangência necessária demandada por essa ontologia, em um primeiro momento (Estágio I), o projeto OMCS-Br foi utilizado para:

- Explorar a variedade do vocabulário adicionalmente pronunciado por outros especialistas da área;
- Validar o vocabulário, as associações perigosas e os novos processos de avaliação de riscos químicos utilizados pelo PNCRC.

Os resultados alcançados no Estágio I possibilitaram a produção de um léxico adicional, denominado *Léxico OCMS-Br*. Em um segundo momento (Estágio II), os resultados alcançados a partir da aplicação das técnicas de associação e associação preditiva sobre os resultados das

diversas análises laboratoriais e avaliações de risco armazenadas no banco de dados Análise de Risco Químico Agropecuário – ARQA, possibilitaram a produção de um segundo léxico adicional, denominado *Léxico dos Dados*. Após a necessária validação pelos especialistas do domínio, os termos e associações perigosas de analitos eleitos a partir de exame minucioso nos léxicos adicionais, foram agregados aos Léxicos da Aplicação (LA) e do Domínio (LD), tradicionalmente utilizados pelo método UPON para composição do Léxico de Referência (LR).

O detalhamento do processo de engenharia utilizado pela abordagem ACATO para criação da ontologia Associações Perigosas de Analitos – APA é apresentada a seguir.

Os resultados do primeiro ciclo, que produziram a versão maior da ontologia APA, foram:

## **I – Fase de Concepção**

As primeiras iterações pertencem à fase de concepção da ontologia e concentram-se, basicamente, na captura dos requisitos necessários à ontologia.

### **1) Fluxo de Trabalho – Requisitos**

Em comum acordo com os especialistas do PNCRC e EMBRAPA, aqui entendidos como especialistas do domínio Associações Perigosas de Analitos e também como usuários finais da ontologia, os requisitos levantados foram:

**1.1) Domínio de Interesse:** O fragmento da realidade a ser modelada são os diversos analitos, agentes intoxicantes ou entidades químicas em associação perigosa capaz de causar dano ao sistema biológico humano, animal ou vegetal, alterando uma função ou mesmo levando-o à morte, sob certas condições de exposição;

**1.2) Escopo:** As associações perigosas que envolvem os analitos relacionados no Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC, animal e vegetal, do Ministério da Agricultura, Pecuária e Abastecimento – MAPA;

**1.3) Objetivo do negócio:** Viabilizar a análise das associações perigosas de analitos passíveis de serem encontradas nas amostras das matrizes agropecuárias animais e vegetais, monitoradas no âmbito do PNCRC. As análises podem envolver ou não os fatores de risco ao nível dos analitos e amostras, citados na seção 7.2.1;



**1.4) Termos utilizados no âmbito do PNCRC e Codex Alimentarius:** O primeiro conjunto de termos, abaixo citado, representa uma primeira parte do vocabulário controlado que compõe o Léxico da Aplicação (LA) da ontologia APA. Este conjunto inicial de termos foi obtido a partir do conhecimento específico dos especialistas do PNCRC e EMBRAPA, por meio de entrevistas e dos “storyboards” (seqüência das atividades que acontecem em um cenário específico), relacionados por estes especialistas, que abordam o processo de análise de riscos químicos em amostras de matrizes dos produtos agropecuários. Cita-se, a seguir, uma parte dos termos inicialmente fornecidos que compõem o Léxico da Aplicação da ontologia APA:

*Analito, resíduo, contaminante, grupo de analitos, amostra laboratorial, matriz agropecuária, produto agropecuário, limite de referência, limite máximo de resíduo, teor máximo de contaminante, risco, risco químico, análise de risco, gerência de risco, controle de risco, avaliação de riscos químicos, fator de risco, fator de risco químico, matriz de risco, pesticida, agrotóxico, metais pesados, nitratos, organofosforados, organoclorados, antiparasitários, contaminantes orgânicos, contaminantes inorgânicos, antimicrobianos, carbamatos, anticoccidianos, betagonistas, anti-inflamatórios, piretróides, sedativos, corantes, HPA, HCH, PCB, DDE, DDT, DDD, toxinas, etc.*

**1.5) Associações perigosas envolvendo os analitos monitorados pelo PNCRC e Codex Alimentarius:** Adicionalmente aos termos citados no item 1.4, o primeiro conjunto das associações perigosas de analitos, abaixo relacionadas, compõem uma outra parte do vocabulário controlado pela ontologia APA, inicialmente inserido no banco de dados OCMS-Br para validação das associações inseridas e sugestão de novas associações perigosas, por especialistas externos. As associações perigosas relacionadas são aquelas detectadas em uma mesma amostra de matriz agropecuária. Inicialmente foram inseridas um total de 128 associações, entre as quais, cita-se:

**1.5.1)** Associações perigosas envolvendo dois ou mais analitos pertencentes ao grupo de analitos organofosforados, particularmente por meio de efeitos tóxicos aditivos e algumas vezes sinérgicos, ocasionando problemas graves no sistema nervoso central, inibindo a atividade muscular voluntária, etc. A título de exemplo, para um total de 55 associações perigosas modeladas para este grupo, cita-se:

Paration e Acefato; Paration e Metidation; Mevinfós e Diazinon; Azinfós Etil e Azinfós Metil; Metamidofós e Pirimifós Metil; etc;

**1.5.2)** Associações perigosas envolvendo os grupos de analitos organofosforados e organoclorados, apresentando características antagônicas que podem mascarar o nível real da concentração de determinados analitos detectados na amostra, caso o outro analito não estivesse presente na amostra. Para um total de 24 associações perigosas envolvendo os grupos, cita-se: Aldrin e Paration; Dieldrin e pp-DDT(h); Eldrin e PCB 101; Azinfós Etil e PBC 180; Trans Clordane e Acefato; Cis Clordane e Clorpirifós Etil, etc;

**1.5.3)** Associações perigosas envolvendo tipos sinérgicos de aflatoxinas, detectados principalmente no amendoim, milho e feijão, apresentando alto poder cancerígeno. Para um total de 19 associações perigosas, cita-se: aflatoxina B1 e aflatoxina G1 para a matriz arroz, castanha-do-brasil e feijão, aflatoxina G1 e aflatoxina G2 para amendoim, aflatoxina B2 e aflatoxina G2 para o milho, etc.

**1.5.4)** Metais pesados, como por exemplo, a associação entre chumbo e arsênico gerando aumento aditivo de coproporfirinogênio urinário, zinco e cádmio, arsênico e selênio;

Etc;

**1.6) Questões de Competência (QC):** As questões solicitadas aos especialistas do domínio, orientadas à descoberta de recursos e recuperação da informação semântica que a ontologia deve ser capaz de realizar, são apresentadas a seguir. Mais à frente, estas questões serão validadas no fluxo de trabalho de teste para avaliar os resultados alcançados, a cobertura da ontologia e os detalhes modelados:

**1.6.1) (QC1):** Quais são as atividades principais que antecedem e participam diretamente do processo de avaliação da existência de associações perigosas de analitos em uma amostra de matriz agropecuária monitorada pelo PNCRC ?

*As atividades principais compreendem:*

*1.6.1.1) A detecção dos analitos e concentrações (CA) existentes em uma amostra de matriz agropecuária. Esta detecção é realizada exclusivamente pelos laboratórios credenciados do PNCRC e discriminadas no COA (Certificado Oficial de Análise) emitido por estes laboratórios;*

*1.6.1.2) A avaliação das tendências das concentrações realizada pelos especialistas do PNCRC;*

*1.6.1.3) As avaliações de risco realizadas ao nível de cada analito detectado na amostra da matriz analisada e ao primeiro nível de risco aferido para a amostra (NRAO);*

**1.6.2) (QC2):** Quais as atividades principais que compõem o processo de avaliação de associações perigosas de analitos em uma amostra de matriz agropecuária?

*A partir do armazenamento dos dados citados na questão anterior, a abordagem Chem-Risk deverá ser capaz de detectar cada possível tipo de interação entre agentes tóxicos existente na amostra, que ofereça riscos à saúde humana, animal e vegetal, sugerindo um nível de risco para cada associação detectada, de acordo com o poder de toxicidade da associação;*

**1.6.3) (QC3):** Em que se aplicam os resultados das avaliações de risco aferidas a partir das associações detectadas em uma amostra? *Estas informações deverão influenciar o Nível de Intervenção e as Ações Regulatórias ao processo produtivo do alimento, estabelecido pelo PNCRC e estarem discriminadas na matriz de risco da amostra analisada;*

**1.6.4) (QC4):** Quais associações perigosas deverão fazer parte da ontologia ? *Inicialmente todas as associações relacionadas pelos especialistas do PNCRC / EMBRAPA, e, posteriormente, aquelas sugeridas por especialistas externos;*

**1.6.5) (QC5):** Quais os tipos de associações perigosas de analitos que devem ser analisadas em uma amostra ? *As interações dos agentes tóxicos, aqui entendidos como analitos, são aquelas que causam algum tipo de enfermidade ou a falência completa de alguma função vital do organismo humano, animal ou vegetal. As interações ditas benéficas entre uma ou mais substâncias não necessitam ser analisadas por esta aplicação. As interações tóxicas de interesse imediato são: aditivas, sinérgicas, potenciação e antagonismo;*

**1.6.6) (QC6):** Quem deverá validar as associações sugeridas pelos especialistas externos? *Toda associação perigosa sugerida ou mesmo validada por especialistas externos ao PNCRC deve ser, obrigatoriamente, validada pelos especialistas do PNCRC e EMBRAPA;*

**1.6.7) (QC7):** Quais documentos deverão ser utilizados como apoio à execução destas atividades? *Inicialmente a tabela de associações perigosas de analitos (ver tabela 15, sub-etapa 1.3.3) elaborada por estes especialistas e, posteriormente, as associações constantes em documentos possivelmente sugeridos por especialistas externos;*

**1.6.8) (QC8):** Quais são as substâncias que devem ser monitoradas nas associações perigosas controladas pelo PNCRC? *Todos os analitos pré-determinados pelo PNCRC e Codex Alimentarius, especificados no PNCRC Animal e Vegetal, são passíveis de monitoramento;*

**1.6.9) (QC9):** Observa-se que a grande maioria das associações perigosas envolvem dois ou mais analitos, no caso, diretamente especificados pelo PNCRC Animal e Vegetal. Além dos analitos, as associações perigosas devem adicionalmente envolver algum outro tipo de informação? *Existem alguns poucos casos em que as interações tóxicas de analitos dizem respeito a uma matriz particular, por exemplo, envolvendo Aflatoxinas\_M1 monitoradas somente para a matriz leite ou o corante Verde Malaquita monitorado somente para uma matriz de peixe cultivado ou camarão. Entretanto, existem situações em que a tendência da concentração é importante, facilitando, em muitos casos, antecipar uma intervenção ao processo produtivo que está sendo desenvolvido pelo produtor ou instituição;*

**1.6.10) (QC10):** Quem deverá realizar as avaliações sobre a existência de associações perigosas em uma amostra? *Todo tipo de avaliação de risco, ou seja, as estimativas de risco ao nível dos analitos, amostra (inclusive associações perigosas) e lote de amostras, são realizadas unicamente pelos especialistas de risco do PNCRC.*

**1.6.11) (QC11):** Em suma, quais são os profissionais envolvidos neste processo de avaliação da existência de associações perigosas em uma amostra? *Conforme explicado anteriormente, os especialistas do domínio (peritos do PNCRC, EMBRAPA e eventualmente consultores externos) são responsáveis tecnicamente por relacionar as associações perigosas a serem monitoradas, bem como, validar aquelas outras sugeridas por especialistas externos. Os especialistas ou analistas de risco possuem maiores conhecimentos para desenvolver e documentar os riscos aferidos nos processos das avaliações de risco atualmente adotados pelo PNCRC / MAPA;*

**1.7) Casos de Uso:** Os casos de uso servem como base para especificar a utilização prevista para a ontologia Associações Perigosas de Analitos – APA. A título de exemplo, cita-se, a seguir, o caso de uso respectivo à questão de competência (QC1) ilustrado na figura 34:

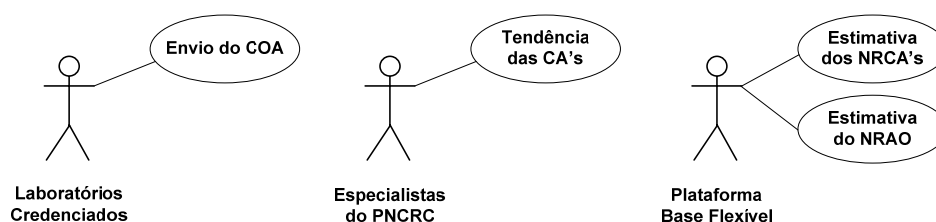
---

**Questão de Competência**

*(QC1) : Quais são as atividades principais que antecedem e participam diretamente do processo de avaliação da existência de associações perigosas de analitos em uma amostra de matriz agropecuária monitorada pelo PNCRC ?*

---

**Modelo de caso de uso**



**Figura 34 – Questão de Competência (QC1) e Caso de Uso Correspondente**

## 2) Fluxo de Trabalho – Análise

Neste fluxo refinamos e estruturamos os requisitos da ontologia APA identificados no fluxo de trabalho Requisitos:

**2.1) Glossário de Referência (GR):** De acordo com o método UPON, o Léxico da Aplicação (LA) é enriquecido através de um léxico de domínio mais geral ou Léxico do Domínio (LD) visando a construção de um Léxico de Referência (LR) o qual, posteriormente, será usado para criar o Glossário de Referência (GR) que, adicionalmente, incorpora as definições dos termos utilizados no Léxico de Referência. As etapas utilizadas para estruturação dos léxicos utilizados pela abordagem ACATO, são apresentadas a seguir:

**2.1.1) Léxico do Domínio (LD) :** O Léxico do Domínio (LD) foi construído a partir da colheita da terminologia adicional utilizada no domínio de interesse, desta vez,

extraídas principalmente por meio da análise documental dos recursos existentes e sugeridos pelos especialistas do domínio, quais sejam:

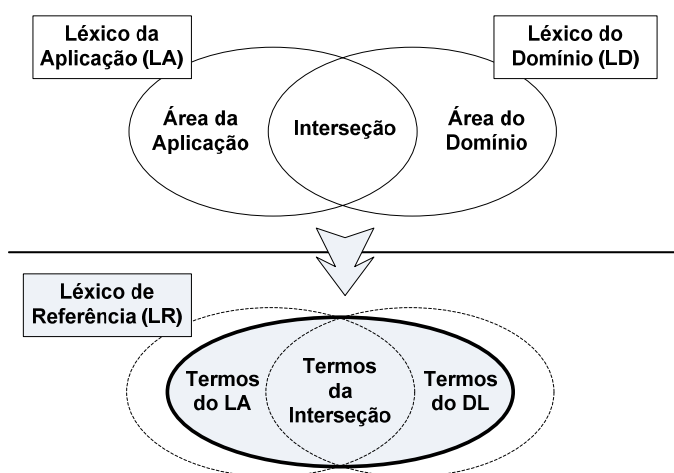
- Documentos técnicos gentilmente oferecidos:
  - Instrução Normativa Nº 21 de 02 de Setembro de 2010, promulgando o PNCRC Vegetal 2010/2011 – MAPA, SDA;
  - Instrução Normativa Nº 08 de 29 de Abril de 2010, promulgando o PNCRC Animal 2010 – MAPA, SDA;
- Certificado Oficial de Análise (COA) contendo, entre outras informações:
  - Produto agropecuário analisado
  - Matriz agropecuária analisada;
  - Identificação da amostra analisada;
  - Relação de analitos analisados;
  - Relação dos analitos detectados;
  - Concentrações dos analitos (CA) detectados;
  - Dados administrativos: data recebimento, data análise, operador, etc;
- Manuais de procedimentos técnicos do PNCRC para laboratórios credenciados – Área Animal, Versão 1, Dezembro/2008 – Abordando: recebimento de amostras, embalagens, validação, emissão do certificado oficial de análise, rejeição, violação, descrição dos resultados, etc;
- Apresentação do PNCRC para o leite durante visita técnica do coordenador do PNCRC – Animal à Embrapa Gado de Leite;
- Reunião nacional de gestores regionais do PNCRC / 2010;
- Informações e links disponibilizados oficialmente a partir do portal <http://www.agricultura.gov.br> por autoridades do MAPA, SDA e PNCRC.

Exemplifica-se, a seguir, uma parte dos termos extraídos:

*Micotoxinas, dioxinas, agente tóxico, ingrediente ativo, NPC, cultivares, espécie cultivada, espécie alvo, associação perigosa de analitos, interação de agentes químicos, interação aditiva, interação sinérgica, potenciação, antagonismo, antídoto, efeito aditivo, efeito sinérgico, sinergismo, conformidade, não conformidade, toxidade, ensaios, itens de ensaio, anabolizantes, resultado analítico, metabólitos, certificado*

oficial de análise, compostos halogenados, limite de referência para tomada de ação regulatória, etc.

**2.1.2) Léxico de Referência (LR) :** Devido não pré-existir outra ontologia modelando semanticamente as associações perigosas de analitos, o Léxico de Referência foi inicialmente construído diretamente a partir da fusão seletiva do Léxico da Aplicação (termos diretamente relacionados pelos especialistas do domínio) com o Léxico do Domínio (termos extraídos a partir de documentos e recursos em fontes externas). A figura 35 ilustra o processo de geração do Léxico de Referência recomendado pelo método UPON:



**Figura 35 – Atividade de Construção do Léxico de Referência (LR)**  
Baseada no Método UPON (DE NICOLA; MISSIKOFF; NAVIGLI, 2009)

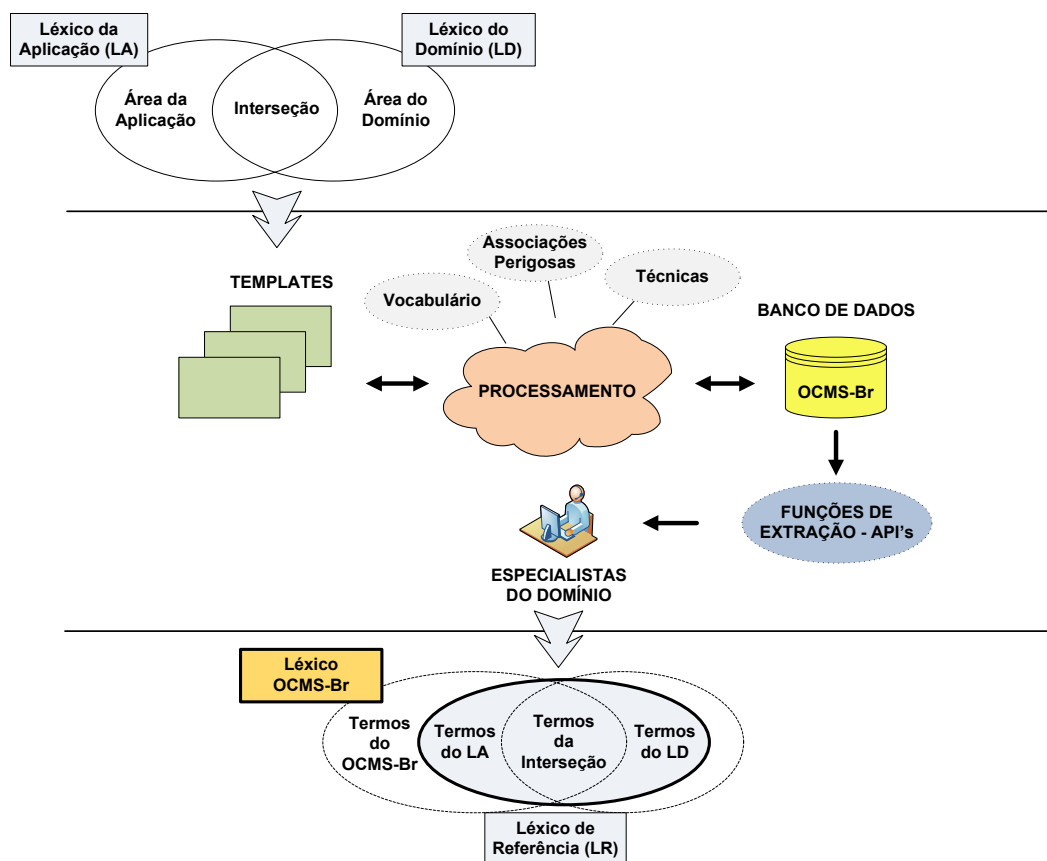
### **Produção do Léxico OCMS-Br – Abordagem ACATO (Estágio I)**

Adicionalmente a este processo de criação do Léxico de Referência, recomendado pelo método UPON, os diversos termos e Associações Perigosas de Analitos – APA's que o compõe, foram inseridos no banco de dados OCMS-Br, visando sua *validação* por parte dos especialistas externos, e, também, como motivação para *sugestão* de novos termos e associações perigosas.

Diferentemente do método UPON, os novos termos e associações perigosas sugeridos pelos especialistas externos formaram o Léxico OCMS-Br, o qual

conjuntamente com os termos e associações validados pelos especialistas de domínio do PNCRC / EMBRAPA, foram incorporados à ontologia APA.

A figura 36 ilustra o esquema de incorporação das técnicas de validação/sugestão, por meio do projeto OCMS-Br, ao processo de engenharia da ontologia APA adotado pela Abordagem ACATO. De acordo com esta figura, pode-se observar que os termos, técnicas e associações perigosas de analitos, coletados durante o fluxo de trabalho Requisitos, são inseridos diretamente no banco de dados OCMS-Br por meio de templates elaborados pelos engenheiros da ontologia APA. As templates também são utilizadas, em um segundo momento, para validação e coleta dos termos e associações sugeridas.



**Figura 36 – Processo de Produção e Incorporação do Léxico OCMS-Br ao Léxico de Referência (LR) - Abordagem ACATO (Estágio I)**



Alguns exemplos dos termos sugeridos e aprovados pelos especialistas são:

*Agente químico, agente tóxico, agente intoxicante, agente toxicante, interação de agentes tóxicos, ação tóxica, classes toxicológicas, intoxicação química, grau de toxicidade, receptor biológico, toxicidade, substância endógena, substância exógena, margem de segurança, carcinogênese química, hepatotoxicidade, antagonismo químico, antagonismo funcional, antagonismo competitivo, perigo químico.*

Exemplos de associações perigosas sugeridas e validadas pelos especialistas do PNCRC / EMBRAPA por meio da técnica OCMS-Br:

*Mercurio produzindo efeito tóxico redutor no Selênio;*

*Hidrocarbonetos e Nitratos em ação tóxica aditiva;*

*Dimercaprol produzindo efeito tóxico redutor no Mercúrio;*

*Toxicidade sinérgica do mercúrio e chumbo, chumbo e cádmio, cádmio e mercúrio (SCHUBERT et al; 1978);*

*Toxicidade sinérgica dos piretróides ocasionando enfermidades no sistema nervoso central e periférico (MARASCHIN, 2003).*

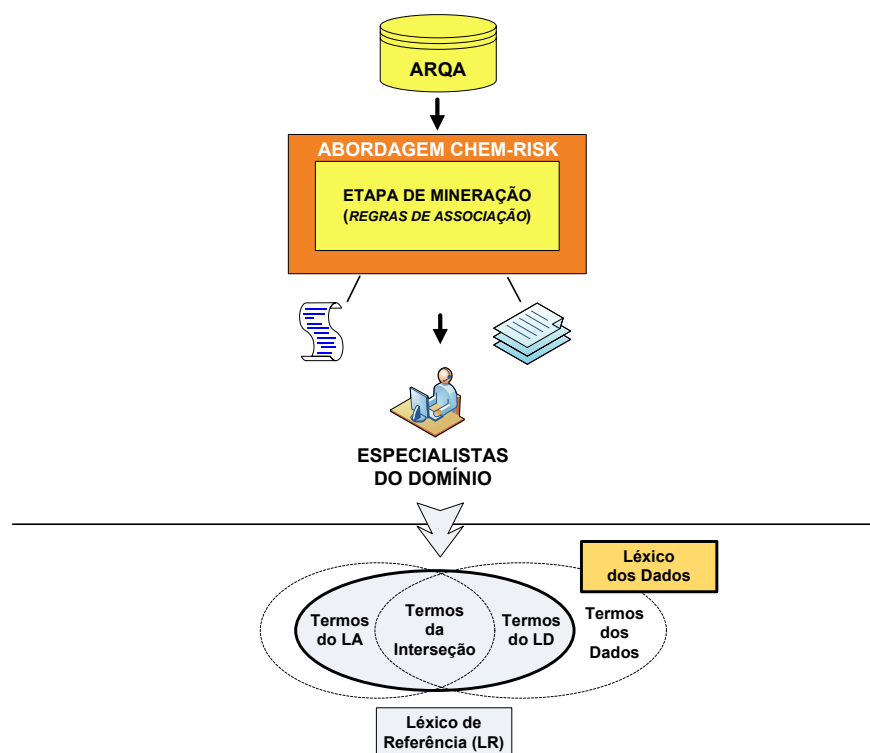
### **Produção do Léxico dos Dados – Abordagem ACATO (Estágio II)**

Como um segundo recurso adicional, visando que os termos e associações perigosas utilizados na ontologia APA alcancem uma maior completude, a abordagem ACATO utiliza as técnicas de associação e associação preditiva, para formar um Léxico de Dados, a partir dos resultados da etapa de mineração sobre o banco de dados ARQA, utilizando:

- Os resultados das análises laboratoriais nos analitos das amostras das matrizes agropecuárias analisadas pelo PNCRC;
- As avaliações de riscos químicos em analitos, amostras e lotes de amostras;

O Léxico dos Dados, contendo as associações de analitos sugeridas por este processo, deverá ser validado pelos especialistas do PNCRC / EMBRAPA e, em seguida, incorporado ao processo de formação do Léxico de Referência, recomendado pelo método UPON.

A figura 37 ilustra o esquema de incorporação do Léxico dos Dados ao processo de engenharia da ontologia APA, adotado pela Abordagem ACATO:



**Figura 37 – Processo de Produção e Incorporação do Léxico dos Dados ao Léxico de Referência (LR) - Abordagem ACATO (Estágio II)**

As associações perigosas de analitos, sugeridas por meio das técnicas de associação e associação preditiva e aprovadas pelos especialistas de domínio, são apresentadas a seguir. Deve ser ressaltado que a sugestão realizada a partir dos dados não é capaz de informar o tipo da associação perigosa, ou seja, Efeito Tóxico Aditivo, Efeito Tóxico Sinérgico, etc. Por esta razão, o processo de validação destas associações sugeridas baseia-se unicamente no conhecimento dos especialistas de domínio e no conhecimento diretamente coletado a partir dos especialistas externos por meio das técnicas do projeto OCMS-Br.

*Metamidofós e Mevinfós, Mirex e Mevinfós, Mirex e Metamidofós, Moxidectina e Albendazol, Abamectina e Albendazol, Albendazol e Mivinfós, Azinfós-Etil e Azinfós-Metil, Azinfós-Metil e Sulfatiazol, Azinfós-Etil e Sulfatiazol, Aldicarb e Sulfatiazol.*

A figura 38 resume o processo de produção do Léxico OCMS-Br e Léxico dos Dados utilizados na Abordagem ACATO para produção do Léxico de Referência:

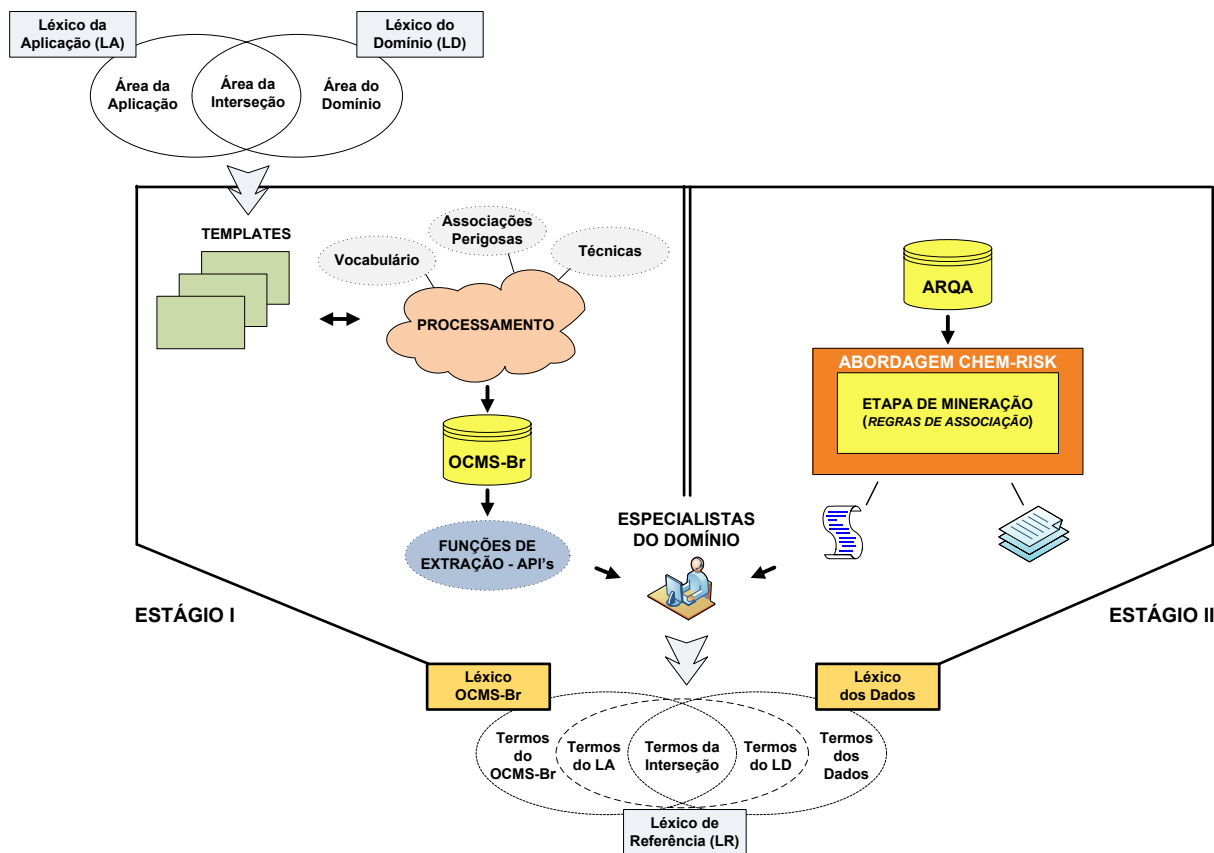


Figura 38 – Processo de Produção do Léxico OCMS-Br, Léxico dos Dados e Incorporação ao Léxico de Referência (LR) – Abordagem ACATO (Estágio I e II)

**2.1.3) Glossário de Referência (GR) :** Após o processo de filtragem dos novos termos e associações sugeridos por especialistas externos, apresenta-se, a seguir, uma parte dos termos e conceitos associados que compõem o Glossário de Referência (GR) da ontologia APA:

**Analito:** Substância analisada passível de ser encontrada nas amostras das matrizes agropecuárias. São pré-determinados pelo Codex Alimentarius e PNCRC de acordo com a matriz analisada, por exemplo: Ivermectina, Aflatoxina, Oxitetraciclina, Dioxinas, Chumbo, Mercúrio, etc. No âmbito do PNCRC subdivide-se em:

**Resíduo:** No domínio das associações perigosas de analitos, resíduo é uma especialização de analito, consistindo na matéria ou parte desta que consideramos indesejável e sem utilidade. Esta modelagem considera os resíduos químicos passíveis

de serem detectados nos alimentos monitorados pelo PNCRC, que apresentam periculosidade à saúde humana e animal;

**Contaminante:** No domínio das associações perigosas de analitos, contaminante é uma especialização de analito e refere-se a qualquer substância adicionada intencionalmente ou não aos alimentos, como resultado da produção, fabrico, transformação, preparação, tratamento, embalagem, transporte ou armazenagem do referido alimento ou em resultado de contaminação ambiental. Nesta modelagem considera-se os alimentos agropecuários e contaminantes monitorados pelo PNCRC;

**Grupo de Analitos:** Grupo de analitos pré-determinado pelo Codex Alimentarius e PNCRC de acordo com sua classe química, por exemplo: Antimicrobianos, Organofosforados, Antiparasitários, etc;

**Produto Agropecuário:** No domínio das associações perigosas de analitos, o termo produto agropecuário refere-se ao bem acabado e pronto para comercialização, originado a partir de um processo de produção agropecuária monitorado pelo PNCRC / MAPA.

**Matriz Agropecuária:** Fração do produto agropecuário a ser analisado na forma de amostras laboratoriais, por exemplo: Carne-Bovina-Músculo, Carne Bovina-Rim, Aves-Fígado, Leite, Alface, Milho, etc;

**Amostra laboratorial:** parte ou fração de uma matriz agropecuária submetida à análise laboratorial;

**Limite de Referência:** Corresponde ao limite máximo permitido para a concentração do analito detectado em uma amostra de matriz analisada. São estabelecidos e publicados mundialmente pela Comissão do Codex Alimentarius, podendo existir algumas diferenças com relação aos diferentes planos de monitoramento adotados por alguns países;

**Limite Máximo de Resíduo (LMR):** Limite de referência pré-estabelecido para determinado resíduo detectado em uma amostra de matriz agropecuária;

**Teor Máximo de Contaminante (TMC):** Limite de referência pré-estabelecido para determinado contaminante detectado em uma amostra de matriz agropecuária;

**Risco Químico:** Probabilidade de um fator de risco químico, por exemplo, a Concentração do Analito – CA, alcançar um valor prejudicial e inviabilizar o processo de comercialização de um produto agropecuário;

**Avaliação de Risco Químico:** Método utilizado para avaliar o risco químico produzido por um fator de risco;

**Fator de Risco Químico:** Evento químico que possa prejudicar o processo de comercialização de um produto agropecuário;

**Pesticidas:** Termo genérico das substâncias inseticidas empregadas para proteger as plantas do ataque de parasitas. Os termos, pesticida, biocida, defensivo agrícola, praguicida e agrotóxico são comumente usados com o mesmo significado;

**Metais Pesados:** Elementos químicos constituintes naturais da crosta terrestre estando amplamente espalhados na natureza, por exemplo: Chumbo, Cádmio, Mercúrio, etc. Alguns elementos são tóxicos em concentrações elevadas e como tendem a acumular-se nos organismos vivos ao longo do tempo, constituem um risco para a saúde humana e animal;

**Nitratos:** Os nitratos (NO<sub>3</sub><sup>-</sup>) são constituintes azotados cuja presença é natural no meio ambiente em consequência do ciclo do azoto. São adicionados como fertilizantes com o objetivo de promover o crescimento rápido e obtenção de folhas maiores e mais vistosas dos produtos hortícolas. O uso abusivo conduz ao aumento de teor de nitratos nas plantas e solos, além de contaminar os lençóis freáticos e águas superficiais;

**Micotoxinas:** A contaminação química dos alimentos pode resultar de processos naturais, que envolvem a ocorrência de toxinas produzidas pelos próprios produtos alimentares ou por outros organismos vivos, que por serem altamente tóxicas, constituem um sério risco para a saúde humana e animal. Por exemplo: Zearalenona (Milho), Aflatoxina\_M1 (Leite), Aflatoxina\_B1 (Amendoim), Ocratoxina\_A (Café), Patulina (Maçã), etc;

**Dioxinas e PCB's:** Dioxinas e policlorodibenzodioxinas são os compostos químicos mais tóxicos resultantes da ação do homem. Resulta da combustão ou incineração de resíduos e processamento de metais;

**Organofosforados:** são substâncias químicas (pesticidas) que contém carbono e fósforo, sendo geralmente obtidas através do uso de sais orgânicos do ácido fosfórico.

São extremamente tóxicos, sendo fatais na proporção de alguns miligramas para um homem de cerca de 70 kg;

**Organoclorados:** Composto orgânico de ação cancerígena contendo ao menos um átomo de cloro ligado covalentemente. Sinônimos: organocloro, organocloreto, clorocarbono, hidrocarboneto clorado e solvente clorado;

**Carbamatos:** Os inseticidas carbamatos são compostos anticolinestecásicos com variado grau de toxicidade para o ser humano. Estas substâncias vêm sendo amplamente utilizadas como inseticidas, fungicidas e parasiticidas na agricultura. A exposição ao carbamato no sistema respiratório leva diminuição da acetilcolinesterase e aumento acentuado das secreções brônquicas e dispnéias sibilias. A forma mais comum de morte é falência por broncoconstrição, broncorreia, paralisia dos respiratórios e de pressão dos centros respiratórios.

**Agente Tóxico:** No contexto das associações perigosas de analitos é sinônimo de analito. Entidade química capaz de causar dano a um sistema biológico, alterando uma função ou levando-o à morte, sob certas condições de exposição;

**Associações Perigosas de Analitos:** São as interações entre analitos, onde o efeito final de tóxicos combinados pode ser maior ou menor que a soma dos efeitos individuais. Deve-se, na exposição a dois ou mais agentes intoxicantes, considerar a possibilidade de um analito interferir na ação do outro;

**Efeito Aditivo:** O efeito tóxico final é igual à soma dos efeitos tóxicos de cada um dos agentes envolvidos;

**Efeito Sinérgico:** O efeito tóxico final é maior que a soma dos efeitos tóxicos causados por cada agente em separado;

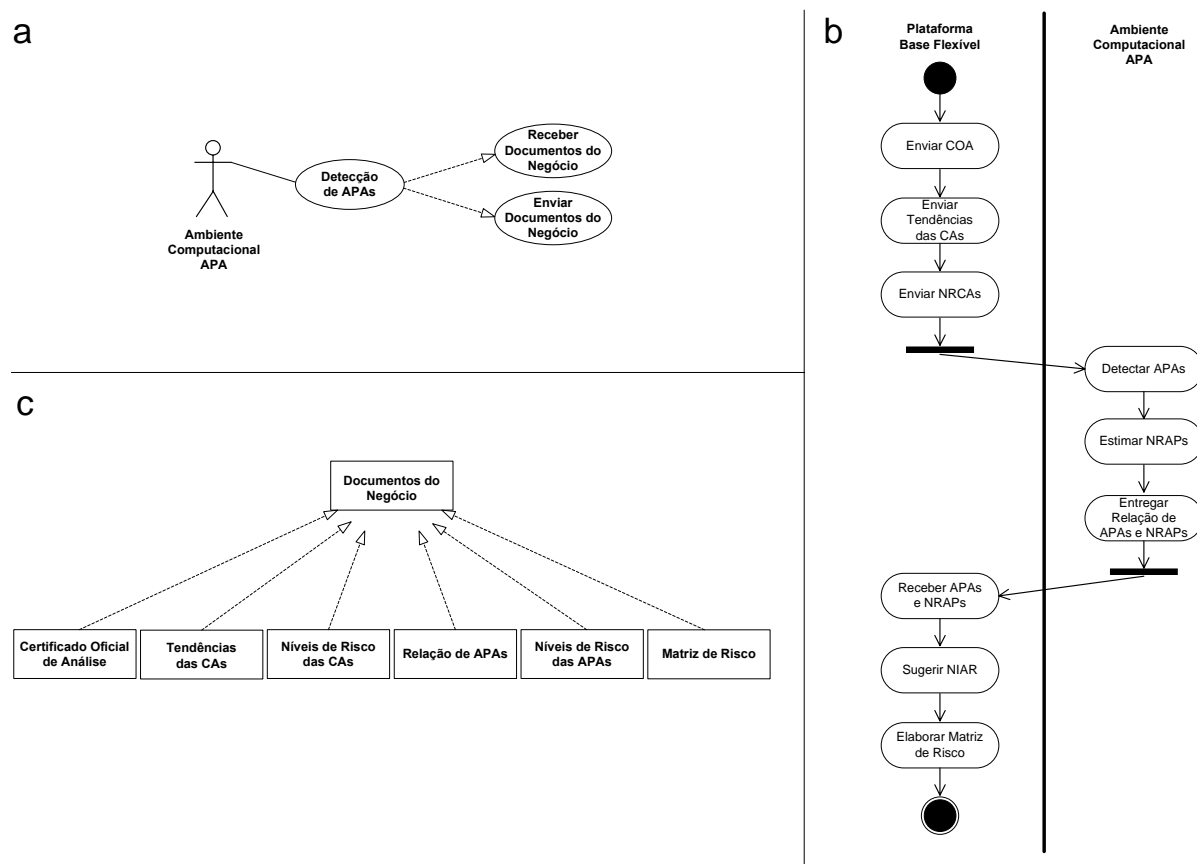
**Potencialização:** O efeito tóxico de um agente é aumentado quando em combinação com outro agente;

**Antagonismo:** O efeito de um agente tóxico é diminuído, inativado ou eliminado quando em combinação com outro agente;

## 2.2) Modelagem do Cenário da Aplicação

Nesta atividade o cenário da aplicação é modelado, adicionando as atividades e os diagramas de classe aos diagramas de caso de uso desenhados durante o fluxo de trabalho. Requisitos Os diagramas UML ilustrados na figura 39 apresentam parte do modelo de aplicação

usado para validar a ontologia APA. O Diagrama de Caso de Uso ilustrado na letra “a” ilustra o processo de Detecção das APA’s enfocando o trâmite dos documentos envolvidos neste negócio, citados no diagrama apresentado na letra “c”. O Diagrama de Atividades, ilustrado na letra “b”, apresenta o fluxo das atividades necessárias para inserção das APA’s detectadas na matriz de risco elaborada pela Plataforma Base Flexível. Informações técnicas acerca do processo de elaboração de diagramas UML podem ser encontradas em <http://www.uml.org>.



**Figura 39 – Diagramas UML Modelando uma Parte do Cenário da Aplicação: a) Diagrama de Caso de Uso; b) Diagrama de Atividade; c) Diagrama de Classe**

### 3) Fluxo de Trabalho – Projeto

Neste fluxo deixaremos a dimensão linguística da ontologia APA e entraremos na dimensão conceitual, visando fornecer uma estrutura ontológica ao conjunto de entradas referenciadas no Glossário de Referência (GR). Isto é essencialmente alcançado por meio da

organização de termos, de acordo com as hierarquias conceituais e sua estruturação por meio de atributos e axiomas.

### **3.1) Modelagem dos Conceitos**

Os principais conceitos identificados nas principais categorias da ontologia APA são:

#### **Atores**

Os atores que operam, produzem, atualizam ou consomem os objetos de negócio, de acordo com os especialistas do domínio são:

*(AT1) Os Especialistas do Domínio do PNCRC;*

*(AT2) Os Analistas de Risco do PNCRC;*

*(AT3) Laboratórios Credenciados pertencentes à rede LANAGRO;*

*(AT4) Plataforma Base Flexível;*

*(AT5) Ambiente Computacional APA (Regras SWRL, Linguagem SQWRL e API Java);*

#### **Objetos do Negócio:**

Os principais objetos ou documentos pertencentes ao domínio de negócio sobre os quais o processo de negócio opera são:

*(OB1) Certificado Oficial de Análise – COA emitido pelos laboratórios credenciados do PNCRC*

*(OB2) PNCRC – Animal (Ver Quadro 1, Seção 3.4)*

*(OB3) PNCRC – Vegetal (Ver Quadro 2, Seção 3.4)*

*(OB4) Associações Perigosas de Analitos – APA detectadas na amostra (Ver item 1.5 – Fluxo de Trabalho, Requisitos)*

#### **Processos**

As atividades principais que operam sobre os objetos de negócio e contribuem para satisfazer o objetivo do negócio, citado no item 1.3, ou seja, a detecção de associações perigosas em amostras de matrizes agropecuárias, são:

*(PR1): Obtenção do Certificado Oficial de Análise (COA) contendo, entre outras informações, o produto, a matriz agropecuária, a identificação da amostra, os analitos detectados e as concentrações detectadas;*

*(PR2): Detecção das associações perigosas de analitos;*

*(PR3): Inferir os níveis de risco das associações perigosas detectadas;*



*(PR4): Relacionar as associações perigosas detectadas na amostra;*

*(PR5): Disponibilização dos níveis de risco das associações perigosas;*

### **Mensagens**

Exemplos das mensagens trocadas durante as interações entre processos são:

*(MSG1): Matriz agropecuária analisada*

*(MSG2): Amostra da matriz agropecuária analisada*

*(MSG3): Analitos detectados na amostra da matriz agropecuária analisada*

*(MSG4): Concentrações dos Analitos (CA) detectados na amostra da matriz agropecuária analisada*

*(MSG5): Nível de Risco da Concentração do Analito (NRCA)*

*(MSG6): Tendência da concentração (TDCA) detectada*

*(MSG7): Grupo do analito detectado*

*(MSG8): Associações perigosas de analitos detectadas na amostra analisada*

*(MSG9): Níveis de risco das associações perigosas de analitos (NRAP) detectadas na amostra analisada*

### **Atributos**

Exemplos de atributos complexos e atômicos ou estruturas de informação dos conceitos trabalhados pela ontologia APA são:

*(Atributo Complexo): Endereço Entrega subdividido em:*

*(Atributo Atômico): Região*

*(Atributo Atômico): Estado*

*(Atributo Atômico): Cidade*

*(Atributo Atômico): CEP*

*Etc.*

### **3.2) Modelar hierarquia de conceitos e relacionamentos específicos do domínio**

A organização dos conceitos em uma hierarquia taxonômica traz como resultado uma rede semântica representada por diagramas de classe UML usando generalização (é-um), agregação (parte-de) e associação. A figura 40 ilustra em forma diagramática, o processo (PR2) identificado para a aplicação, apresentando estrutura “parte-de”. A figura 41 ilustra um exemplo de relacionamentos que envolvem a aplicação Associações Perigosas de Analitos – APA:

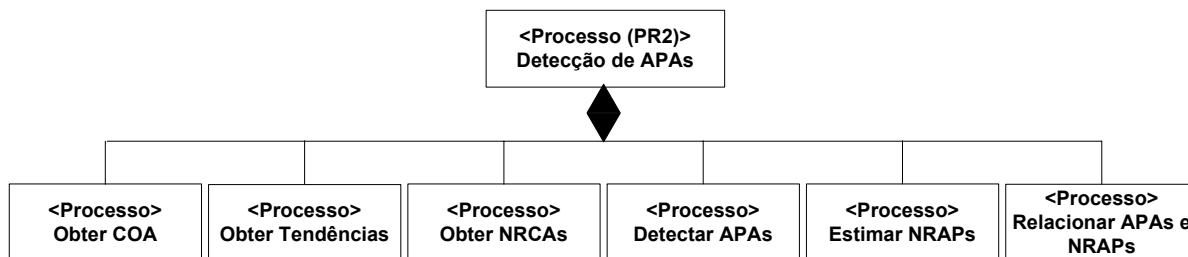


Figura 40 – Processo Identificado (PR2) para a Aplicação APA com Estrutura Parte-de

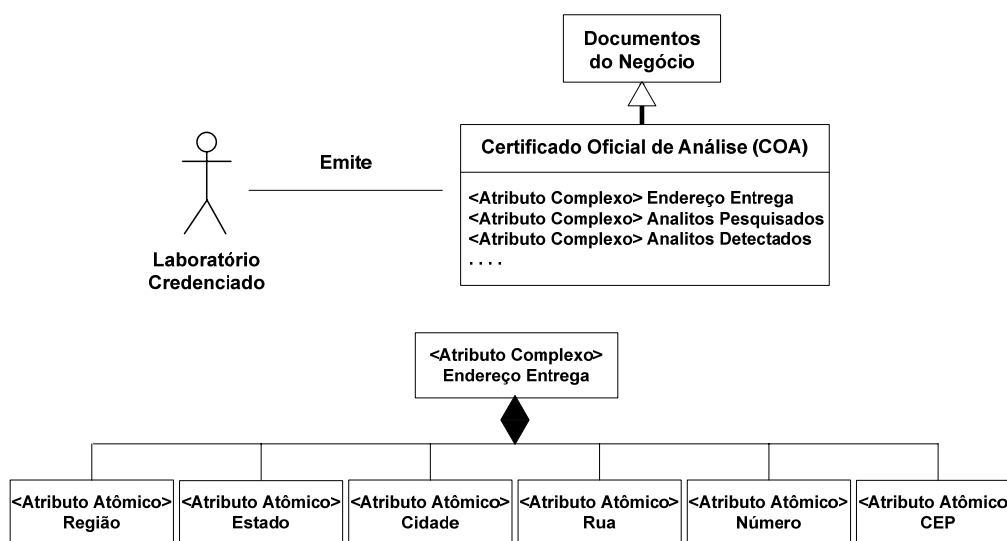


Figura 41 – Exemplos de Relacionamentos Ontológicos para Aplicação APA

#### 4) Fluxo de Trabalho – Implementação

Neste fluxo de trabalho a codificação da ontologia foi realizada utilizando uma linguagem formal rigorosa, considerando o poder expressivo, a complexidade computacional do método de raciocínio e o nível de aceitação na comunidade científica. Além de conceitos e relações, uma ontologia é caracterizada por um conjunto de axiomas que, essencialmente, são restrições impostas pela realidade modelada. Os axiomas podem estar associados à componente estrutural (por exemplo, restrições de cardinalidade em um atributo), à hierarquia (por exemplo, a disjunção de dois subconceitos), ou stand-alone, expressos por linguagem específica (por exemplo, uma expressão booleana). Adicionalmente a estas e outras características, apresentadas na seção 5.3, como resultado de um longo esforço de padronização, a Ontology Web Language (OWL), particularmente o dialeto OWL-DL, foi escolhido para codificar a ontologia APA.

Utilizando o dialeto OWL-DL, de acordo com suas características e finalidade, a ontologia APA foi implementada como uma aplicação web podendo ser acessada remotamente por meio de um navegador utilizado como cliente. Associadas ao método de construção UPON, citado na seção 5.4, as ferramentas utilizadas para sua construção e utilização, citadas nas seções 5.5 e 5.6, foram: Protegé 3.4.4; Semantical Web Rule Language (SWRL); Semantical Query Web Rule Language (SQWRL); Jess Rule Engine e SWRL Query Java API;

*Nota.* A versão 3.4.4 do Protegé foi escolhida para realizar a modelagem semântica da ontologia APA, devido às questões de compatibilidade e sincronismo com as demais ferramentas.

A figura 42 apresenta um recorte do código em OWL-DL gerado para a ontologia APA, focando na parte de inserção de axiomas associados à hierarquia (disjunção de conceitos) e na parte de inserção das regras SWRL:

```

<?xml version="1.0"?>
...
<owl:Class rdf:ID="Analito">
  <owl:disjointWith rdf:resource="#AssociacaoPerigosaAnalitos"/>
  <owl:disjointWith rdf:resource="#MatrizAgropecuaria"/>
  <owl:disjointWith rdf:resource="#NivelRiscoAssociacaoPerigosa"/>
  <owl:disjointWith rdf:resource="#NivelRiscoConcentracaoAnalito"/>
  <owl:disjointWith rdf:resource="#TendenciaConcentracaoAnalito"/>
  <owl:disjointWith rdf:resource="#&swrl;Entity"/>
</owl:Class>
...
<MatrizAgropecuaria rdf:ID="Feijao"/>
<Analito rdf:ID="Fenvarelato"/>
<Analito rdf:ID="Gamacialotrina"/>
<Analito rdf:ID="Gentamicina"/>
<owl:Class rdf:ID="GrupodoAnalito"/>
<owl:ObjectProperty rdf:ID="hasAnalitoAmostra"/>
<swrl:Imp rdf:about="http://www.cnppl.embrapa.br/ontologias/apa-001">
  <swrl:body>
    <swrl:AtomList>
      <rdf:first>
        <rdf:Description>
          <rdf:type rdf:resource="#&swrl;IndividualPropertyAtom"/>
          <swrl:argument2>
            <rdf:Description rdf:about="http://www.cnppl.embrapa.br/ontologias/ana-001"/>
          </swrl:argument2>
          <swrl:argument1>
            <rdf:Description rdf:about="http://www.cnppl.embrapa.br/ontologias/amo-001"/>
          </swrl:argument1>
          <swrl:propertyPredicate rdf:resource="#hasAnalitoAmostra"/>
        </rdf:Description>
      </rdf:first>
    </swrl:AtomList>
  </swrl:body>
...

```

**Figura 42 – Fragmento de Código em OWL-DL da Ontologia Associações Perigosas de Analitos - APA**

## 5) Fluxo de Trabalho – Teste

No método UPON o fluxo de teste é concebido para verificar a qualidade semântica e pragmática da ontologia. De acordo com (BURTON-JONES et al, 2005) a qualidade de uma ontologia é uma característica multi-dimensional e, por esta razão, uma ontologia deve ser avaliada com relação a quatro diferentes características multi-dimensionais:

- **Qualidade sintática:** mede a qualidade da ontologia de acordo com seu estilo formal, do jeito que está escrito;

*No caso da ontologia APA, existe uma concordância entre a forma sintática dos termos armazenados no banco de dados Análise de Risco Químico Agropecuário – ARQA e que compõem o Léxico de Referência (LR) projetado no Fluxo de Análise do método UPON para utilização na ontologia APA. Desta forma, a qualidade sintática foi verificada no fluxo de trabalho anterior e garantida durante a codificação da OWL-DL, gerada automaticamente pela ferramenta Protégé 3.44;*

- **Qualidade semântica:** segundo os autores, a principal preocupação é com a ausência de conceitos contraditórios. Desta forma, a coerência pode ser verificada pela consistência da ontologia por meio de um raciocinador, como Racer ou Pellet.

*A coerência e consistência da ontologia APA foi verificada pelo raciocinador Pellet que acompanha a versão 3.4.4 do Protégé;*

A coerência também exige que a modelagem semântica seja corretamente utilizada, por exemplo, caso exista ausência de ciclos de construção ou o fato de que classes e propriedades sejam disjuntas.

*Conforme recomendado pelo método de construção UPON, os ciclos, fases, fluxos de trabalho e iterações recomendados, foram utilizados para construção da ontologia APA;*

*Adicionalmente, na ontologia APA todas as classes, subclasses e propriedades são disjuntas de forma a não causar resultados inesperados durante a execução do raciocinador (Observar o termo *disjointWith* na figura 42);*

- **Qualidade pragmática:** refere ao conteúdo da ontologia e utilidade para os usuários, independentemente da sua sintaxe e semântica. Está relacionada com três características:

- **Fidelidade:** pode ser medida verificando se as reivindicações que a ontologia faz são verdadeiras no domínio considerado. Esta tarefa pode ser alcançada, por exemplo, verificando-se as referências às fontes utilizadas nas descrições dos termos;

*No caso da ontologia APA, reportar-se ao fluxo de trabalho Análise, seção 2.1.1 – Léxico do Domínio (LD);*

- **Relevância:** verificada em conjunto com a integralidade, a verificação da correta aplicação dos requisitos da ontologia, reuniram-se no primeiro fluxo de trabalho, ou seja, o fluxo de trabalho Requisitos;
- **Integralidade:** segundo os autores, uma ontologia é completa e eficaz quando satisfaz os requisitos e atende as restrições do problema a que se propõe resolver. Por esta razão, a integralidade da ontologia pode ser avaliada verificando se os objetivos do domínio, o escopo e as questões de competência, definidos no fluxo de trabalho Requisitos, foram alcançados.

*De acordo com as recomendações do método UPON, pode-se afirmar que a ontologia APA atende ao objetivo do negócio, escopo e questões de competência definidos no fluxo de trabalho Requisitos, retornando corretamente todas as associações perigosas de analitos e os níveis de risco destas associações, porventura existentes em uma amostra de matriz agropecuária, a partir de consultas a nível semântico, disparadas pela Plataforma Base Felxível da Abordagem Chem-Risk;*

- **Qualidade social:** reflete critérios mais gerais, por exemplo:
  - O número de outras ontologias que apontam para ela, por exemplo, através da definição de seus termos e definições;
  - O número de vezes que é acessada (quando pública) a partir de dentro e fora da comunidade que a gerencia;

No caso da ontologia APA, a qualidade social somente poderá ser verificada após sua publicação;

### 8.4.2 Estimativa do Nível de Risco de Associação Perigosa - NRAP

A título de exemplo, a figura 43 apresenta o esquema conceitual para estimativa do NRAP relacionado à associação perigosa APA-003. Observe que esse exemplo didático apenas ilustra a participação de ontologias no processo de Análise de Risco Químico Agropecuário, desenvolvido pelo Ministério da Agricultura, Pecuária e Abastecimento – MAPA, e representa um pequeno fragmento da ontologia Associações Perigosas de Analitos – APA, apresentando-se, por esta razão, muito simplificado em relação ao contexto maior da modelagem semântica necessária. O acrônimo “i-o” refere-se ao termo “instance-of”:

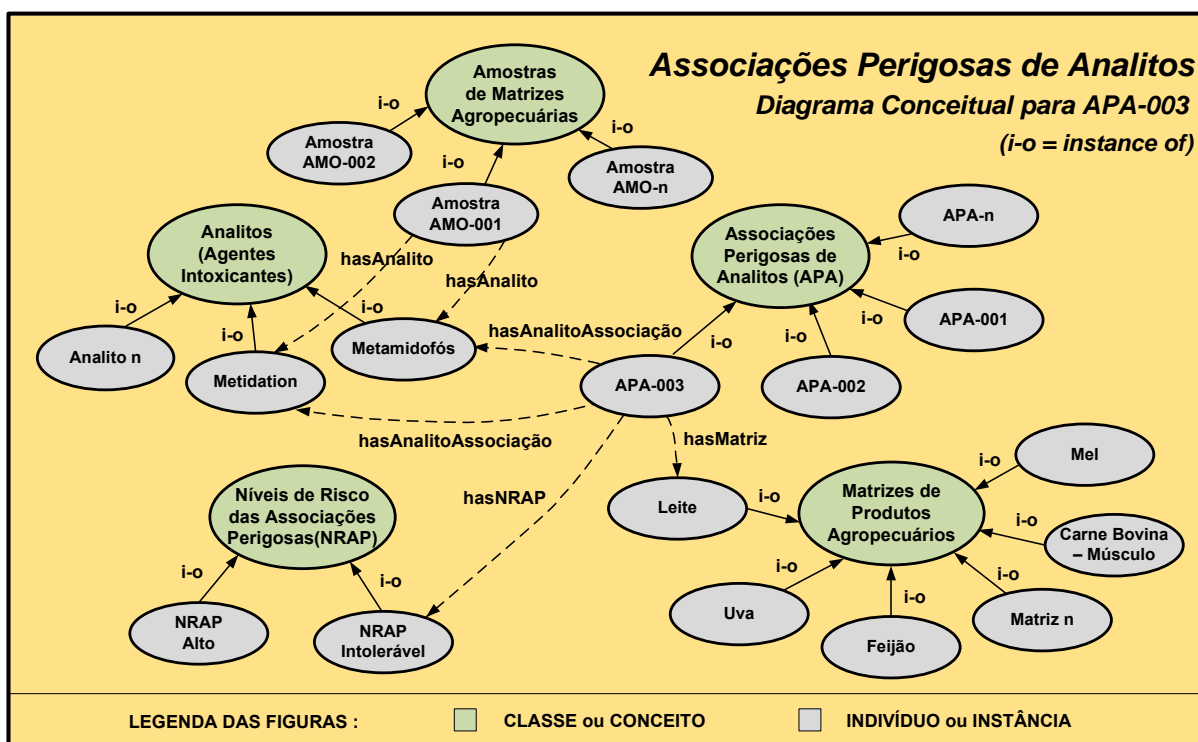


Figura 43 – Fragmento da Ontologia Associação Perigosa de Analitos – APA Ilustrando a Associação Perigosa APA-003 para Matriz Agropecuária Leite

## **8.5 Resultados da Incorporação da Lógica *Fuzzy* e Ontologias**

Os resultados apresentados nesta seção foram gerados a partir da massa de testes laboratoriais, simulada de acordo com os padrões técnicos e unidades de medição adotados pelo PNCRC. Os estudos comparativos envolveram as abordagens Tradicional, Vetor de Risco Ponderado e Chem-Risk, consideradas, segundo autoridades do PNCRC, como abordagens únicas no contexto das políticas públicas brasileiras que visam a maximização da segurança alimentar química dos produtos agropecuários brasileiros, em nível nacional. De acordo com o item 1 da seção 8.2.1 – Medidas Qualitativas de Desempenho, apesar da Abordagem Chem-Risk permitir a realização de avaliações de riscos químicos agropecuários incrementais, ou seja, em analitos, grupos de analitos, amostras, grupos especiais de amostras, lotes de amostras, matrizes, produtos, cidades, estados, regiões, zonas geográficas e país, as comparações limitaram-se às medições realizadas somente ao nível dos analitos (CA's), amostras e lotes de amostras, de acordo com o escopo também adotado pelas abordagens Tradicional e Vetor de Risco Ponderado. Os resultados aferidos demonstraram a superioridade da Abordagem Chem-Risk, mediante o ganho de informação proporcionado pela maior exatidão de dados e resultados, e, também, por permitir melhor expressar a realidade da segurança química do agronegócio brasileiro, de forma mais detalhada, abrangente e fidedigna.

### **8.5.1 Resultados das Avaliações de Riscos nas Concentrações de Analitos**

Na primeira bateria de testes envolvendo as abordagens Tradicional, Vetor de Risco Ponderado e Chem-Risk, os resultados das análises laboratoriais correspondentes à cada Concentração de Analito (CA) detectada, foram submetidos aos diferentes métodos de avaliação de riscos químicos, adotados pelas abordagens. Em seguida, os resultados das avaliações foram armazenados nos bancos de dados particulares das abordagens. Em virtude da Abordagem Chem-Risk estabelecer o nível de risco (NRCA) “Alto” em 5 e as abordagens Tradicional (ACA) e Vetor de Risco Ponderado (NRC) estabelecê-lo em 2 e 3, respectivamente, foi realizada uma equivalência dos níveis de risco altos aferidos pelo métodos ACA e NRC ao NRCA. Semelhantemente, para efeito de comparação, o cálculo do Resultado Percentual (RP) inferido a partir do confronto entre as Concentrações dos Analitos (CA) e seus Limites de Referência (LR) respectivos, utilizado para estimativa do NRCA, foi estendido aos métodos para avaliação da ACA, sem, contudo, causar qualquer interferência nos resultados apresentados pelo método.

Associada à figura 44, a tabela 17 apresenta os resultados das avaliações de risco armazenados nos bancos de dados particulares das abordagens. Os Resultados Percentuais, colocados em ordem crescente, visam agrupar e melhor evidenciar os ganhos de informação obtidos a partir da incorporação dos fatores de risco adicionais, ao nível dos analitos, e dos conceitos da lógica *fuzzy* às avaliações de risco em analitos, realizadas durante a execução da Macro-Etapa 1 de pré-processamento de dados da Abordagem Chem-Risk.

*Nota.* O NRAA realiza uma previsão para cada concentração detectada, incorporando a esta, de acordo com a experiência de especialistas, uma tendência estimada, uma estimativa de prazo e de custo de ajustamento do processo produtivo aos índices de segurança mundiais. A coluna NRAA da tabela 17 exhibe, entre parêntesis, da esquerda para a direita, os fatores de risco tendência, prazo de ajustamento e custo de ajustamento atribuídos pelos especialistas. Para encontrar o NRAA, os valores atribuídos aos fatores de risco devem ser somados ao NRCA.

**TABELA 17 – Avaliações de Risco Químico em Carne Bovina – Músculo – Amostra 5/ Lote 229**

| Analito | Resultado Percentual (RP) | NRAA         | NRCA | NRC | ACA |
|---------|---------------------------|--------------|------|-----|-----|
| 77      | 11%                       | 2.00 (0+0+1) | 1    | 1   | 1   |
| 71      | 15%                       | 3.00 (1+0+1) | 1    | 1   | 1   |
| 35      | 22%                       | 1.10 (0+0+0) | 1.10 | 2   | 1   |
| 74      | 25%                       | 1.25 (0+0+0) | 1.25 | 2   | 1   |
| 79      | 34%                       | 3.70 (1+1+0) | 1.70 | 2   | 1   |
| 75      | 44%                       | 3.00 (0+1+0) | 2    | 2   | 1   |
| 22      | 53%                       | 3.00 (0+1+0) | 2    | 2   | 1   |
| 73      | 81%                       | 3.10 (0+1+0) | 2.10 | 2   | 1   |
| 78      | 86%                       | 2.60 (0+0+0) | 2.60 | 2   | 1   |
| 23      | 93%                       | 3.00 (0+0+0) | 3    | 2   | 1   |
| 21      | 100%                      | 4.00 (0+0+0) | 4    | 2   | 1   |
| 76      | 107%                      | 6.00 (0+0+1) | 5    | 3   | 2   |
| 36      | 115%                      | 7.00 (0+1+1) | 5    | 3   | 2   |
| 70      | 122%                      | 6.20 (0+0+1) | 5.20 | 3   | 2   |
| 34      | 124%                      | 7.40 (0+1+1) | 5.40 | 3   | 2   |
| 72      | 124%                      | 7.40 (0+1+1) | 5.40 | 3   | 2   |
| 20      | 129%                      | 7.90 (0+1+1) | 5.90 | 3   | 2   |
| 37      | 133%                      | 8.00 (0+1+1) | 6    | 3   | 2   |
| 24      | 141%                      | 9.00 (1+1+1) | 6    | 3   | 2   |



Gerada na Macro-Etapa 3 da Abordagem Chem-Risk, a figura 44 ilustra os resultados comparativos para 19 Concentrações de Analitos (CA) detectadas na amostra número 5 do lote 229, respectivo à matriz “Carne Bovina – Músculo” analisada. As áreas achurradas e as retas de equivalência da ACA e NRC ao NRCA foram adicionadas posteriormente à figura original:

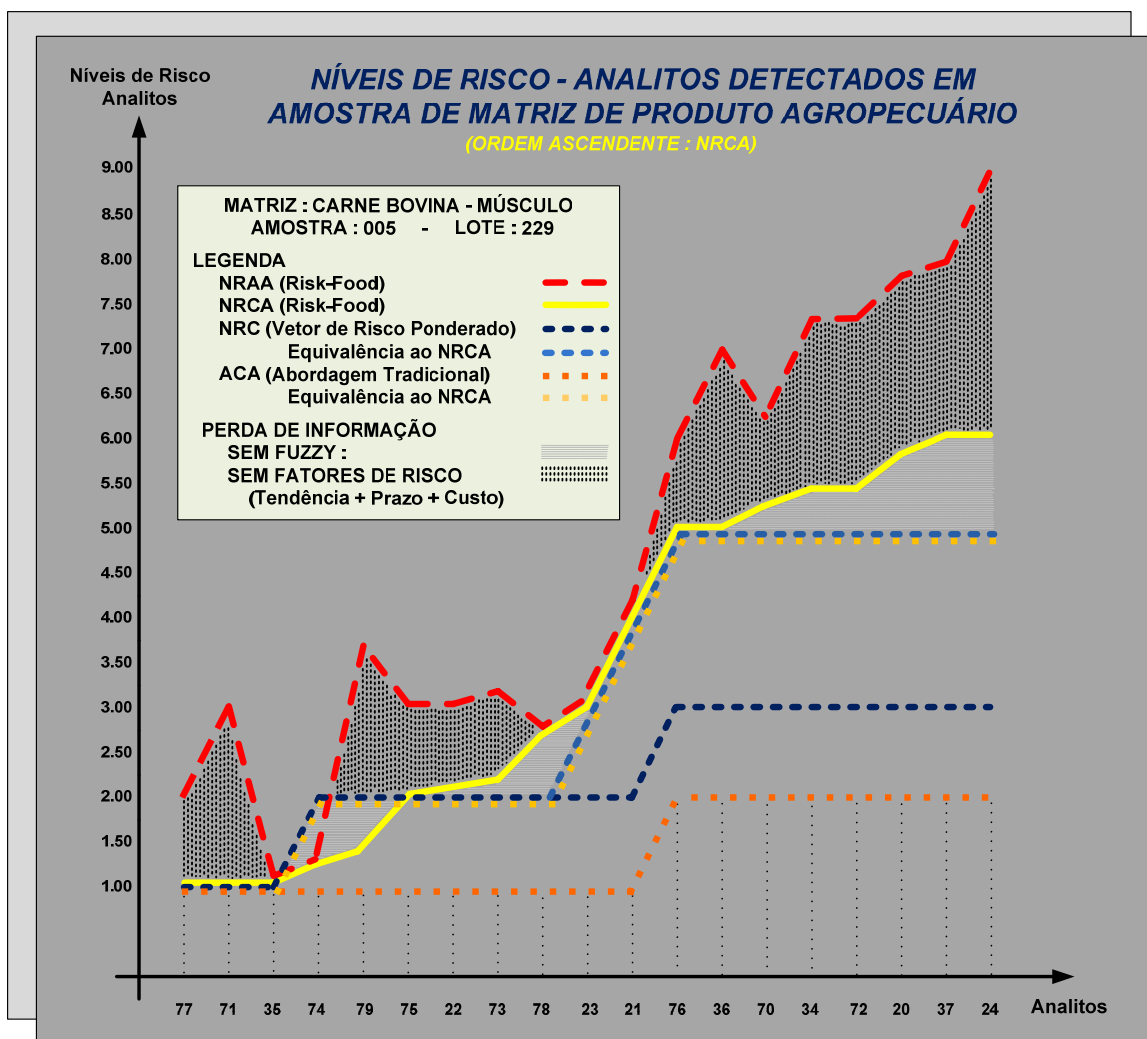


Figura 44 – Resultados Comparativos dos Níveis de Riscos Químicos das Concentrações de Analitos - Abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk

### 8.5.2 Resultados das Avaliações de Risco em Amostras

Na segunda bateria de testes foi realizado um estudo comparativo, desta vez, com os níveis de riscos químicos aferidos para as amostras analisadas, considerando os diferentes métodos de

avaliação, quais sejam, “Avaliação da Amostra (AA)”, “Nível de Risco da Amostra (NRA)” e “Nível de Risco da Amostra Original (NRAO)” adotados, respectivamente, pela abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk. Semelhantemente às comparações das Concentrações dos Analitos (CA’s), os resultados das avaliações em amostras foram recuperados a partir dos bancos de dados particulares utilizados pelas abordagens.

Associada à figura 45, a tabela 18 apresenta os resultados das avaliações de risco ao nível das amostras, armazenados nos bancos de dados particulares das abordagens. A título de comparação, a técnica dos percentuais de conformidade e de não conformidade, utilizada na Abordagem Chem-Risk, foi estendida aos métodos para estimativa da AA e NRA sem interferir nos resultados destas estimativas. Desta forma, os resultados das avaliações de risco para as amostras foram calculados de acordo com os Percentuais de Conformidade (PC) aos Intervalos de Conformidade (ICA) das amostras conformes e com os Percentuais de Não Conformidade (PNC) aos Intervalos de Não Conformidade (INCA) das amostras não conformes. Os percentuais de conformidade e não conformidade foram colocados em ordem crescente, separadamente considerados, visando agrupar e melhor evidenciar os ganhos de informação obtidos a partir da incorporação dos conceitos da lógica *fuzzy* e ontologias às avaliações de risco em amostras, realizadas durante a execução da Macro-Etapa 1 de pré-processamento de dados da Abordagem Chem-Risk.

*Nota.* Observe que o NRAO reflete os ganhos alcançados através da aplicação da lógica *fuzzy* e o NRAF resume os ganhos obtidos a partir da aplicação da lógica *fuzzy* e ontologia APA à Macro-Etapa 1.

**TABELA 18 – Avaliações de Risco Químico em Amostras de Leite**

| Amostra | PC<br>(ICA) | PNC<br>(INCA) | NRAF | NRAO | NRA | AA |
|---------|-------------|---------------|------|------|-----|----|
| 14      | 9.71%       | ---           | 3    | 1    | 1   | 1  |
| 4       | 10.83%      | ---           | 1    | 1    | 1   | 1  |
| 7       | 17.92%      | ---           | 1    | 1    | 1   | 1  |
| 13      | 23.54%      | ---           | 1.18 | 1.18 | 2   | 1  |
| 1       | 27.03%      | ---           | 1.35 | 1.35 | 2   | 1  |
| 5       | 31.76%      | ---           | 4    | 1.59 | 2   | 1  |
| 16      | 34.28%      | ---           | 3    | 1.71 | 2   | 1  |
| 3       | 35.91%      | ---           | 1    | 1.79 | 2   | 1  |

|    |        |        |   |      |   |   |
|----|--------|--------|---|------|---|---|
| 15 | 38.13% | ---    | 1 | 1.82 | 2 | 1 |
| 2  | ---    | 2.32%  | 4 | 3    | 3 | 2 |
| 9  | ---    | 4.61%  | 1 | 3.26 | 3 | 2 |
| 10 | ---    | 6.97%  | 1 | 3.57 | 3 | 2 |
| 12 | ---    | 9.30%  | 3 | 3.87 | 3 | 2 |
| 8  | ---    | 11.62% | 3 | 4    | 3 | 2 |
| 11 | ---    | 16.27% | 4 | 4    | 3 | 2 |
| 6  | ---    | 20.93% | 1 | 4    | 3 | 2 |

Gerada na Macro-Etapa 3 da Abordagem Chem-Risk, a figura 45 ilustra os resultados comparativos para 16 amostras de leite analisadas. As áreas achurradas e a reta de equivalência da AA ao NRAO foram adicionadas à figura original:

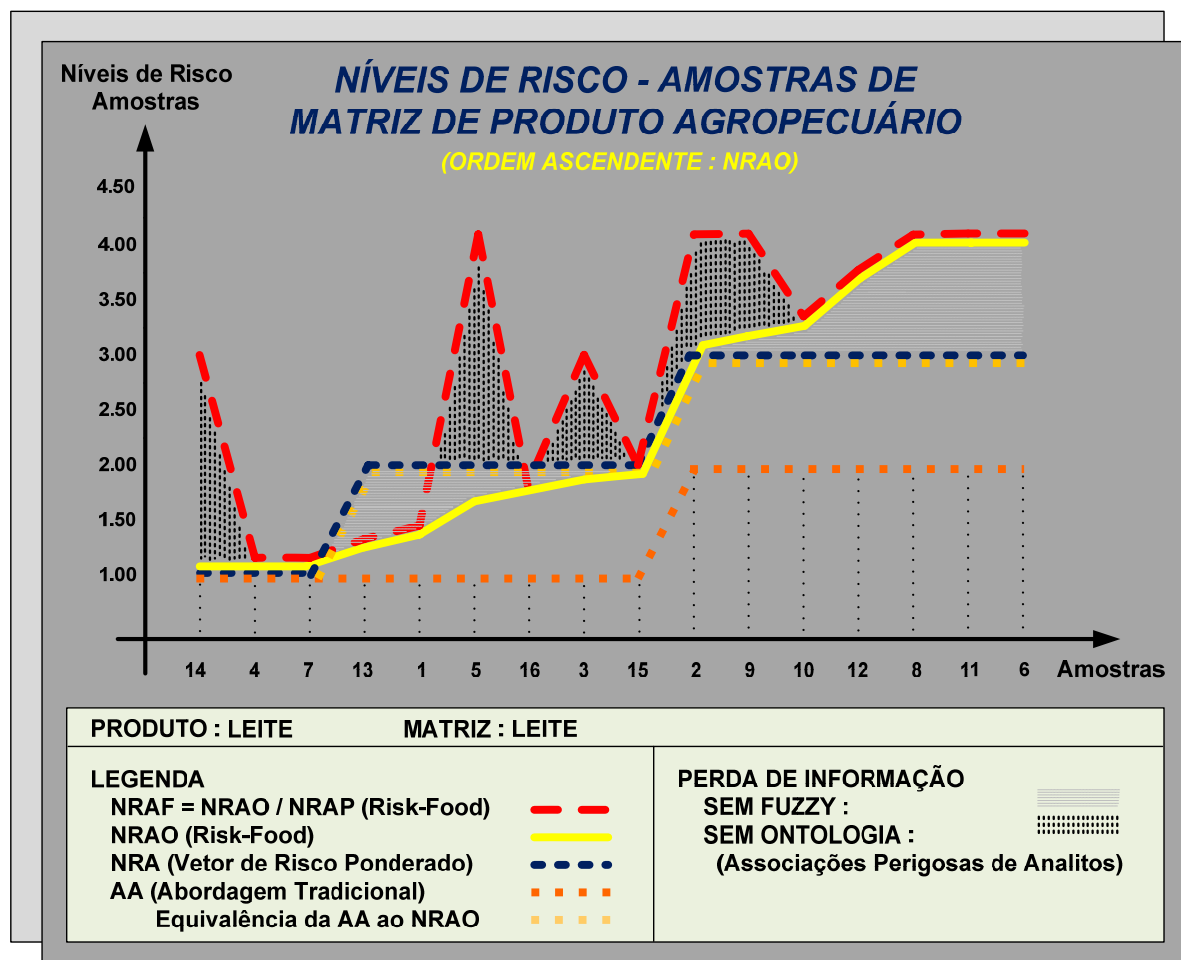


Figura 45 – Resultados Comparativos dos Níveis de Riscos Químicos de Amostras - Abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk

### 8.5.3 Resultados das Avaliações de Risco em Lotes de Amostras

Na terceira bateria de testes foi realizado um estudo comparativo com os níveis de riscos químicos aferidos para os lotes de amostras analisados, considerando os diferentes métodos de avaliação, quais sejam, “Risco do Lote (RL)”, “Nível de Risco do Lote (NRL)” e “Nível de Risco do Lote de Amostras (NRLA)” adotados, respectivamente, pela abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk. Semelhantemente às comparações anteriores os resultados das avaliações nos lotes foram recuperados a partir dos bancos de dados particulares utilizados pelas abordagens.

Associada à figura 46, a tabela 19 apresenta os resultados das avaliações de risco ao nível dos lotes de amostras, armazenados nos bancos de dados particulares das abordagens. Os níveis de risco foram colocados em ordem crescente, separadamente considerados, visando agrupar e melhor evidenciar os ganhos de informação obtidos a partir da incorporação dos conceitos da lógica *fuzzy* e ontologias às avaliações de risco em lotes de amostras, realizadas durante a execução da Macro-Etapa 1 de pré-processamento de dados da Abordagem Chem-Risk.

*Nota.* Observe que o NRLA resume os ganhos obtidos a partir da aplicação da lógica *fuzzy* e ontologia APA à Macro-Etapa 1.

TABELA 19 – Avaliações de Risco Químico em Lotes de Amostras de Mel

| Lote de Amostras | NRLA | NRL | RL |
|------------------|------|-----|----|
| 19               | 1.00 | 1   | 1  |
| 10               | 1.00 | 1   | 1  |
| 1                | 1.00 | 1   | 1  |
| 20               | 1.70 | 1   | 1  |
| 29               | 1.71 | 1   | 1  |
| 11               | 1.71 | 1   | 1  |
| 3                | 2.00 | 2   | 1  |
| 59               | 2.00 | 2   | 1  |
| 42               | 2.00 | 2   | 1  |
| 13               | 3.00 | 3   | 2  |
| 4                | 3.00 | 3   | 2  |
| 226              | 3.80 | 3   | 2  |
| 227              | 3.80 | 3   | 2  |

|     |      |   |   |
|-----|------|---|---|
| 228 | 3.80 | 3 | 2 |
| 54  | 4.00 | 3 | 2 |
| 27  | 4.00 | 3 | 2 |
| 61  | 4.00 | 3 | 2 |

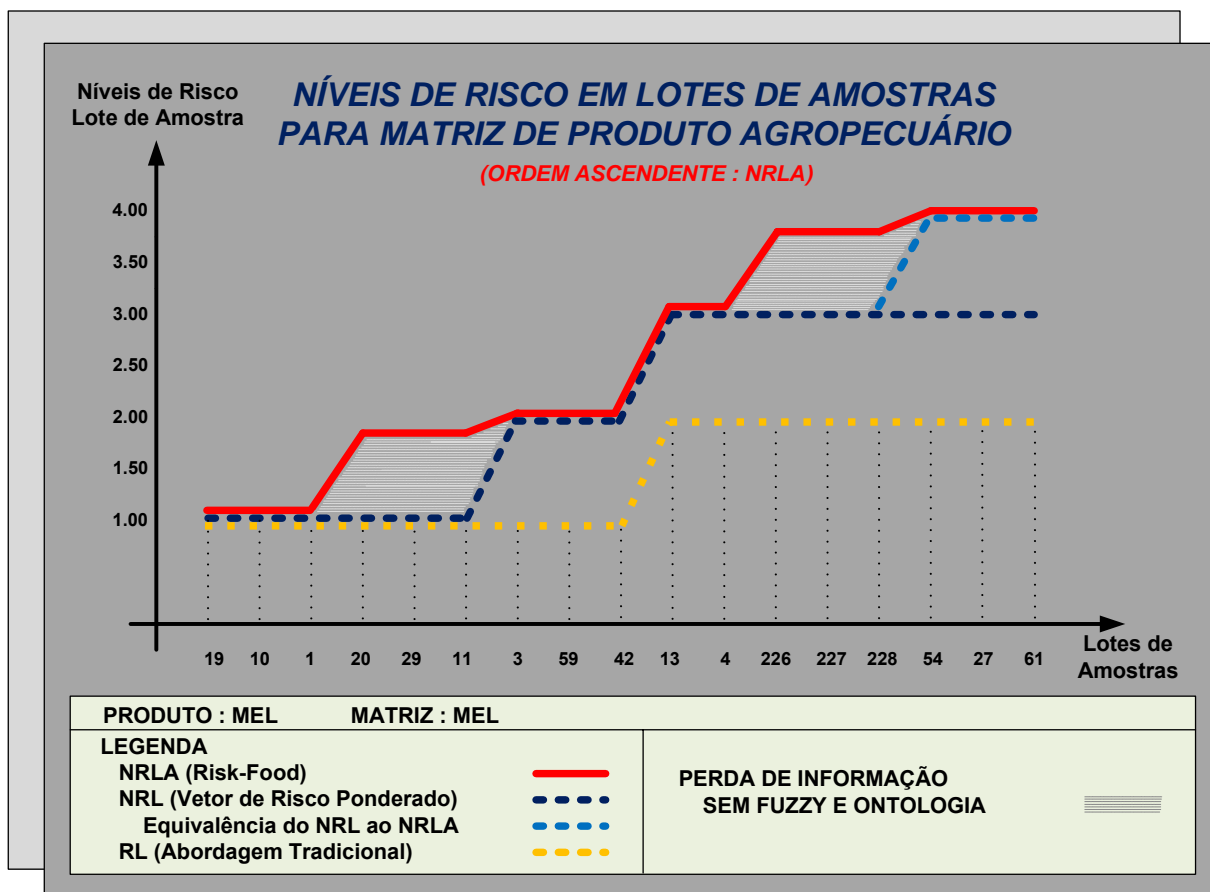


Figura 46 – Resultados Comparativos dos Níveis de Riscos Químicos de Lotes de Amostras - Abordagem Tradicional, Vetor de Risco Ponderado e Chem-Risk

## 8.6 Resultados da Aplicação das Técnicas de Mineração de Dados

Esta seção ilustra uma parte dos resultados obtidos a partir dos procedimentos computacionais realizados na Macro-Etapa 2 da Abordagem Chem-Risk, inerentes à atividade de seleção e formatação de dados a serem submetidos aos processos de mineração de dados. Ressalta-se que, na Macro-Etapa 2, a aplicação das técnicas de mineração sobre os resultados obtidos na Macro-Etapa 1, permite que as autoridades do PNCRC, especialmente os especialistas

de risco, consigam extrair maiores níveis de conhecimentos não-triviais, a partir de um conhecimento superior incorporado aos dados, por meio da aplicação da lógica *fuzzy* e ontologia durante a etapa de pré-processamento dos resultados das análises laboratoriais.

### 8.6.1 Migração dos Níveis de Risco das Amostras

Comparativamente com a abordagem Vetor de Risco Ponderado, observa-se existir uma migração, originada a partir da reclassificação dos níveis de risco das amostras, para as novas faixas de risco (clusters principais) sugeridas pela abordagem Chem-Risk, semanticamente mais fidedignas à realidade mundial da segurança química alimentar. A figura 47 ilustra esta migração:

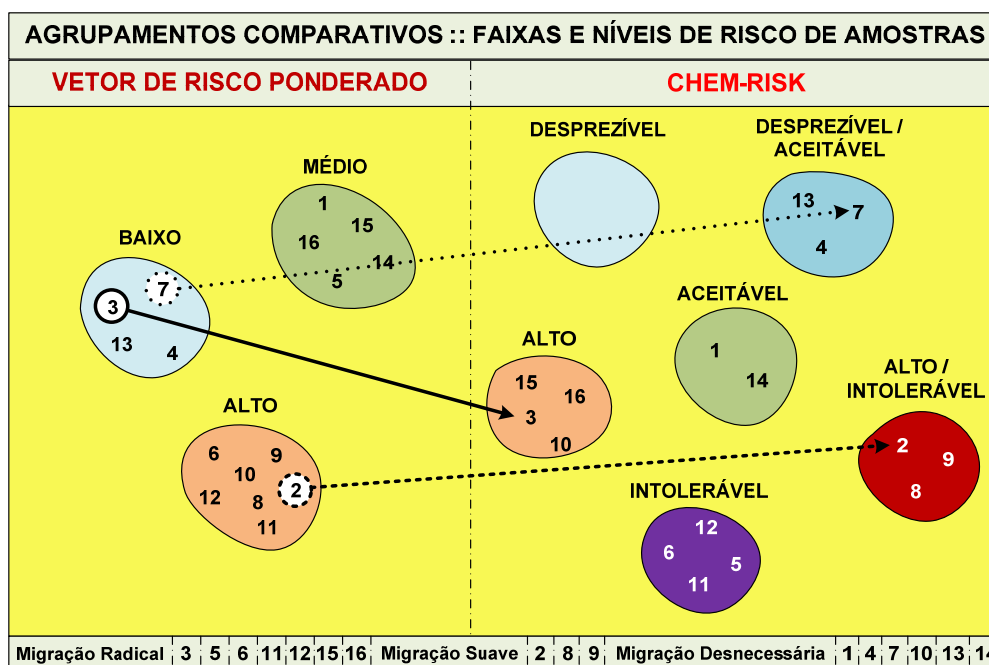


Figura 47 – Migração de Níveis de Riscos de Amostras Aferidos na Abordagem Vetor de Risco Ponderado para os Novos Agrupamentos de Risco Sugeridos pela Abordagem Chem-Risk

Os resultados apresentados a seguir, foram gerados a partir do software Weka 3.6.1 (WITTEN; FRANK, 2005).

### 8.6.2 Aplicação da Técnica de Associação e Associação Preditiva

A técnica de associação de dados tem como objetivo representar padrões interessantes entre itens do domínio de uma aplicação, desde que possam ser verificados com frequência em uma

base de dados. A figura 48 apresenta os resultados originados a partir da aplicação da técnica de associação, por meio do algoritmo *Apriori*, sobre a massa de amostras de leite armazenada no banco de dados Análise de Risco Químico Agropecuário – ARQA, sugerindo a investigação de novas Associações Perigosas de Analitos – APA, além daquelas anteriormente modeladas por meio de ontologias na Macro-Etapa 1 da Abordagem Chem-Risk. Conforme recomendado na abordagem ACATO, as associações perigosas sugeridas, após estudo e aprovação dos especialistas do PNCRC, irão compor o Léxico dos Dados a ser incorporado à ontologia APA.

Associator

Choose **Apriori** -N 200 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1

Start Stop

Associator output

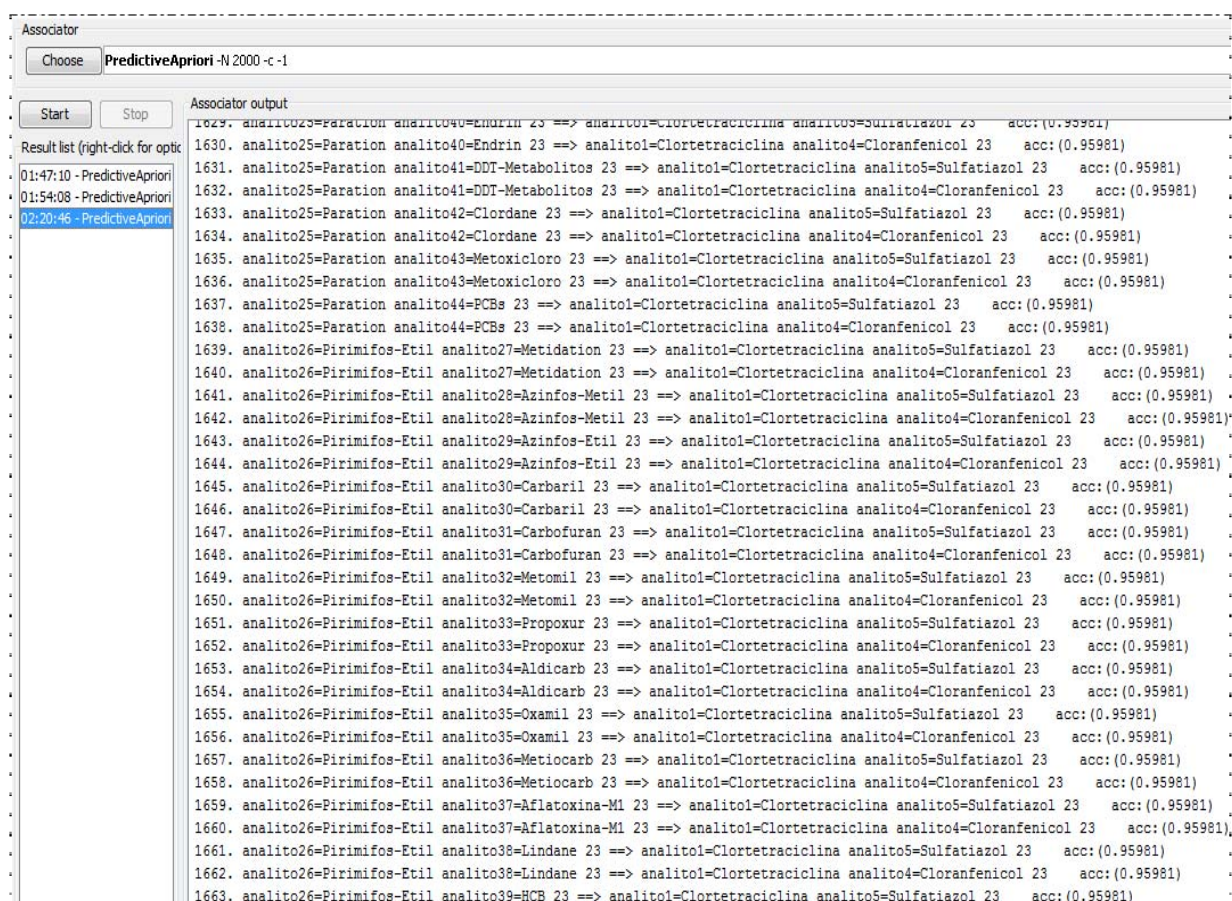
Result list (right-click for c  
08:27:46 - Apriori

|      |  |          |
|------|--|----------|
| 96.  | analito17=Metamidofos analito18=Mevinfos 3 ==> analito9=Mirex 3          | conf:(1) |
| 97.  | analito9=Mirex analito18=Mevinfos 3 ==> analito17=Metamidofos 3          | conf:(1) |
| 98.  | analito9=Mirex analito17=Metamidofos 3 ==> analito18=Mevinfos 3          | conf:(1) |
| 99.  | analito18=Mevinfos 3 ==> analito9=Mirex analito17=Metamidofos 3          | conf:(1) |
| 100. | analito17=Metamidofos 3 ==> analito9=Mirex analito18=Mevinfos 3          | conf:(1) |
| 101. | analito9=Mirex 3 ==> analito17=Metamidofos analito18=Mevinfos 3          | conf:(1) |
| 102. | analito13=Moxidectina analito14=Albendazol 3 ==> analito10=Abamectina 3  | conf:(1) |
| 103. | analito10=Abamectina analito14=Albendazol 3 ==> analito13=Moxidectina 3  | conf:(1) |
| 104. | analito10=Abamectina analito13=Moxidectina 3 ==> analito14=Albendazol 3  | conf:(1) |
| 105. | analito14=Albendazol 3 ==> analito10=Abamectina analito13=Moxidectina 3  | conf:(1) |
| 106. | analito13=Moxidectina 3 ==> analito10=Abamectina analito14=Albendazol 3  | conf:(1) |
| 107. | analito10=Abamectina 3 ==> analito13=Moxidectina analito14=Albendazol 3  | conf:(1) |
| 108. | analito13=Moxidectina analito17=Metamidofos 3 ==> analito10=Abamectina 3 | conf:(1) |
| 109. | analito10=Abamectina analito17=Metamidofos 3 ==> analito13=Moxidectina 3 | conf:(1) |
| 110. | analito10=Abamectina analito13=Moxidectina 3 ==> analito17=Metamidofos 3 | conf:(1) |
| 111. | analito17=Metamidofos 3 ==> analito10=Abamectina analito13=Moxidectina 3 | conf:(1) |
| 112. | analito13=Moxidectina 3 ==> analito10=Abamectina analito17=Metamidofos 3 | conf:(1) |
| 113. | analito10=Abamectina 3 ==> analito13=Moxidectina analito17=Metamidofos 3 | conf:(1) |
| 114. | analito13=Moxidectina analito18=Mevinfos 3 ==> analito10=Abamectina 3    | conf:(1) |
| 115. | analito10=Abamectina analito18=Mevinfos 3 ==> analito13=Moxidectina 3    | conf:(1) |
| 116. | analito10=Abamectina analito13=Moxidectina 3 ==> analito18=Mevinfos 3    | conf:(1) |
| 117. | analito18=Mevinfos 3 ==> analito10=Abamectina analito13=Moxidectina 3    | conf:(1) |
| 118. | analito13=Moxidectina 3 ==> analito10=Abamectina analito18=Mevinfos 3    | conf:(1) |
| 119. | analito10=Abamectina 3 ==> analito13=Moxidectina analito18=Mevinfos 3    | conf:(1) |
| 120. | analito14=Albendazol analito17=Metamidofos 3 ==> analito10=Abamectina 3  | conf:(1) |
| 121. | analito10=Abamectina analito17=Metamidofos 3 ==> analito14=Albendazol 3  | conf:(1) |
| 122. | analito10=Abamectina analito14=Albendazol 3 ==> analito17=Metamidofos 3  | conf:(1) |
| 123. | analito17=Metamidofos 3 ==> analito10=Abamectina analito14=Albendazol 3  | conf:(1) |
| 124. | analito14=Albendazol 3 ==> analito10=Abamectina analito17=Metamidofos 3  | conf:(1) |
| 125. | analito10=Abamectina 3 ==> analito14=Albendazol analito17=Metamidofos 3  | conf:(1) |
| 126. | analito14=Albendazol analito18=Mevinfos 3 ==> analito10=Abamectina 3     | conf:(1) |
| 127. | analito10=Abamectina analito18=Mevinfos 3 ==> analito14=Albendazol 3     | conf:(1) |
| 128. | analito10=Abamectina analito14=Albendazol 3 ==> analito18=Mevinfos 3     | conf:(1) |
| 129. | analito18=Mevinfos 3 ==> analito10=Abamectina analito14=Albendazol 3     | conf:(1) |

Figura 48 – Técnica de Associação Sugerindo Novas Associações Perigosas de Analitos - APA

De acordo com a figura 48, observa-se que as associações *Metamidofós e Mevinfós, Mirex e Mevinfós, Abamectina e Albendazol* foram incorporadas ao Léxico dos Dados.

A proposta do algoritmo *Predictive Apriori* consiste em buscar uma relação entre os valores de suporte e confiança que possam maximizar a chance de uma correta predição de dados não analisados, ou seja, dados futuros ou que ainda não foram analisados no processo de mineração. A figura 49 ilustra sua aplicação em amostras de leite armazenadas no banco de dados ARQA, apresentando 44 analitos pré-determinados pelo *Codex Alimentarius* e PNCRC. Conforme explicado anteriormente, toda associação perigosa sugerida, mesmo na forma de tendência observada, devem ser analisadas e validadas pelos especialistas de domínio do PNCRC. Observa-se a sugestão de inúmeras tendências de associações perigosas para o analito Clortetraciclina:



**Figura 49 – Técnica de Associação Preditiva Sugerindo Inúmeras Tendências de Novas Associações Perigosas de Analitos - APA Envolvendo o Analito Clortetraciclina**



### 8.6.3 Aplicação da Técnica de Classificação

Como vimos na seção 6.4 os algoritmos de mineração que pertencem à tarefa de classificação possuem dois objetivos principais: extrair modelos para descrever os dados e realizar previsões utilizando os modelos criados. A figura 50 ilustra a aplicação do algoritmo *C4.5 (J48 no Weka)*, escolhido em razão da possibilidade de trabalhar com dados contínuos, além de permitir expressar os resultados na forma de árvore de decisão, que facilita a interpretação dos resultados por parte dos especialistas do PNCRC. Os resultados demonstram o alto grau de acerto na classificação dos níveis de risco aferidos para as Concentrações de Analitos (CA), de acordo com as novas faixas de risco trabalhadas na abordagem Chem-Risk:

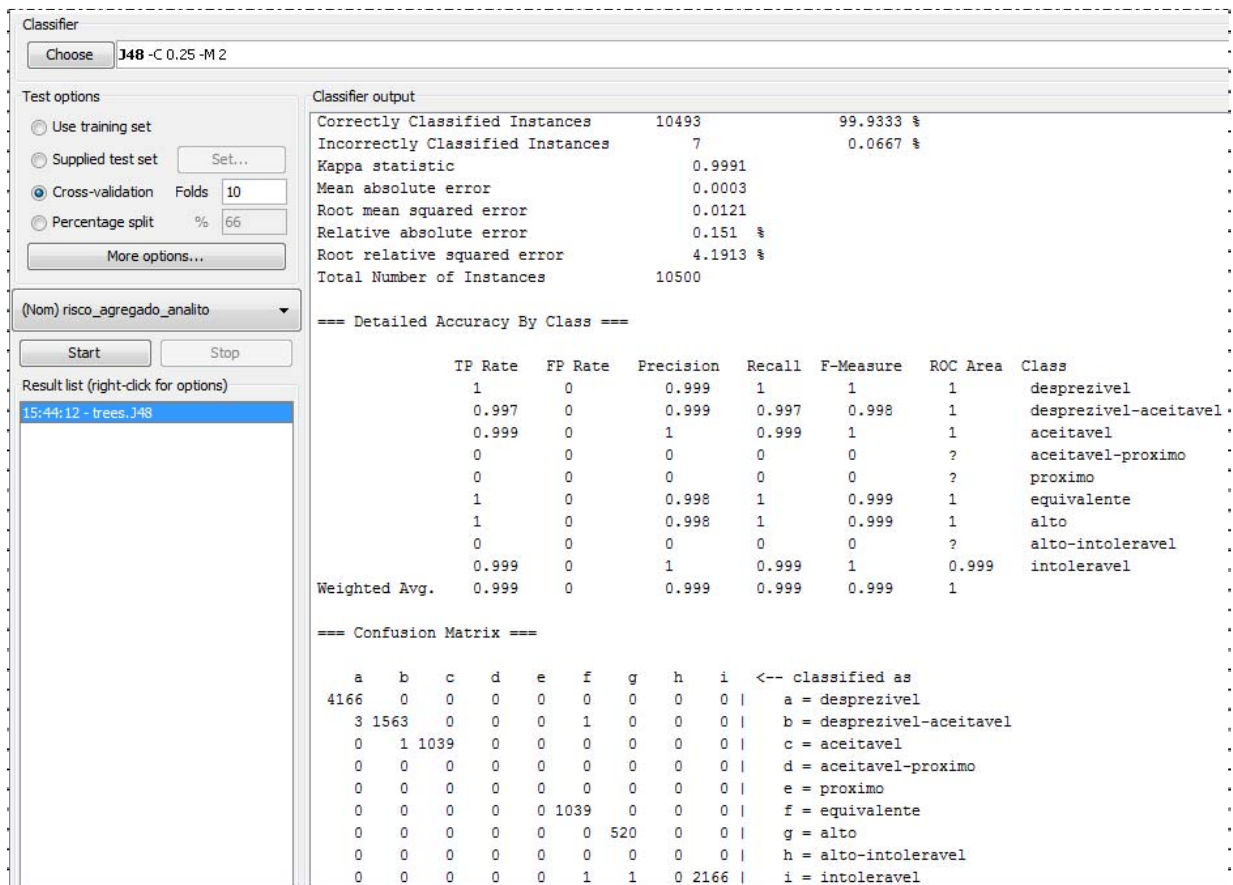
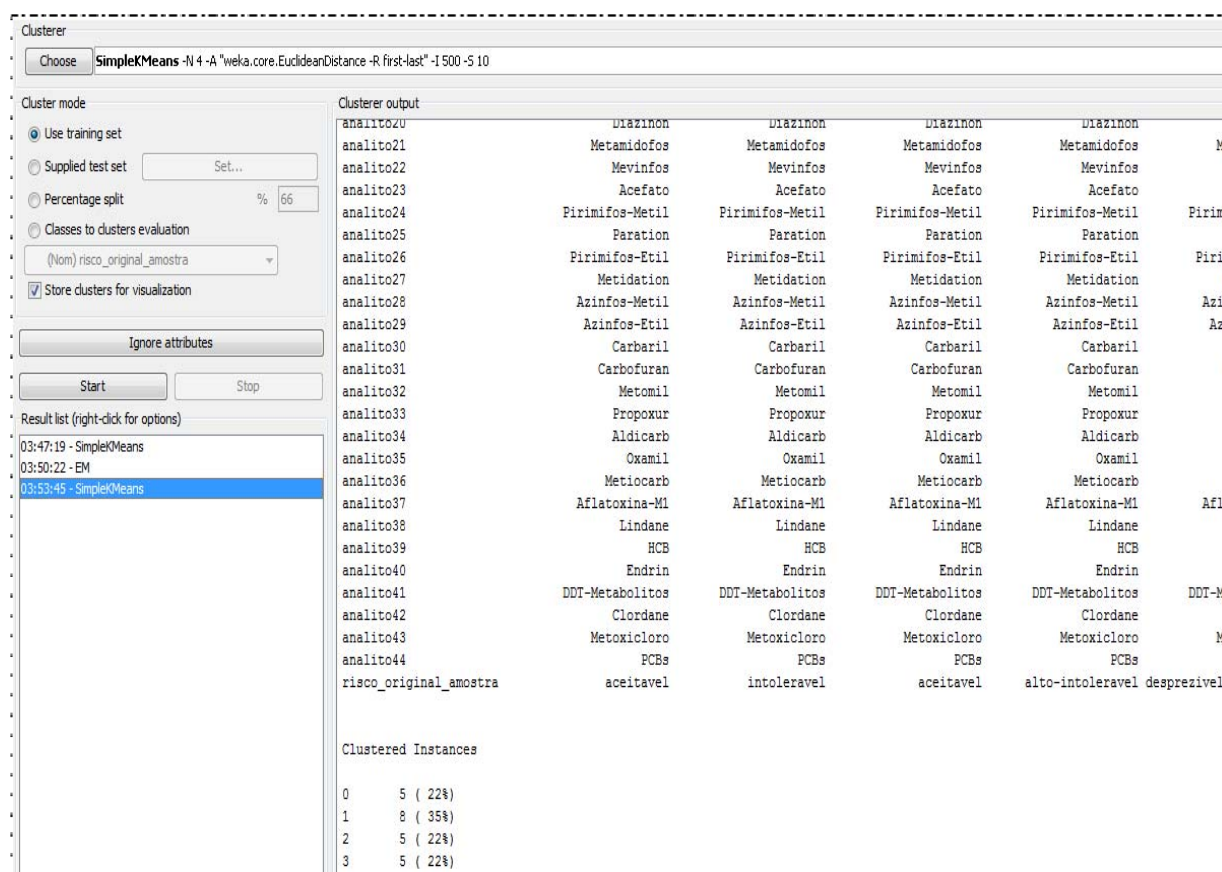


Figura 50 – Resultado da Aplicação da Técnica de Classificação para as Concentrações de Analitos (CA) – Algoritmo J48 (WEKA 3.6.1)

## 8.6.4 Aplicação da Técnica de Agrupamento

Apesar da existência de faixas de risco ou agrupamentos bem definidos pela Abordagem Chem-Risk para os níveis de riscos de analitos, amostras e lotes de amostras, a aplicação das técnicas de agrupamento de dados podem sugerir novas faixas ou agrupamentos a serem avaliados pelos analistas de risco do PNCRC.

A figura 51 apresenta uma distribuição bastante uniforme dos dados de amostras de leite submetidos a um processo de mineração de dados. Sugeridos pelo algoritmo *Simple k-means*, os resultados consideram quatro agrupamentos distintos construídos com base na distância euclidiana dos dados:



| Clusterer              | SimpleKMeans    | -N 4 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10 |                 |                              |
|------------------------|-----------------|--|-----------------|------------------------------|
| analito20              | Diazinon        | Diazinon   | Diazinon        | Diazinon                     |
| analito21              | Metamidofofos   | Metamidofofos  | Metamidofofos   | Metamidofofos                |
| analito22              | Mevinfos        | Mevinfos   | Mevinfos        | Mevinfos                     |
| analito23              | Acefato         | Acefato  | Acefato         | Acefato                      |
| analito24              | Pirimifos-Metil | Pirimifos-Metil  | Pirimifos-Metil | Pirimifos-Metil              |
| analito25              | Paration        | Paration   | Paration        | Paration                     |
| analito26              | Pirimifos-Etil  | Pirimifos-Etil   | Pirimifos-Etil  | Pirimifos-Etil               |
| analito27              | Metidation      | Metidation   | Metidation      | Metidation                   |
| analito28              | Azinfos-Metil   | Azinfos-Metil  | Azinfos-Metil   | Azinfos-Metil                |
| analito29              | Azinfos-Etil    | Azinfos-Etil   | Azinfos-Etil    | Azinfos-Etil                 |
| analito30              | Carbaril        | Carbaril   | Carbaril        | Carbaril                     |
| analito31              | Carbofuran      | Carbofuran   | Carbofuran      | Carbofuran                   |
| analito32              | Metomil         | Metomil  | Metomil         | Metomil                      |
| analito33              | Propoxur        | Propoxur   | Propoxur        | Propoxur                     |
| analito34              | Aldicarb        | Aldicarb   | Aldicarb        | Aldicarb                     |
| analito35              | Oxamil          | Oxamil   | Oxamil          | Oxamil                       |
| analito36              | Metiocarb       | Metiocarb  | Metiocarb       | Metiocarb                    |
| analito37              | Aflatoxina-M1   | Aflatoxina-M1  | Aflatoxina-M1   | Aflatoxina-M1                |
| analito38              | Lindane         | Lindane  | Lindane         | Lindane                      |
| analito39              | HCB             | HCB  | HCB             | HCB                          |
| analito40              | Endrin          | Endrin   | Endrin          | Endrin                       |
| analito41              | DDT-Metabolitos | DDT-Metabolitos  | DDT-Metabolitos | DDT-Metabolitos              |
| analito42              | Clordane        | Clordane   | Clordane        | Clordane                     |
| analito43              | Metoxicloro     | Metoxicloro  | Metoxicloro     | Metoxicloro                  |
| analito44              | PCBs            | PCBs   | PCBs            | PCBs                         |
| risco_original_amostra | aceitavel       | intoleravel  | aceitavel       | alto-intoleravel desprezivel |

Clustered Instances

|   |   |        |
|---|---|--------|
| 0 | 5 | ( 22%) |
| 1 | 8 | ( 35%) |
| 2 | 5 | ( 22%) |
| 3 | 5 | ( 22%) |

Figura 51 – Técnica de Agrupamento de Dados Utilizando o Algoritmo Simple K-means Sugerindo Novos Agrupamentos de Dados para a Abordagem Chem-Risk

## 8.7 Análises estratégicas

Como consequência natural de um adequado processo de preparação de dados, permitido pela Abordagem Chem-Risk, séries históricas com maior precisão poderão ser construídas, apresentando para diferentes níveis de risco incrementais (citados na seção 8.2.1), análises de risco mais fidedignas associadas ao processo de produção dos alimentos agrícolas brasileiros, de acordo com a granularidade desejada para a informação. Nos quadros 3 e 4 encontram-se exemplificadas as análises de risco realizadas para determinado proprietário:

QUADRO 3 - Análise de Risco Periódica: Analitos e Amostras de Leite por Proprietário

| <b>Análise de Risco Periódica por Analito / Amostra</b>      |      |      |      |      |    |       |       |       |       |      |       |
|--|------|------|------|------|----|-------|-------|-------|-------|------|-------|
| <b>Proprietário: Associação dos Produtores de Leite. . .</b> |      |      |      |      |    |       |       |       |       |      |       |
| <b>Localidade: Santo Antonio do Monte - MG</b>               |      |      |      |      |    |       |       |       |       |      |       |
| <b>Exercício: 2010</b>                                       |      |      |      |      |    |       |       |       |       |      |       |
| <b>Nível de Risco Agregado do Analito - NRAA</b>             |      |      |      |      |    |       |       |       |       |      |       |
| Analitos   | Jan  | Fev  | Mar  | Abr  | .. | 1ºTri | 2ºTri | 3ºTri | 4ºTri | .... | Anual |
| Analito 1  | 7.27 | 6.13 | 5.73 | 6.87 | .. | 6.32  |       |       |       | .... | 5.25  |
| Analito 2  | 5.12 | 4.26 | 6.01 | 3.93 | .. | 5.13  |       |       |       | .... | 6.25  |
| ....   | ..   | ..   | ..   | ..   | .. | ....  | ....  | ....  | ....  | .... | ....  |
| <b>Nível de Risco da Amostra Final - NRAF</b>                |      |      |      |      |    |       |       |       |       |      |       |
| Amostras   | Jan  | Fev  | Mar  | Abr  | .. | 1ºTri | 2ºTri | 3ºTri | 4ºTri | .... | Anual |
| Lote 01  | 1.07 | 1.89 | 1.36 | 1.61 | .. | 1.60  |       |       |       | .... | 1.72  |
| Lote 02  | 1.45 | ..   | 1.27 | ...  |    |       |       |       |       |      |       |
| Lote 03  | 1.74 | ...  | 1.91 | .... |    |       |       |       |       |      |       |
| Média  | 1.42 | ...  | 1.51 | ...  |    |       |       |       |       |      |       |

QUADRO 4 – Análise de Risco Acumulada: Analitos e Amostras de Leite por Proprietário

| <b>Análise de Risco Acumulada por Analito / Amostra</b>      |      |      |      |      |      |           |
|--|------|------|------|------|------|-----------|
| <b>Proprietário: Associação dos Produtores de Leite. . .</b> |      |      |      |      |      |           |
| <b>Localidade: Santo Antonio do Monte - MG</b>               |      |      |      |      |      |           |
| <b>Nível de Risco Agregado do Analito - NRAA</b>             |      |      |      |      |      |           |
| Analitos   | 2007 | 2008 | 2009 | 2010 | .... | Acumulado |
| Analito 1  | 5.50 | 4.25 | 4.50 | 5.25 | .... | 4.87      |
| Analito 2  | 3.13 | 3.74 | 4.08 | 3.91 | .... | 3.71      |
| ....   | .... | .... | .... | .... | .... | ....      |
| <b>Nível de Risco da Amostra Final - NRAF</b>                |      |      |      |      |      |           |
| Amostras   | 2007 | 2008 | 2009 | 2010 | .... | Acumulado |
| NRA  | 3.44 | 3.06 | 2.88 | 1.72 | .... | 2.91      |

Técnicas de visualização de dados podem ser empregadas para facilitar o entendimento dos dados armazenados, em menor ou maior granularidade, ilustrando séries históricas ao longo de determinado período. A figura 52 ilustra uma série histórica para as análises de risco desenvolvidas para o analito “Aflatoxina M1” monitorado para a matriz Leite:

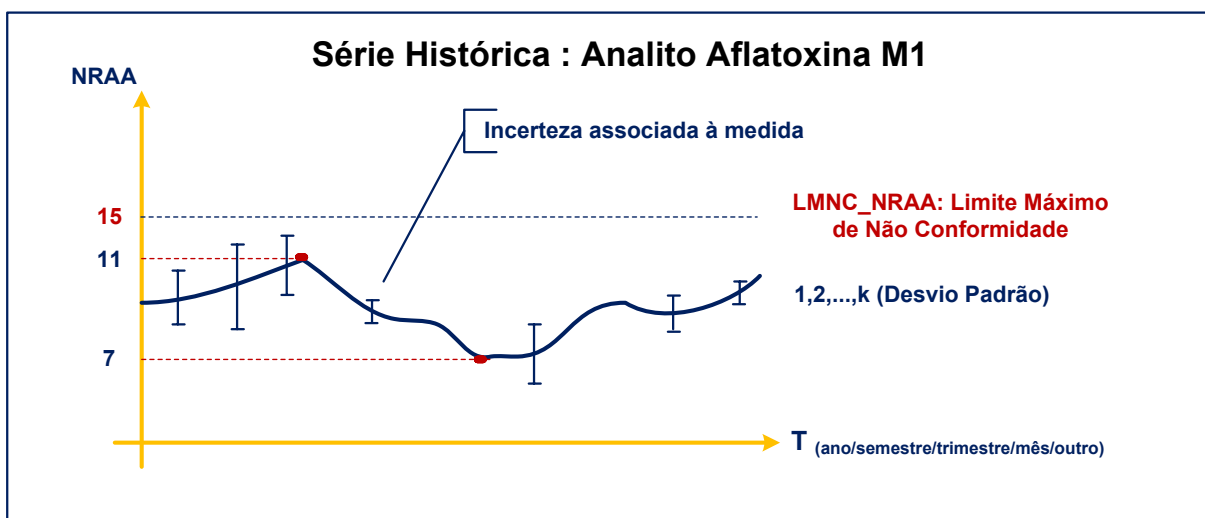


Figura 52 - Ferramenta de Visualização para Série Histórica das Análises de Risco Associadas ao Analito “Aflatoxina M1”

## 8.8 Considerações Finais

Os resultados do estudo de viabilidade envolvendo o Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC / Animal e Vegetal, demonstram a superioridade técnica, qualitativa e quantitativa da Abordagem Chem-Risk, aplicada no contexto das políticas públicas direcionadas à segurança alimentar química mundial.

Com relação à aplicação da lógica *fuzzy*, observa-se que os conjuntos *crisp* tendem a generalizar os níveis de risco nas faixas de risco estabelecidas para analitos e amostras, enquanto que os conjuntos *fuzzy* são mais específicos nesta questão. Paralelamente aos benefícios alcançados a partir da aplicação de ontologias, na computação dos níveis de risco NRCA, NRAA e NRAO a lógica *fuzzy* também contribui para que a Abordagem Chem-Risk ofereça uma semântica mais apropriada ao domínio da segurança química alimentar, permitindo:

- Inferir níveis de risco mais exatos e precisos para o universo das análises de risco incrementais;

- Gerar séries históricas mais exatas a serem utilizadas em avaliações de risco futuras, possibilitando visualizar tendências de comportamento e previsões em geral, de forma mais precisa;
- Representar de forma mais fidedigna a realidade da segurança química produtos agropecuários brasileiros, quando individualmente considerados, ou, em maior nível de abstração, representar a realidade da segurança química vivenciada no dia-a-dia do agronegócio brasileiro.

A ontologia APA permite que a expressividade do domínio “Associações Perigosas de Analitos” seja semanticamente enriquecida por meio da aplicação das Regras de Horn ao nível dos indivíduos da ontologia, permitindo estender o contexto maior da análise de riscos químicos agropecuários desenvolvido pelo PNCRC / MAPA. Tecnicamente, a existência da ontologia APA é uma pré-condição para o fornecimento de novos e avançados serviços oferecidos pelo PNCRC, como, por exemplo, a busca e recuperação semântica das associações perigosas de analitos nas matrizes agropecuárias analisadas.

A qualidade dos resultados obtidos durante a etapa de mineração de dados comprovam a existência de um pré-processamento de dados adequado, disponibilizando dados mais exatos e precisos que apresentam uma semântica mais fidedigna ao domínio da segurança química dos alimentos agropecuários. Por meio das técnicas de associação, associação preditiva, agrupamento e classificação empregadas, os resultados da etapa de mineração comprovam e complementam a realidade da segurança química dos alimentos agrícolas revelada pela Abordagem Chem-Risk. Os resultados ampliam o nível de conhecimento especializado das autoridades e especialistas de risco do PNCRC, favorecendo a existência de um processo de tomada de decisões não triviais e inéditas no âmbito do MAPA. Por outro lado, conforme comumente observado nos diferentes contextos da aplicação das técnicas de mineração, constata-se que, caso não houvesse adequado pré-processamento de dados a nível semântico, os resultados obtidos durante a etapa de mineração, apesar de aparentemente corretos, estariam encobrendo bases falsas e inadequadas, gerando, por esta razão, conhecimentos completamente equivocados perante a realidade mundial da segurança química dos alimentos agrícolas.

# Capítulo 9

## Conclusões

---

### 9.1 Considerações Iniciais

Esta pesquisa desenvolveu uma nova abordagem, denominada Chem-Risk, para avaliação de riscos químicos agropecuários a ser empregada no contexto das políticas públicas que visam maximizar a segurança alimentar química de alimentos “in natura”. A hipótese inicial foi que a incorporação de técnicas administrativas inerentes à análise de risco e técnica de Kepner-Tregoe, e, principalmente, a incorporação de técnicas computacionais inerentes à lógica *fuzzy* e ontologias associadas às técnicas do projeto OMCS-Br e mineração de dados, pudessem contribuir para aprimorar a segurança alimentar química dos produtos agropecuários, proporcionando melhorias na qualidade dos dados e resultados alcançados, quais sejam: maior exatidão, precisão e fidedignidade à realidade química vivenciada por sistemas de produção agropecuária desenvolvidos em países membros do Codex Alimentarius - FAO/OMS/ONU. O objetivo desta pesquisa visou fornecer dados e resultados aprimorados para que especialistas, autoridades e tomadores de decisão sejam capazes de qualificar e quantificar o impacto das práticas e intervenções governamentais ao longo da cadeia produtiva. Os resultados alcançados almejam apoiar decisões estratégicas que tenham como objetivo principal reduzir os perigos químicos associados aos processos de produção alimentar, auxiliando a criar uma inteligência de negócio que ofereça maiores níveis de conhecimentos não-triviais a especialistas, tomadores de decisão e autoridades governamentais.

A Plataforma Base Flexível é o coração desta abordagem e determina uma sequência de passos para adequadamente promover avaliações de riscos químicos incrementais, partindo de uma estimativa de riscos ao nível dos analitos e chegando ao nível do risco químico agropecuário aferido para um país. Sua característica essencial é ser flexível de forma a permitir a incorporação de novos fatores de risco ao nível dos analitos, amostras e lotes de amostras, além de permitir a incorporação de novos analitos ao grupo de analitos pré-determinados para as matrizes agropecuárias, tanto pela Comissão do Codex Alimentarius e como pelo PNCRC / MAPA.

A lógica *fuzzy* permite construir séries históricas mais fidedignas e, portanto, mais confiáveis, através do aprimoramento da qualidade dos dados trabalhados na etapa de pré-processamento, especificamente com relação aos adjetivos exatidão, precisão, fidedignidade e abrangência. Consegue-se, assim, que os resultados aferidos na etapa de mineração de dados também estejam mais exatos, fidedignos e realistas.

A ontologia é pré-condição para representar semanticamente as associações perigosas de analitos, padronizar o vocabulário utilizado neste domínio de conhecimento, reduzir a dependência dos especialistas, a subjetividade, o tempo e custo das análises desenvolvidas, além de habilitar análises complexas comumente desprezadas que comprometem a qualidade dos dados e resultados finais. Através do senso comum (OCMS-Br) é possível contar com a estimada participação de numerosos especialistas em centros de pesquisa, universidades, instituições públicas relacionadas à segurança alimentar química, peritos agropecuários, entre outros especialistas. A incorporação das técnicas deste projeto possibilita estender e enriquecer significativamente as associações perigosas de analitos, a modelagem semântica das avaliações de riscos químicos, as técnicas de análise empregadas e o vocabulário composto por termos técnicos usados na ontologia APA. Adicionalmente, os resultados alcançados na etapa de mineração de dados permitem complementar os termos e associações perigosas gerenciados pela ontologia APA.

## **9.2 Sinopse dos Resultados Alcançados pela Abordagem Chem-Risk**

Com base no estudo de viabilidade desenvolvido a partir do Plano Nacional de Controle de Resíduos e Contaminantes – PNCRC / Animal e Vegetal, constata-se que ao aprimorar a qualidade dos dados e resultados, a abordagem produz um ganho de informação, originado a partir da aplicação das técnicas administrativo-computacionais, que conjuntamente os torna mais fidedignos à realidade da segurança química alimentar mundial. No contexto das políticas públicas, esse ganho de informação representa um diferencial qualitativo superior que contribui para que as autoridades governamentais possam alcançar o triângulo virtuoso “*Dados-Conhecimento-Inteligência*”, apoiando decisões estratégicas que visam reduzir os riscos químicos associados aos processos de produção e comercialização de alimentos “in natura”. Por meio do diferencial qualitativo, as autoridades podem agora mudar os rumos do processo de

produção e comercialização dos alimentos agrícolas, impedindo, em muitas ocasiões, que diversos alimentos cheguem à mesa dos consumidores brasileiros e mundiais, os quais, anteriormente, circulavam livremente pelos mercados consumidores apresentando qualidades químicas indesejáveis e ocultas aos métodos tradicionais utilizados para avaliação de riscos.

Por fim, com base nos resultados alcançados, demonstrados no capítulo 8, pode-se afirmar que a abordagem Chem-Risk contribui para que os programas nacionais de controle alimentar estejam alinhados à realidade mundial da segurança alimentar química, às necessidades prioritárias da população local e mundial e, assim, em favor da saúde humana. Por esta razão, afirmamos anteriormente que o que torna nosso estudo de caso mais notável é que ele ajuda a melhorar a saúde da população brasileira e mundial e, por extensão, ajuda a salvar nossas próprias vidas.

*O conhecimento é um estágio da sabedoria. Classicamente sua origem está nos dados, que se convertem em informação, que quando adquirida e acumulada se torna um conhecimento, que sendo qualificado é consolidado como competência (que não é autoproclamada e sim conferida por terceiros). É o conhecimento (saber) no sentido amplo que constitui nosso mais importante ativo e, além de ser uma conexão com o passado, uma forma de capital e de relação com o poder, sua aquisição é um investimento obrigatório no futuro, uma memória e, notadamente, uma esperança social sem limitações ao seu avanço e ao acesso democrático. O conhecimento depende de obstinada dedicação, sendo transmissível, mas nunca transferível. É alvo permanente de mudanças oriundas de variadas fontes e, por esta razão, universidades e demais instituições precisam enfrentar os rápidos avanços e obsoletismos, entendendo a necessidade de novas ações e do convencimento de setores que são mais resistentes à modernidade, mesmo quando cuidadosamente preservadas a tradição, a qualidade a identidade institucional. Deve, principalmente, integrar áreas que estão monopolizando especialidades e construir novos caminhos para que o conhecimento interdisciplinar ou translacional deixe de ser apenas um discurso propositivo ainda com poucos resultados e obtenha boas idéias e transformações inovadoras, que são indispensáveis e impactantes para o crescimento social e econômico.*

*Flavio Fava de Moraes*

*Ex-reitor da USP (1993 – 1997)*

*Jornal da Fundação de Faculdade de Medicina (Dez/2010)*



### 9.3 Trabalhos Futuros

Existe muito trabalho e pesquisa a serem conduzidos, especialmente com relação aos processos de mineração de dados que podem ser melhor estendidos e enriquecidos com o passar do tempo. Indubitavelmente, em futuro próximo, os resultados de mineração de dados irão contribuir de forma ainda melhor para aprimorar os resultados alcançados pela Plataforma Base Flexível e, assim, enriquecerem o processo de avaliação de riscos químicos realizado em produtos agropecuários, especialmente com relação às associações perigosas de analitos semanticamente modeladas pela ontologia APA.

Estima-se também que os resultados da mineração de dados passem a contribuir com o processo de estimativa dos fatores de risco tendência da concentração do analito, prazos e custos estimados de adaptação do sistema produtivo aos índices de segurança alimentar química, pré-determinados pelo Codex Alimentarius e PNCRC. Estuda-se a possibilidade de estender a ontologia para atender a esta finalidade.

Por outro lado, o armazenamento contínuo de dados no banco de dados Análise de Risco Químico Agropecuário – ARQA tecnicamente sugere a construção de um modelo multidimensional de dados, construído a partir dos dados aprimorados pela Abordagem Chem-Risk, formando, assim, um Data Warehouse (KIMBALL; ROSS, 2002) com capacidade para otimizar a geração de análises estratégicas diferenciadas, solicitadas por especialistas de risco, autoridades e tomadores de decisão do PNCRC / MAPA. A figura 53 apresenta um esboço inicial, a ser aprimorado, de um esquema estrela para o Data Warehouse ARQA:

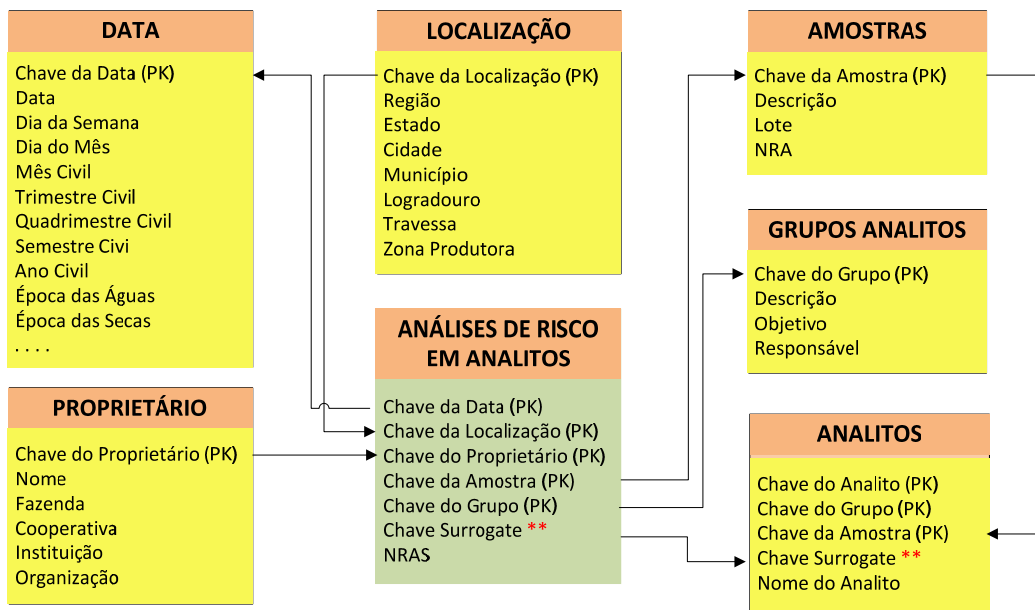


Figura 53 - Exemplo de Esquema Estrela Inicial para o Data Warehouse - ARQA

## Referências

---

ALENCAR, A. J.; SCHMITZ, E. A. **Análise de Risco em Gerência de Projetos**. Rio de Janeiro: Ed. Brasport, 2006. 196 p.

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Revista Ciência da Informação**, v. 32, n. 3, p. 7-20, 2003.

ANACLETO J. C. et al. **Machines with good sense: How can computers become capable of sensible reasoning?** In: Proc. of WCC, p. 1-10, 2008. In press.

BOBILLO F. et al. A semantic fuzzy expert system for a fuzzy balanced scorecard, **Expert Systems with Applications: An International Journal**, v. 36, p. 423-433, Ed. Pergamon Press, New York – USA, 2009.

BONNET, M.; SILVA, A. E. **Núcleo de Análise de Risco Agropecuário: Projeto NARAGRO**. Brasília, DF: Secretaria de Defesa Agropecuária / MAPA, 2007.

BORGES, K. A. V. **Uso de uma Ontologia de Lugar Urbano para Reconhecimento e Extração de Evidências Geo-espaciais na Web**. 2006. 181 f. Tese (Doutorado) - Universidade Federal de Minas Gerais, Belo Horizonte, MG

BRAGA, L. P. V. **Introdução à Mineração de Dados**. 2. ed. Rio de Janeiro: E-Papers Serviços Editoriais Ltda., 2005. 202 p.

BURTON-JONES et al., **A semiotic metrics suite for assessing the quality of ontologies**. Data & Knowledge Engineering, Volume 55, Issue 1, Ed. Elsevier, 2005, Pages 84-102

CABRAL, J. R. A.; VALENTE, M. E. R. **Rastreabilidade na indústria de laticínios: relatório parcial**. Viçosa: UFV, 2008. 50 p. (Projeto final de curso).

CHAKRABART, S. et al. **Data Mining Know it all**. Burlington, MA: Morgan Kaufmann, 2009. 460 p.

CHAPMAN, P. et al. **CRISP-DM Step-by-Step Data Mining Guide**. Crisp-dm homepage, 2000. Disponível em: <<http://www.crisp-dm.org/>>. Acesso em: 20 fev. 2010.

CLARK, D. Mad cows, metathesaurim and meaning. **IEEE Intelligent Systems**, v. 14, n. 1, p. 75-77, jan. 1999.

COSTA, T.; GUSMÃO, C. Definição de Ontologia para Identificação de Riscos de Projetos de Software – **Seminário de Pesquisa em Ontologia no Brasil**, 2008 - Universidade Federal Fluminense, Niterói, RJ.

CULLOT, N. et Al. **Ontologies: A contribution to the DL/DB debate**. Proc. of the 1st International Workshop on the Semantic Web and Databases, 29th International Conf. on Very Large Data Bases, 2003, 109-129.

D'AMORE, R.; SAOTOME, O.; KIENITZ, K. H. Controlador Nebuloso com Detecção de Regras Ativas. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 3., 1997, Vitória. **Anais...** Vitória, 1997. p. 313-318.

DE NICOLA, A.; MISSIKOFF, M.; NAVIGLI R. **A software engineering approach to ontology building**, 2009 DOI=10.1016/j.is.2008.07.002.

ELMARSI, R.; NAVATHE, S. B. **Sistemas de Banco de Dados**. 4. ed. São Paulo: Pearson Education do Brasil, 2005. 724 p.

ESCOVAR, E. L. G.; BIAJIZ, M.; VIEIRA, M. T. P. SSDM: A Semantically Similar Data Mining Algorithm. In: BRAZILIAN SYMPOSIUM OF DATABASES, 20., 2005, Uberlândia, MG. **Anais...** Uberlândia, 2005. 15 p.

FALBO, R.; RUY, F.; PEZZIN J.; MORO, R. D. **Ontologias e ambientes de desenvolvimento de software semântico**. In: JORNADAS IBEROAMERICANAS DE INGENIERIA DEL SOFTWARE E INGENIERIA DEL CONOCIMIENTO, 4., 2004, Madrid. **Actas...** Madrid : Servicio de Publicaciones de la Facultad de Informática de la U.P.M., 2004. v. 1, p. 277–292.

FAOSTAT database. FAO, Rome, 2010. Available at: <<http://faostat.fao.org/>>. Acesso em: 27 jan. 2010.

FAYYAD, U. et al. From Data Mining to Knowledge Discovery: an Overview. In: FAYYAD, U. et al. (Ed.). **Advances in Knowledge Discovery and Data Mining**. Cambridge, Mass.: AAAI/MIT Press, 1996. p. 1-30.

FERNANDEZ, M.; GOMEZ-PÉREZ, A.; JURISTO, H. **Methontology: from ontological art towards ontological engineering**, 1997. Disponível em: <<http://citeseer.ist.psu.edu/context/544607/0/>>. Acesso em: 20 jan. 2010.

GILCHRIST, Alan. Thesauri, taxonomies and ontologies - an etymological note. *Journal of Documentation*, v.59, n. 1, 2003, p. 7-18. GOMES-PÉREZ, A. Tutorial on Ontological Engineering, In: INTERNACIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE – IJCAI, 1999, Estocolmo. **Anais...** Estocolmo, 1999.

GOMES-PÉREZ, A.; FERNÁNDEZ-LOPES, M.; CORCHO, O. **Ontological Engineering**. London: Springer-Verlag, 2003.

GRUBER, T. R. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. **International Journal of Human and Computer Studies**, v. 43, n. 5-6, p. 907-928, 1995.

GRUBER, T. R. **Ontology in the Encyclopedia of Database Systems**. London: Springer-Verlag, 2009.

GUARINO, N. Formal Ontology and Information Systems. In: INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGIES IN INFORMATION SYSTEMS, 1998, Trento, Italy. **Proceedings..** Amsterdam: IOS Press, 1998. p. 3-15.

HAIR, J. F. et al. **Análise multivariada de dados**. São Paulo: Bookman. Cap. 3, p. 23-45, 2005.

HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. 2. ed. San Francisco: Elsevier / Morgan Kaufmann, 2006. 800 p.

HORROCKS, I. Ontologies and the semantic web. **Communications of the ACM**, ACM, New York, NY, USA, v. 51, p. 58–67, 2008.

HORROCKS, I. et al. **SWRL: A Semantic Web Rule Language Combining OWL and RuleML**, 2004. Available at: <<http://www.w3.org/Submission/SWRL/>>.

IBGE. **Censo Agropecuário 2006**: resultados preliminares. Available at: <<http://www.ibge.gov.br/home/estatistica/economia/agropecuaria/censoagro/2006/default.shtm>>. Acesso em: 07 April 2008.

KEPNER, C. H.; TREGOE, B. B. **O Novo Administrador Racional**. São Paulo: Makron Books, 1991.

KIMBALL R.; ROSS, M. **The Data Warehouse Toolkit**. Rio de Janeiro: Campus, 2002. 494 p.

KLIR, G. J.; YUAN, B. **Fuzzy Sets and Fuzzy Logic: Theory and Applications**. Upper Saddle River: Prentice Hall, 1995.

LEITE, M. A. A. **Modelo fuzzy para recuperação de informação utilizando múltiplas ontologias relacionadas**. 2009. 183 f. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, SP.

LIU H.; SINGH H. P. **ConceptNet: A Practical Common Sense Reasoning Toolkit**. BT Technology Journal, v. 22, n. 4, 211-226.

MARASCHIN, L. **Avaliação do grau de contaminação por pesticidas na água dos principais rios formadores do pantanal mato-grossense**. 2003. 90f. Tese (Doutorado) –Universidade Federal de Mato Grosso, Cuiabá, MT.

PRADO, H. A. **Orpheo: uma estrutura de trabalho para integração dos paradigmas de aprendizado supervisionado e não-supervisionado**. 2001. 154 f. Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS.

NARDI, D.; BRACHMAN, R. J. An Introduction to Description Logics. In: BAADER, F. et al. **The Description Logic Handbook**. New York: Cambridge University Press, 2003. p. 5-44.

NEVES, M. F.; CÔNSOLI, M. A. **Estratégias para o leite no Brasil**. São Paulo: Atlas/Pensa, 2006. 304 p.

NOY, N. F. **Semantic integration: a survey of ontology-based approaches**. SIGMOD Record, ACM Press, 2004, 33, 65-70.

NOY, N. F.; McGUINNESS, D. L. **Ontology Development 101: A Guide to Creating your First Ontology**, Stanford Medical Informatics Technical Report SMI. 2001. 7 p.

O'CONNOR, M. J.; DAS. A. **SQWRL: a Query Language for OWL**, OWL: Experiences and Directions (OWLED 2009), Fifth International Workshop, Chantilly, VA, 2009.

OLIVEIRA, K. R. **AdequaS: Ferramenta Fuzzy para Avaliação da Qualidade de Software**. 2002. 117 f. Tese (Mestrado) – Universidade de Fortaleza, Fortaleza, CE.

OLSON, D. L.; DELEN, D. **Advanced data Mining Techniques**. Berlin, Heidelberg: Springer-Verlag, 2008. 182 p.

PRADO, H. A. **Orpheo**: uma estrutura de trabalho para integração dos paradigmas de aprendizado supervisionado e não-supervisionado. 2001. 154 f. Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS.

SANDRI, S.; CORREA, C. Lógica Nebulosa. In: ESCOLA DE REDES NEURAIAS, 5., 1999, São José dos Campos. **Anais...** São José dos Campos: ITA, 1999. p. c073-c090.

SEMPREBOM, T.; CAMADA M. Y.; MENDONÇA, I. **Ontologias e Protegé**. 2007. Programa de Pós-graduação em Engenharia de Automação e Sistemas – Universidade Federal de Santa Catarina, Florianópolis, SC.

SCHUBERT, J.; RILEY, E.J.; TYLER, S.A. **Efeitos combinados em Toxicologia - A Sistemática Rapid** Procedimento de Teste: cádmio, mercúrio e chumbo, *Toxicologia e Saúde Ambiental*, 4: 763, 776 p., 1978.

SILVA, D. L. da; SOUZA, R. R.; ALMEIDA, M. B. Ontologias e vocabulários controlados: comparação de metodologias para construção. / *\*Ciência da Informação\**, Brasília, v. 37, n. 3, p. 60-75, 2008. Disponível em: <<http://www.scielo.br/pdf/ci/v37n3/v37n3a05.pdf>>

SKOGLUND, T.; DEJEMEK, P. Fuzzy traceability: a process simulation derived extension of the traceability concept in continuous food processing. – **Food and Bioproducts Processing** Volume 85, Issue 4, 2007, p 354-359.

SMITH, M. K.; WELTY, C.; MCGUINNESS, D. L. **W3C Proposed Recommendation: OWL Web Ontology Language Guide**. Disponível em: <<http://www.w3.org/TR/2004/REC-owl-guide-20040210>>. Acesso em: 20 fev. 2010>

TAN, P.; STEINBACH, M.; KUMAR, V. **Introdução ao DATAMINING Mineração de Dados**. Rio de Janeiro: Ciência Moderna, 2009. 900 p.

USCHOLD, M.; GRÜNINGER, M. Ontologies and semantics for seamless connectivity. **SIGMOD Record**, v.33, n. 4, p. 58-64, 2004.

YAGUINUMA, C. A. **Sistema FOQuE para Expansão Semântica de Consultas baseada em Ontologias Difusas**. 2007. 114 f. Tese (Mestrado) – Universidade Federal de São Carlos, São Carlos, SP.

WITTEN, I. H.; FRANK, E. Transformations: Engineering the input and output. In: WITTEN, I. H.; FRANK, E. **Data Mining**: practical machine learning tools and techniques with Java implementations. San Francisco: Elsevier / Morgan Kaufmann, 2005. Cap. 7, p. 288-343.

ZADEH, L. A. Fuzzy Sets. **Fuzzy Sets, Information and Control**, v. 8, p. 338-353, 1965.

ZADEH, L. A. Fuzzy Sets as a Basis for a Theory of Possibility. **Fuzzy Sets and Systems**, v. 1, p. 3-28, 1978.

ZADEH, L. A. Fuzzy Sets. In: YAGER, R. R. et al. **Fuzzy sets and applications**: Selected Papers by L.A. Zadeh. New York: Wiley-Interscience, 1987a. p. 29-44

ZADEH, L. A. Similarity Relations and Fuzzy Orderings. In: YAGER, R. R. et al. **Fuzzy sets and applications**: Select Papers by L. A. Zadeh. New York: Wiley-Interscience, 1987b. p. 81-104.



# Apêndice A

## Banco de Dados - ARQA

---

### A.1 Considerações Iniciais

Este apêndice descreve alguns tópicos essenciais relacionados ao projeto, criação e utilização do banco de dados Análise de Risco Químico Agropecuário – ARQA. Como parte do modelo conceitual, a seção A.2 apresenta o diagrama ER utilizado para modelar o banco de dados ARQA; a seção A.3 apresenta o esquema conceitual utilizado para descrever as tabelas, relacionamentos e atributos do banco de dados ARQA. A seção A.4 apresenta o *script* utilizado para criação das tabelas, relacionamentos e atributos neste banco de dados, e, finalmente, a seção A.5 discute em maiores detalhes a *Análise das Atualizações Estruturais* apresentando informações e restrições essenciais ao comportamento dos programas que compõem a Plataforma Base Flexível de pré-processamento, os quais interagem em sincronismo com os dados armazenados no banco de dados ARQA, de acordo com o diagrama ER ilustrado na seção A.2.

Em virtude de sua robustez, da existência de inúmeros *plugins*, da sua abordagem como software livre e da sua larga utilização no mercado da pesquisa e profissional, o Sistema Gerenciador de Banco de Dados (SGBD) *PostgreSQL versão 8.4* foi escolhido para gerenciar as ações deste banco de dados.

### A.2 Diagrama ER

Segundo os autores (ELMARSI; NAVATHE, 2005) a modelagem conceitual é uma fase muito importante no planejamento de uma aplicação de banco de dados que, no caso desta pesquisa, corresponde ao banco de dados Análise de Risco Químico Agropecuário – ARQA e aos programas associados que implementam as consultas, atualizações, etc. Como parte integrante do sistema de informações utilizado nesta pesquisa, o diagrama E-R apresentado na figura A.1 descreve o modelo de dados utilizado para modelar conceitualmente o processo de Análise de Riscos Químicos Agropecuários desenvolvido pelo Ministério da Agricultura, Pecuária e Abastecimento – MAPA:

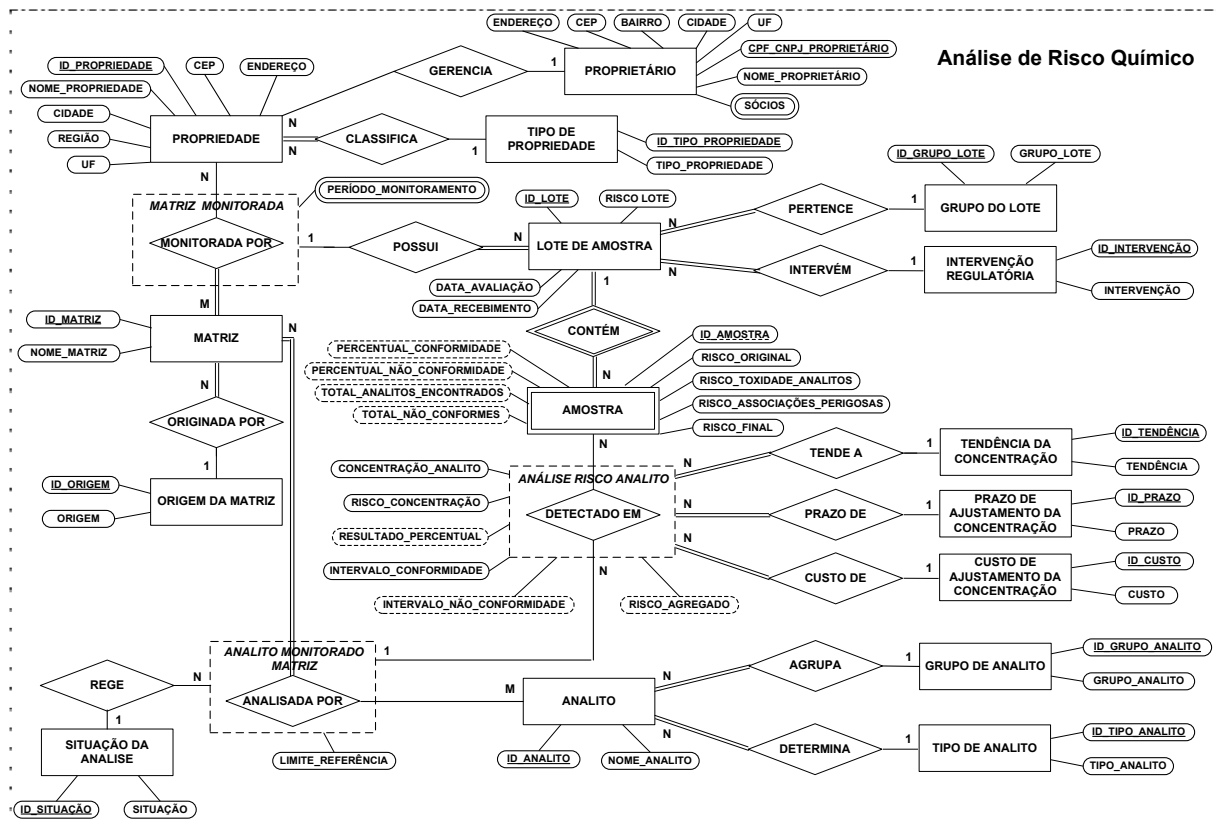


Figura A.1 - Diagrama ER para o Banco de Dados Análise de Risco Químico Agropecuário - ARQA

### A.3 Esquema Conceitual

A figura A.2 apresenta o esquema conceitual do banco de dados Análise de Risco Químico Agropecuário, de acordo com o diagrama ER apresentado na figura A.1:

propriedade (ID\_propriedade, nome\_propriedade, endereço\_propriedade, CEP\_propriedade, cidade\_propriedade, UF\_propriedade, região\_propriedade,  
 FK\_tipo\_propriedade(ID\_tipo\_propriedade), FK\_proprietário(CPF\_CNPJ\_proprietário))  
 tipo\_de\_propriedade (ID\_tipo\_propriedade, tipo\_propriedade)  
 proprietário (CPF\_CNPJ\_proprietário, nome\_proprietário, endereço\_proprietário,  
 CEP\_proprietário, bairro\_proprietário, cidade\_proprietário, UF\_proprietário,  
 sócios\_proprietário)  
 matriz\_monitorada (ID\_matriz\_monitorada, FK\_propriedade(ID\_propriedade),  
 FK\_matriz(ID\_matriz), data\_início, data\_termino)  
 matriz (ID\_matriz, nome\_matriz, FK\_origem\_matriz(ID\_origem))

origem\_matriz (ID\_origem, origem)  
 lote\_de\_amostra (ID\_lote, FK\_matriz\_monitorada(ID\_matriz\_monitorada),  
 data\_recebimento, data\_avaliação, risco\_lote, FK\_grupo\_do\_lote(ID\_grupo),  
 FK\_intervenções\_regulatórias(ID\_intervenção))  
 grupo\_do\_lote (ID\_grupo\_lote, grupo\_lote)  
 intervenções\_regulatórias (ID\_intervenção, intervenção)  
 amostra (FK\_grupo\_do\_lote(ID\_lote), ID\_amostra, percentual\_conformidade,  
 percentual\_não\_conformidade, total\_analitos\_encontrados, total\_não\_conformes,  
 risco\_original, risco\_toxidade\_analitos, riscoassociações\_perigosas, risco\_final)  
 análise\_risco\_analito (FK\_amostra(ID\_lote, ID\_amostra),  
 FK\_analito\_monitorado\_matriz (ID\_analito, ID\_matriz),  
 concentração\_analito, risco\_concentração, resultado\_percentual,  
 intervalo\_conformidade, intervalo\_não\_conformidade,  
 FK\_tendência\_da\_concentração(ID\_tendência),  
 FK\_prazo\_de\_ajustamento\_da\_concentração(ID\_prazo),  
 FK\_custo\_de\_ajustamento\_da\_concentração(ID\_custo), risco\_agregado)  
 analito (ID\_analito, nome\_analito, FK\_tipo\_analito(ID\_tipo\_analito),  
 FK\_grupo\_analito(ID\_grupo\_analito))  
 tipo\_de\_analito (ID\_tipo\_analito, tipo\_analito)  
 grupo\_de\_analito (ID\_grupo\_analito, grupo\_analito)  
 tendência\_da\_concentração (ID\_tendência, tendência)  
 prazo\_de\_ajustamento\_da\_concentração (ID\_prazo, prazo)  
 custo\_de\_ajustamento\_da\_concentração (ID\_custo, custo)  
 situação\_da\_analise (ID\_situacao\_analise, situação\_analise)  
 analito\_monitorado\_matriz (ID\_analito, ID\_matriz, limite\_referência,  
 FK\_situacao\_da\_analise (ID\_situacao\_analise))

**Figura A.2 - Esquema Conceitual do Banco de Dados Análise de Risco Químico Agrpecuário - ARQA**

## A.4 Script CriarTabelas.sql

O script *CriarTabelas.sql* apresentado a seguir na figura A.3, segue os padrões utilizados pelo SGBD *PostgreSQL 8.4* para criação das tabelas, relacionamentos e atributos que compõem o banco de dados ARQA, ilustradas no Diagrama ER da figura A.1:

```

CREATE TABLE administrativa (
    intervalo_conformidade_risco_agregado_analito CHAR(10),
    intervalo_nao_conformidade_risco_agregado_analito CHAR(10)
);
CREATE TABLE proprietario (
    CPF_CNPJ_proprietario CHAR(20) PRIMARY KEY,
    nome_proprietario VARCHAR(60),
    endereco_proprietario VARCHAR(80),
    CEP_proprietario CHAR(10),
    bairro_proprietario VARCHAR(40),
    cidade_proprietario VARCHAR(50),
    UF_proprietario CHAR(2),
    socios_proprietario VARCHAR(200)
);
CREATE TABLE tipo_propriedade (
    ID_tipo_propriedade INTEGER PRIMARY KEY,
    tipo_propriedade VARCHAR(30)
);
CREATE TABLE propriedade (
    ID_propriedade INTEGER PRIMARY KEY,
    nome_propriedade VARCHAR(80),
    endereco_propriedade VARCHAR(80),
    CEP_propriedade CHAR(10),
    cidade_propriedade VARCHAR(50),
    UF_propriedade CHAR(2),
    regioao_propriedade VARCHAR(50),
    CPF_CNPJ_proprietario CHAR(20) REFERENCES proprietario (CPF_CNPJ_proprietario),
    ID_tipo_propriedade INTEGER REFERENCES tipo_propriedade (ID_tipo_propriedade)
);
CREATE TABLE origem_matriz (
    ID_origem_matriz INTEGER PRIMARY KEY,
    origem_matriz CHAR(10)
);
CREATE TABLE matriz (
    ID_matriz INTEGER PRIMARY KEY,
    nome_matriz VARCHAR(30),
    ID_origem_matriz INTEGER REFERENCES origem_matriz (ID_origem_matriz)
);
CREATE TABLE matriz_monitorada (
    ID_matriz_monitorada INTEGER PRIMARY KEY,
    ID_propriedade INTEGER REFERENCES propriedade (ID_propriedade),
    ID_matriz INTEGER REFERENCES matriz (ID_matriz),
    data_inicio DATE,
    data_termino DATE
);
CREATE TABLE grupo_lote (
    ID_grupo_lote INTEGER PRIMARY KEY,
    grupo_lote VARCHAR(40)
);
CREATE TABLE intervencao_regulatoria (
    ID_intervencao_regulatoria INTEGER PRIMARY KEY,
    intervencao_regulatoria CHAR(20)
);

```

```

CREATE TABLE lote_amostra (
  ID_lote_amostra INTEGER PRIMARY KEY,
  ID_matriz_monitorada INTEGER REFERENCES matriz_monitorada (ID_matriz_monitorada),
  data_recebimento DATE,
  data_avaliacao DATE,
  risco_lote_amostra NUMERIC(3,2),
  media_riscos_finais_amostras NUMERIC (4,3),
  ID_grupo_lote INTEGER REFERENCES grupo_lote (ID_grupo_lote),
  ID_intervencao_regulatoria INTEGER REFERENCES intervencao_regulatoria
(ID_intervencao_regulatoria)
);
CREATE TABLE amostra (
  ID_lote_amostra INTEGER REFERENCES lote_amostra (ID_lote_amostra),
  ID_amostra INTEGER,
  total_analitos_encontrados INTEGER,
  total_analitos_nao_conformes INTEGER,
  limite_minimo_conformidade_risco_original_amostra NUMERIC(6,2),
  limite_maximo_conformidade_risco_original_amostra NUMERIC(6,2),
  percentual_conformidade_risco_original_amostra NUMERIC(5,2),
  limite_minimo_nao_conformidade_risco_original_amostra NUMERIC(6,2),
  limite_maximo_nao_conformidade_risco_original_amostra NUMERIC(6,2),
  percentual_nao_conformidade_risco_original_amostra NUMERIC(5,2),
  risco_original_amostra NUMERIC(3,2),
  risco_toxidade_analitos NUMERIC(3,2),
  risco_associacoes_perigosas_analitos NUMERIC(3,2),
  risco_final_amostra NUMERIC(3,2),
  PRIMARY KEY (ID_lote_amostra, ID_amostra)
);
CREATE TABLE grupo_analito (
  ID_grupo_analito INTEGER PRIMARY KEY,
  grupo_analito VARCHAR(50)
);
CREATE TABLE tipo_analito (
  ID_tipo_analito INTEGER PRIMARY key,
  tipo_analito CHAR(12)
);
CREATE TABLE analito (
  ID_analito INTEGER PRIMARY KEY,
  nome_analito VARCHAR(60),
  ID_tipo_analito INTEGER REFERENCES tipo_analito (ID_tipo_analito),
  ID_grupo_analito INTEGER REFERENCES grupo_analito (ID_grupo_analito)
);
CREATE TABLE tendencia_concentracao (
  ID_tendencia INTEGER PRIMARY KEY,
  tendencia VARCHAR(50)
);
CREATE TABLE prazo_ajustamento (
  ID_prazo_ajustamento INTEGER PRIMARY KEY,
  prazo_ajustamento VARCHAR(20)
);
CREATE TABLE custo_ajustamento (
  ID_custo_ajustamento INTEGER PRIMARY KEY,
  custo_ajustamento VARCHAR(20)
);

```

```

CREATE TABLE situacao_analise (
  ID_situacao_analise INTEGER PRIMARY KEY,
  situacao_analise VARCHAR(60)
);
CREATE TABLE analito_monitorado_matriz (
  ID_analito INTEGER REFERENCES analito (ID_analito),
  ID_matriz INTEGER REFERENCES matriz (ID_matriz),
  limite_referencia NUMERIC(8,2),
  ID_situacao_analise INTEGER REFERENCES situacao_analise (ID_situacao_analise),
  PRIMARY KEY (ID_Analito, ID_matriz)
);
CREATE TABLE analise_risco_analito (
  ID_lote_amostra INTEGER,
  ID_amostra INTEGER,
  ID_analito INTEGER,
  ID_matriz INTEGER,
  limite_referencia NUMERIC(8,2),
  ID_situacao_analise INTEGER REFERENCES situacao_analise (ID_situacao_analise),
  concentracao_analito NUMERIC(8,2),
  risco_concentracao_analito NUMERIC(3,2),
  resultado_percentual NUMERIC(5,2),
  intervalo_conformidade_risco_agregado_analito CHAR(10),
  intervalo_nao_conformidade_risco_agregado_analito CHAR(10),
  ID_tendencia INTEGER REFERENCES tendencia_concentracao (ID_tendencia),
  ID_prazo_ajustamento INTEGER REFERENCES prazo_ajustamento (ID_prazo_ajustamento),
  ID_custo_ajustamento INTEGER REFERENCES custo_ajustamento (ID_custo_ajustamento),
  risco_agregado_analito NUMERIC(4,2),
  FOREIGN KEY (ID_lote_amostra, ID_amostra) REFERENCES amostra (ID_lote_amostra,
  ID_amostra),
  FOREIGN KEY (ID_analito, ID_matriz) REFERENCES analito_monitorado_matriz (ID_analito,
  ID_matriz),
  PRIMARY KEY (ID_lote_amostra, ID_amostra, ID_analito, ID_matriz)
);

```

**Figura A.3 - Script CriarTabelas.sql para Geração de Tabelas, Relacionamentos e Atributos - Banco de Dados Análise de Risco Químico Agropecuário - ARQA**

## A.5 Análise das Atualizações Estruturais

Esta seção apresenta uma sequência de informações de caráter geral, complementares ao diagrama ER, que auxiliam a manutenção da integridade dos dados manipulados pelos programas da Plataforma Base Flexível que interagem com o banco de dados ARQA. As principais são:

- Um lote de amostras de matrizes de produto agropecuário monitorado pelo PNCRC / MAPA pode conter uma única amostra ou milhões destas, de acordo com o âmbito desejado para a análise de risco;
- Um lote de amostras deve pertencer a um único produto agropecuário;

- Colhem-se amostras provenientes de diferentes matrizes de diferentes propriedades e proprietários;
- Uma amostra sempre pertence a um determinado lote, não sendo possível a existência de amostras sem um lote específico;
- As amostras são identificadas por números sequenciais crescentes de acordo com o lote ao qual pertencem;
- Os analitos de uma amostra são particulares às matrizes do produto agropecuário monitorado pelo PNCRC / MAPA e pré-determinados pelo Codex Alimentarius;
- Todas as datas devem seguir o padrão *aaaa-mm-dd*;
- A parte decimal dos valores numéricos contínuos deve conter 2 (dois) dígitos de precisão, separados por ponto decimal;
- A tabela *analito\_monitorado\_matriz* adota o termo genérico Limite de Referência particular a cada analito monitorado para a matriz do produto agropecuário, de acordo com o Limite Máximo de Resíduo (LMR) pré-estabelecido, caso o analito seja considerado um resíduo, ou com o Teor Máximo de Contaminante (TMC), caso seja considerado um contaminante, ou finalmente de acordo com o Limite Mínimo de Desempenho Requerido (LMDR) caso os índices anteriores não estejam formalmente definidos para o analito;
- A tabela *administrativa* pertence ao projeto físico de banco de dados:
  - O campo *intervalo\_conformidade\_amostra* apresenta os limites mínimo (0) e máximo (7) de conformidade da amostra para valores armazenados no campo *risco\_concentração* da tabela *analise\_risco\_analito* compreendidos no intervalo [0; 4];
  - O campo *intervalo\_nao\_conformidade\_amostra* apresenta os limites mínimo (4) e máximo (15) de não conformidade da amostra para valores armazenados no campo *risco\_concentração* compreendidos no intervalo [5; 6].
- Os termos inseridos nas tabelas do banco de dados, quando originados a partir do Fluxo de Trabalho Requisitos do método UPON são rigorosamente os mesmos termos armazenados na ontologia APA;