

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL**

RICARDO MARTINS TAQUES

**UMA AVALIAÇÃO SOBRE DIFERENTES PADRÕES DE MODELAGENS PARA
ÍNDICES E INDICADORES DE POLUIÇÃO DO AR EM ONTOLOGIAS**

DISSERTAÇÃO DE MESTRADO

CURITIBA

2018

RICARDO MARTINS TAQUES

**UMA AVALIAÇÃO SOBRE DIFERENTES PADRÕES DE MODELAGENS PARA
ÍNDICES E INDICADORES DE POLUIÇÃO DO AR EM ONTOLOGIAS**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Mestre em Ciências” – Área de Concentração: Informática Industrial.

Orientador: Prof. Dr. Cesar Augusto Tacla.

CURITIBA

2018

Dados Internacionais de Catalogação na Publicação

T175a Taques, Ricardo Martins

Uma avaliação sobre diferentes padrões de modelagens para índices e indicadores de poluição do ar em ontologias / Ricardo Martins Taques.-- 2018.

114 f.: il.

Disponível via World Wide Web.

Texto em inglês com resumo em português.

Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. Área de Concentração: Informática Industrial, Curitiba, 2018.

Bibliografia: f. 96-99.

1. Engenharia elétrica - Dissertações. 2. Ontologias (Recuperação da informação) - Avaliação. 3. Representação do conhecimento (Teoria da informação). 4. Sistemas de reconhecimento de padrões. 5. Modelos matemáticos. 6. Lógicas de descrição. 7. Web Semântica. 8. Ar - Poluição - Índices. 9. Indicadores ambientais. 10. Métodos de simulação. I. Tacla, Cesar Augusto, orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. III. Título.

CDD: Ed. 23 – 621.3

Biblioteca Central do Câmpus Curitiba – UTFPR
Bibliotecária: Luiza Aquemi Matsumoto CRB-9/794



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Diretoria de Pesquisa e Pós-Graduação

TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 815

A Dissertação de Mestrado intitulada “**Uma Avaliação Comparativa Entre Diferentes Padrões de Modelagens de índices e Indicadores de Poluição do Ar em Ontologias**” defendida em sessão pública pelo(a) candidato(a) **Ricardo Martins Taques**, no dia 14 de dezembro de 2018, foi julgada para a obtenção do título de Mestre em Ciências, área de concentração Engenharia de Computação, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial.

BANCA EXAMINADORA:

Prof(a). Dr(a) Cesar Augusto Tacla - Presidente (UTFPR)

Prof(a). Dr(a). Josiane Michalak Dall Agnoll - (UNICENTRO)

Prof(a). Dr(a). Milton Borsato - (UTFPR)

Prof(a). Dr(a). Rita Cristina Galarraga Berardi - (UTFPR)

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 14 de dezembro de 2018.

AGRADECIMENTOS

À Universidade Tecnológica Federal do Paraná por conduzir com muito afinco e dedicação, através de seu corpo docente e demais funcionários, os cursos de pós-graduação permitindo assim que profissionais nas áreas de tecnologia renovem conhecimento, se especializem em linhas de pesquisa de ponta e possam ter novas perspectivas no campo profissional.

Ao paciente e dinâmico professor Cesar A. Tacla pelas horas de orientação, de repasse de conhecimento e paciência com o autor dessa dissertação.

À professora Rita C. Berardi pelo espírito criativo e motivador que nos ajuda a seguir em frente e, aos colegas de mestrado Mônica Barazzetti e de graduação Mateus Belizário pelo auxílio dado, ambos sempre prestativos e com muita simpatia e dedicação.

À minha querida esposa Lívia que me apoiou por todo esse período de estudos com muito carinho e compreensão.

RESUMO

Diversos projetos desenvolveram ontologias com o intuito de representar indicadores de poluição ambiental. Entretanto, como visto nos trabalhos correlatos, alguns dos aspectos referentes às formas de representação dos conceitos e relações envolvidos na relação índice-indicador não foram amplamente explorados. Considerando os problemas advindos destas formas de representação, o presente trabalho tem por objetivo propor formas alternativas de modelagem para a relação índice-indicador, pelo uso de padrões apresentados em um formato customizado, com base nas características envolvidas em cenários extraídos das especificações para geração de índices sobre poluição do ar mantidos por agências ambientais. Adicionalmente, para avaliar as duas modelagens propostas nesta dissertação, foram adotados alguns dos critérios usualmente utilizados para análise da qualidade de ontologias: completude, concisão, expansibilidade e qualidade dos resultados obtidos. Conjuntamente, convida aos especialistas no assunto a fazer uma discussão acerca de padrões para definição de construtos relacionados aos indicadores e índices que visam atender uma utilização sob uma perspectiva particular ditada pelas necessidades do usuário final. Por fim, os padrões são validados através de estudos de caso que orientam uma aplicação para a descrição de índices e indicadores em duas ontologias. Destas avaliações, uma discussão é proposta pela análise dos resultados alcançados que antecede as conclusões finais sobre o trabalho.

Palavras-chave: Ontologia. Representação do conhecimento. Modelagem. Padrões. Índices-indicadores. Meio ambiente. Poluição do ar. Critérios. Avaliação.

ABSTRACT

Several projects have developed ontologies to represent indicators of environmental pollution. However, as seen in the correlated workers, some of the aspects regarding the representation of the concepts and relations involved in the index-indicator relationship were not widely explored. Considering the problems arising from these forms of representation, the present work aims to propose alternative forms for modeling the relation between index and indicators, using content patterns presented in a customized format, based on the features involved in scenarios extracted from the air pollution indexes' specifications, provided by some of the most important agencies for environment control and monitoring. In addition, to evaluate the two models proposed in this dissertation, a set of criteria usually utilized to analyze the quality of ontologies are being used, i.e.: completeness, conciseness, expandability and the quality of the outputs. Along with the main objective, this work invites the subject matter experts to discuss about patterns related to indicators and indexes that aim to meet a particular perspective of the end-user needs. Finally, these proposed patterns are validated through case studies that guide an application to describe index and indicators in two ontologies. From these evaluations, a discussion is proposed by analyzing the results achieved prior to outline the conclusions about the work.

Keywords: Ontology. Knowledge representation. Modelling. Design patterns. Index-indicators. Environment. Air pollution. Criteria. Evaluation.

LISTA DE FIGURAS

Figura 1 – Componentes da Ontologia GCI_Foundation	18
Figura 2 – Reusabilidade de componentes no indicador Primary_Student_teacher_ratio.....	19
Figura 3 – Hierarquia de classes da ontologia AIR_POLLUTION_Onto	22
Figura 4 – Pilha de camadas da Web Semântica.	26
Figura 5 – Padrão: Partição de valores.....	30
Figura 6 – Categorias e Faixas de Valores do Índice IMECA	36
Figura 7 – Categorias e Fórmulas para o Índice IMECA referente ao Poluente PM10.....	36
Figura 8 – Categorias e Faixas de Valores do Índice AQHI	37
Figura 9 – Fórmula de Cálculo do Índice AQHI	37
Figura 10 – Etapas da metodologia.....	39
Figura 11 – Diagrama de atividades do sistema fictício de monitoramento da poluição do ar.	46
Figura 12 – Hierarquia das classes de poluentes do ar.	51
Figura 13 – Indicador PSTR com classes das ontologias Units of Measure (OM) e Time (OT)	52
Figura 14 – Classe DurationDescription em sintaxe Manchester.....	53
Figura 15 – Diagrama UML da Classe DurationDescription importada de Time.owl.	53
Figura 16 – Classe Station em sintaxe Manchester	54
Figura 17 – Classe Agent importada de Prov-o.owl.....	54
Figura 18 – Classe DerivationForms com as suas subclasses.	59
Figura 19 – Declaração em sintaxe Manchester utilizando enumeração de classes.....	59
Figura 20 – Subclasses da classe BQ_UserIndicator.	72
Figura 21 – Subclasses disjuntas que completam a classe IndexCategorization.....	77
Figura 22 – Etapas para desenvolvimento do Estudo de Caso.....	78

LISTA DE QUADROS

Quadro 1 – Classes Indicator e IndicatorSet do modelo OSIS – Generic.....	20
Quadro 2 – Equivalência semântica entre construtos básicos de DL e FOL	25
Quadro 3 – Elementos da lógica de descrição (DL) empregados em OWL DL	26
Quadro 4 – Exemplo comparativo das representações para a definição de uma classe/conceito	27
Quadro 5 – Famílias de padrões de modelagem (ODP).....	28
Quadro 6 – Comparativo das abordagens e critérios de avaliação de ontologias.....	32
Quadro 7 – Relação dos índices pesquisados	34
Quadro 8 – Características dos poluentes do Ar.....	38
Quadro 9 – Modelo para enumeração dos termos a serem modelados como classes.....	40
Quadro 10 – Modelo para enumeração dos termos a serem modelados como propriedades...	41
Quadro 11 – Quadro para apresentação da avaliação de uma ontologia.....	43
Quadro 12 – Enumeração dos termos a serem modelados como classes.....	48
Quadro 13 – Enumeração dos termos a serem modelados como propriedades.	49
Quadro 14 – Fontes de organizações e mecanismos para busca de ontologias candidatas a reuso	50
Quadro 15 – Enumeração dos termos a serem modelados como classes.....	54
Quadro 16 – Enumeração dos termos a serem modelados como propriedades.	56
Quadro 17 – Enumeração dos termos a serem modelados como classes.....	66
Quadro 18 – Enumeração dos termos a serem modelados como propriedades	67
Quadro 19 – Enumeração dos termos a serem modelados como classes.....	69
Quadro 20 – Enumeração dos termos a serem modelados como propriedades.	70
Quadro 21 – Classes da ontologia AQ	79
Quadro 22 – Instâncias da ontologia AQ.	80
Quadro 23 – Avaliação dos Padrões da Modelagem “A” pelas Questões de Competência	82
Quadro 24 – Avaliação da ontologia AQ pela modelagem “A”	83
Quadro 25 – Classes da ontologia BQ	84
Quadro 26 – Instâncias da ontologia BQ	85
Quadro 27 – Avaliação dos Padrões da Modelagem “B” pelas Questões de Competência	88
Quadro 28 – Avaliação da ontologia BQ pela modelagem “B”	89
Quadro 29 – Comparativo das avaliações das modelagens.....	90

LISTA DE ABREVIATURAS E SIGLAS

AI	Artificial Intelligence
API	Application Programming Interface
AQI	Air Quality Index (E.U.A.)
AQHI	Air Quality and Health Index (Canadá)
CO	Monóxido de Carbono / Carbon Monoxide
CP	Content Patterns
DAML	Darpa Agent Markup Language
DAQI	Daily Air Quality Index (Reino Unido)
DARPA	Defense Advanced Research Projects Agency
DEFRA	Department for Environment, Food and Rural Affairs (Reino Unido)
DF	Distrito Federal (Cidade do México)
DL	Lógica de Descrição
DP	Design Patterns
EC	Environment Canada
EIL	Enterprise Integration Ontology
EPA	Environmental Protection Agency
FOL	Lógica de Primeira Ordem
GCI	Global City Indicators e Global Cities Institute
GitHub	Repositório distribuído para controle de versão de softwares
IAP	Instituto Ambiental do Paraná
IMECA	Índice Metropolitano de la Calidad del Aire
IQA	Índice de Qualidade do Ar
ISO	International Standards Organization
KP	Knowledge Provenance ontology
MAC	Maximum Admissible Concentration
μm	micrômetros (unidade de medida)
$\mu\text{g}/\text{m}^3$	Microgramas por metro cúbico (unidade de medida para densidade de massa)
MPTS	Material Particulado Total em Suspensão (poluentes naturais)
N/A	Não se aplica
NADF	Norma Ambiental para o Distrito Federal (México)
NO_2	Dióxido de Nitrogênio / Nitrogen Dioxide
ODP	Ontology Design Patterns

OSIS	Ontology for Sustainability Indicator Sets
O ₃	Ozônio / Ozone
OM	Organization and Measurement ontology
OT	Time ontology
OML	Ontology Markup Language
OIL	Ontology Interchange Language
OWL	Web Ontology Language
PI	Partículas Inaláveis
PM ₁₀	Partículas em Suspensão menores do que 10 micrômetros
PM _{2,5}	Partículas em Suspensão menores do que 2.5 micrômetros
PPB	Partes por bilhão
PPM	Partes por milhão
PO	Pollution Ontology
PR	Provenance Ontology
PROV	Provenance ontology
PTS	Partículas Totais em Suspensão
RDF	Resource Description Framework
ROMEO	Metodologia Orientada a Requisitos para Avaliação de Ontologias
Rule ML	Rule Markup Language
SE	Sensor Ontology
SGBD	Sistemas Gerenciadores de Bancos de Dados Relacionais
SO ₂	Dióxido de Enxofre / Sulfur Dioxide
SPARQL	SPARQL Protocol and RDF Query Language
SSN	Semantic Sensor Network
SWRL	Semantic Web Rule Language
UFSC	Universidade Federal de Santa Catarina
UK	United Kingdom
UML	Unified Modelling Language
USA	United States of America
XML	Extensible Markup Language
XSD	XML Schema Definition
W3C	World Wide Web Consortium

SUMÁRIO

1 INTRODUÇÃO.....	13
2 REVISÃO BIBLIOGRÁFICA	17
2.1 ONTOLOGIAS DE INDICADORES.....	17
2.2 ONTOLOGIAS DE POLUENTES COM INDICADORES	21
3 FUNDAMENTAÇÃO TEÓRICA	23
3.1 ONTOLOGIAS.....	23
3.2 LÓGICAS DE DESCRIÇÃO (DLS)	24
3.3 OWL – LINGUAGEM WEB PARA ONTOLOGIAS	25
3.4 PADRÕES DE MODELAGEM	27
3.5 PADRÕES ELEMENTARES	29
3.6 CRITÉRIOS PARA AVALIAÇÃO DE MODELAGENS	31
3.7 ÍNDICES SOBRE POLUIÇÃO DO AR.....	33
3.7.1 Composição do Índice IMECA – Índice Metropolitano de la Calidad del Aire.....	35
3.7.2 Composição do Índice AQHI – Air Quality Health Index.....	36
3.8 CARACTERÍSTICAS DOS POLUENTES DO AR.....	37
4 METODOLOGIA	39
4.1 IDENTIFICAÇÃO DO CENÁRIO COMO REQUERIMENTO PARA MODELAGEM	39
4.2 DETERMINAR QUESTÕES DE COMPETÊNCIA PARA AUXILIAR NA ESPECIFICAÇÃO DOS PADRÕES	40
4.3 ENUMERAR OS COMPONENTES DA MODELAGEM	40
4.4 CONSIDERAR O REUSO DE ELEMENTOS DE ONTOLOGIAS <i>TOP-LEVEL</i> NA MODELAGEM	41
4.5 ESPECIFICAÇÃO DOS PADRÕES DE MODELAGEM.....	42
4.6 AVALIAÇÃO DA MODELAGEM	42
4.6.1 Recursos Utilizados	43
5 PADRÕES PROPOSTOS – MODELAGEM A.....	45
5.1 IDENTIFICAÇÃO DO CENÁRIO COMO REQUERIMENTO PARA MODELAGEM	45
5.1.1 Características do Cenário para a Modelagem A	46
5.2 DETERMINAR AS QUESTÕES DE COMPETÊNCIA PARA AVALIAÇÃO DA MODELAGEM	47
5.3 ENUMERAR OS COMPONENTES DA MODELAGEM	48

5.4 CONSIDERAR O REUSO DE ELEMENTOS DE ONTOLOGIAS TOP-LEVEL NA MODELAGEM	50
5.4.1 Identificação das Classes e Relações Associadas aos Poluentes	51
5.4.2 Identificação das Classes e Relações Associadas aos Indicadores	51
5.4.3 Identificação das Classes e Relações Associadas aos Sensores	53
5.4.4 Identificação das Classes e Relações Associadas à Proveniência dos Dados	54
5.5 ESPECIFICAÇÃO DOS PADRÕES DE MODELAGEM.....	56
6 PADRÕES PROPOSTOS – MODELAGEM B.....	64
6.1 IDENTIFICAÇÃO DO CENÁRIO COMO REQUERIMENTO PARA MODELAGEM	64
6.1.1 Características do Cenário para a Modelagem B	64
6.2 DETERMINAR AS QUESTÕES DE COMPETÊNCIA PARA AVALIAÇÃO DA MODELAGEM	65
6.3 ENUMERAR OS COMPONENTES DA MODELAGEM	66
6.4 CONSIDERAR O REUSO DE ELEMENTOS DE ONTOLOGIAS <i>TOP-LEVEL</i> NA MODELAGEM	68
6.4.1 Identificação das Classes e Relações Associadas aos Poluentes	68
6.4.2 Identificação das Classes e Relações Associadas aos Indicadores	68
6.4.3 Identificação das Classes e Relações Associadas aos Sensores	68
6.4.4 Identificação das Classes e Relações Associadas à Proveniência dos Dados	69
6.5 ESPECIFICAÇÃO DOS PADRÕES DE MODELAGEM.....	71
7 ESTUDOS DE CASO.....	78
7.1 DESENVOLVIMENTO DA ONTOLOGIA AIRQUALITYONTO (AQ) PELA MODELAGEM “A”	78
7.1.1 Introdução ao Cenário de Teste.....	78
7.1.2 Criação das Classes e Hierarquia	79
7.1.3 Criação de Indivíduos Via Instanciação.....	80
7.1.4 Avaliação da Ontologia (AQ).....	82
7.2 DESENVOLVIMENTO DA ONTOLOGIA AIRQUALITYONTO (BQ) PELA MODELAGEM “B”	84
7.2.1 Introdução ao Cenário de Teste.....	84
7.2.2 Criação das Classes e Hierarquia	84
7.2.3 Criação de Indivíduos Via Instanciação.....	85
7.2.4 Avaliação da Ontologia (BQ).....	87
8 RESULTADOS E DISCUSSÃO.....	90

9 CONCLUSÕES	92
9.1 LIMITAÇÕES	93
9.2 TRABALHOS FUTUROS	94
REFERÊNCIAS	96
APÊNDICE A – CONSULTAS EM LINGUAGEM SPARQL	100
APÊNDICE B – ESTATÍSTICAS DAS ONTOLOGIAS	109
APÊNDICE C – CONTEÚDO DA ONTOLOGIA SWRL	110

1 INTRODUÇÃO

No contexto atual, onde se vê uma crescente preocupação com as condições do meio ambiente, surge uma oportunidade para exploração de tecnologia semântica. Com o intuito de se evitar catástrofes na natureza que afetem não só ela própria, mas também as populações que nela habitam, torna-se crucial a monitoração da poluição ambiental.

Existem atualmente várias entidades, majoritariamente governamentais, que realizam o monitoramento sobre os níveis de poluição nos rios, no solo e também no ar. Estes órgãos atuam utilizando estações de coleta com equipamentos especialmente desenvolvidos para obter dados sobre poluentes. Estas coletas ocorrem periodicamente e depois de tratadas, produzem medições que podem ser comparadas com patamares e escalas de valores, cientificamente estabelecidos pelos órgãos ambientais reguladores. Assim, é possível analisar a presença dos agentes poluentes e promover mudanças na rotina das áreas com situação de maior risco ou impacto.

Entretanto, este monitoramento pode ser enriquecido com um sistema que permita viabilizar um compartilhamento em tempo real da situação de cada local em relação às condições ambientais dos poluentes. Deste modo, seria possível antecipar problemas no meio ambiente ou, reagir rapidamente contra alterações inesperadas na emissão de algum poluente. Ainda mais, o sistema poderia ficar disponível para consulta de usuários não especialistas, mas que também precisam estar informados. E poderia permitir que usuários especialistas façam uso de suas funcionalidades, ampliando deste modo seus conhecimentos ao explorarem novas formas de se trabalhar com as informações geradas por ele.

Um sistema com esta missão deve ser capaz de adquirir e tratar dados heterogêneos produzidos e publicados por diferentes organizações, e disponibilizá-los de forma a não exigir dos agentes computacionais que os acessem, conhecimento de cada modelo específico utilizado para publicação dos dados originais. Desta forma, este sistema provê uma camada intermediária de acesso aos dados que esconde suas diferenças de modelagem e de representação. Nesta dissertação, parte-se do princípio que tecnologias relacionadas à Web Semântica são adequadas para resolver o problema apresentado, sendo que ontologias são um dos pilares destas tecnologias. É importante esclarecer que esta dissertação não aborda o problema de eficiência de armazenamento e consulta de grandes volumes de medições (dados) como também, não explora como os dados devem ser recebidos e tratados a partir de diferentes fontes.

Diversas ontologias que permitem a representação de conceitos e propriedades envolvidas no contexto do monitoramento do meio ambiente têm surgido, tais como: Semantic Sensor Network ¹ (SSN); Environment, Pollution e GCI-Foundation ², AIR_POLLUTION_Onto³. Porém, o nível de abstração destas ontologias é variável. A ontologia SSN representa informações em um nível menos abstrato, abrangendo os conceitos de sensores, observadores e medições. As demais possuem nível de abstração mais alto, abrangendo termos, como os tipos de poluentes, níveis de concentração e os ambientes em que se encontram. As ontologias do projeto GCI seguem as normas e padrões da ISO 37120⁴ (2018) que por sua vez tratam de índices e indicadores de serviços urbanos e qualidade de vida. A ontologia AIR_POLLUTION_Onto desenvolvida por Oprea (2009), visa o monitoramento e controle da poluição do ar pela comparação dos níveis de concentração obtidos para os poluentes com seus valores máximos de concentração admitidos.

Além do alto nível de abstração das ontologias citadas, observou-se também que seus conteúdos não são adequados para representar informações como a derivação do valor de um indicador a partir de um conjunto de medições ou, como calcular um índice a partir de seus indicadores. As representações destas derivações quando presentes, são feitas normalmente na forma textual em comentários. Pelos estudos sobre índices de poluição do ar (vide Seção 3) praticados pelas autoridades ambientais, percebeu-se que indicadores são derivados por meio de operações de agregação sobre as medições de poluentes, salvaguardando informações temporais (data/hora de referência do indicador) e de proveniência (qual estação/sensor fez a medição, quem fez sua publicação).

É importante ressaltar que conceitualmente, um indicador é um dado qualitativo, quantitativo ou uma medição descritiva e que um índice pode ser considerado equivalente a um indicador. Contudo, de acordo com Siche et al. (2007), um índice é um valor agregado final de todo um procedimento de cálculo onde se utilizam inclusive, os indicadores. Além disso, um índice pode se tornar um instrumento para tomada de decisão, emissão de alertas ou para a previsão de um evento. Por conseguinte, pode-se concluir que um indicador antecede um índice porque este último é formado por um ou mais indicadores.

Pôde-se ver que esta interpretação descrita sobre a relação entre índices e indicadores é percebida nas especificações dos índices pesquisados. Relação esta que não está sendo tratada pelas ontologias relacionadas à poluição ambiental.

¹Desenvolvida pela World Wide Web Consortium (W3C).

²Desenvolvida no projeto GCI – Global City Indicators.

³Desenvolvida por Mihaela Oprea – University Petroleum-Gas of Ploiesti, Romênia.

⁴ISO37120 – International Organization for Standardization – Indicators for city services.

Então, foi a partir desta constatação que se propôs estender as ontologias utilizadas no domínio da poluição do ar para que índices e indicadores sejam descritos de forma mais refinada. O intuito, portanto, é facilitar a construção de ontologias que tratam o domínio sobre o monitoramento da poluição do ar, por meio de padrões de modelagem para o uso de aplicações que necessitam criar ou utilizar índices distintos daqueles pré-definidos na ISO 37120 (2018) ou dos indicadores propostos por Oprea (2009).

Esta extensão se baseia em construtos que permitem representar: um índice a partir de fórmulas que tratam indicadores que agregam medições por meio de operadores, como: média, mínimo e máximo; e, categorizações de índices em classes definidas por partições de valores (ex. ruim, regular, bom), o que é uma prática comum na definição de índices ambientais.

A maior dificuldade encontrada, está nas formas de representação computacionais aplicáveis em ontologias para a relação índice-indicador. Esta relação apresenta as seguintes características:

- meronímia (parte-todo);
- relações que apresentam propriedades;
- funções de agregação, de medições ou fórmulas matemáticas que permitem calcular valores de indicadores e índices a partir das medições;
- dependência de categorizações identificadas por intervalos de valores.

Destes problemas, surge o objetivo do presente trabalho: desenvolver e avaliar formas alternativas de modelagem por meio de padrões para a relação índice-indicador com base nas características envolvidas em um cenário proposto.

Para avaliar os padrões propostos, foram escolhidos alguns critérios dentre aqueles que são utilizados em estudos sobre avaliação de ontologias. Nesta escolha foi considerado que ontologias sejam: flexíveis; que seus axiomas sejam abrangentes e respeitem a linguagem formal adotada sem serem redundantes; que façam uso de padrões de modelagem; e que tenham qualidade nos resultados gerados para uso de aplicações. Estas características visam atender, respectivamente, os critérios de: expansibilidade, completude, concisão e a qualidade nas saídas geradas.

Para desenvolvimento deste trabalho, buscou-se fundamentação teórica a respeito das linguagens de descrição e seus recursos para uso nas modelagens. Além disso, índices ambientais de poluição do ar foram pesquisados, ontologias sobre o tema foram analisadas, critérios de avaliação foram escolhidos dentre os utilizados em trabalhos que tratam o assunto

e, uma metodologia foi proposta para o desenvolvimento e avaliação das modelagens aqui apresentadas.

O restante deste trabalho está organizado da seguinte forma: a seção 2 contém a revisão bibliográfica, na seção 3 são apresentados os conceitos e definições sobre os quais se fundamenta a dissertação. Na seção 4, é apresentada a metodologia empregada para desenvolvimento dos padrões que são utilizados em duas alternativas de modelagem apresentadas, respectivamente, nas seções 5 e 6. Em seguida, a seção 7 contém os estudos de caso que permitiram aplicar as duas alternativas de modelagem previamente desenvolvidas. Na seção 8 são apresentados e discutidos os resultados alcançados. Por fim, a seção 9 traz as conclusões e perspectivas para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

2.1 ONTOLOGIAS DE INDICADORES

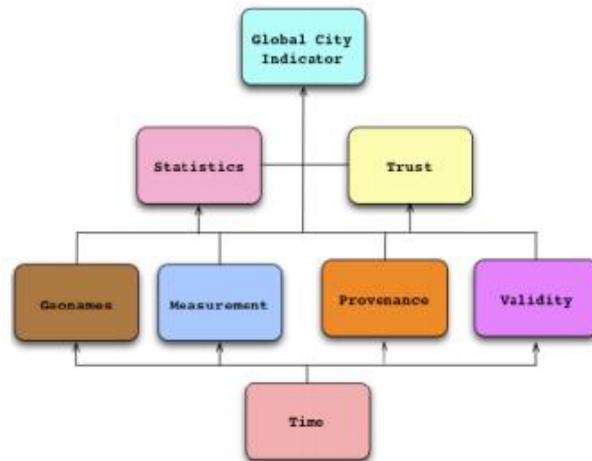
Um indicador, segundo a análise feita por Siche et al. (2007), pode ser um dado individual ou um agrupamento de informações e um índice é o valor agregado final de cálculos onde se utilizam inclusive, os indicadores. Sendo assim, os dois conceitos e suas relações precisam ser definidos nos padrões para construção de uma ontologia sobre a poluição do ar.

Dentre os trabalhos mais importantes em modelagem semântica aplicada ao domínio da poluição ambiental, destacam-se os trabalhos no âmbito do projeto GCI – Global Cities Institute, da Universidade de Toronto (DAHLEH; FOX, 2016), de Fox (2015) sobre como conceitualizar um indicador educacional e, da ontologia AIR_POLLUTION_Onto desenvolvida em Oprea (2009).

No projeto dos indicadores globais para cidades inteligentes conhecido pela sigla GCI, é apresentado um método para criação de ontologias para tratarem de indicadores urbanos. Dentre tantas, algumas ontologias se referem ao domínio do meio-ambiente, como já foram mencionadas na introdução deste trabalho. Deste método pode-se destacar o reaproveitamento de conceitos e relações de ontologias já produzidas sobre o tema e também, da aplicação de questões de competência para definição das classes das ontologias da GCI e para posterior avaliação da qualidade destas.

Da análise feita por Dahleh e Fox (2016) sobre as ontologias existentes e as questões de competência que deviam ser respondidas, surgiu a ontologia base do projeto, chamada de *GCI_Foundation.owl* cujos componentes estão demonstrados na Figura 1 ilustrada a seguir.

Figura 1 – Componentes da Ontologia GCI_Foundation

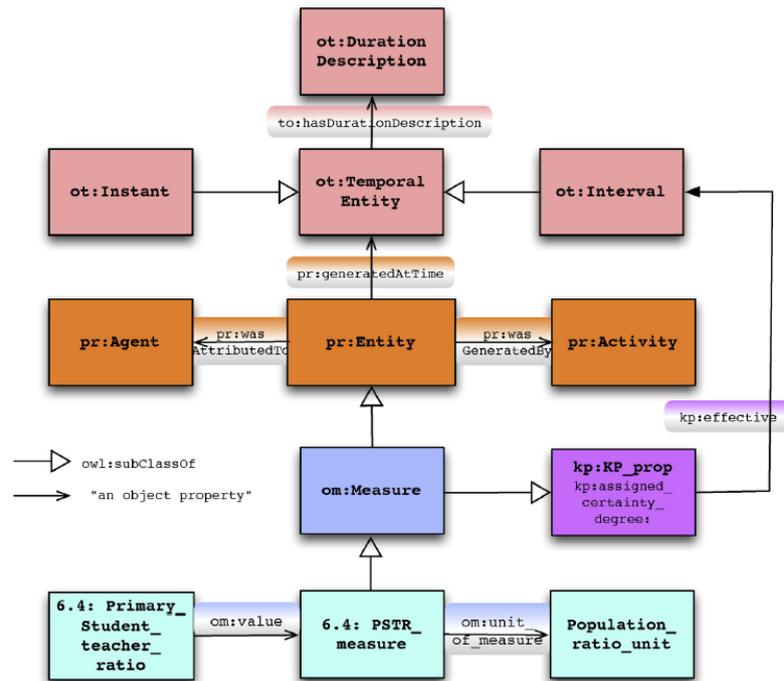


Fonte: Dahleh e Fox (2016)

Segundo os autores, esta ontologia serviu de fundação para o desenvolvimento das demais ontologias de indicadores que fazem parte do projeto. As ontologias que a sucederam tiveram que adicionar suas próprias informações sobre esta ontologia de base conforme os temas específicos que fossem especificar.

Em relação à modelagem de indicadores, Fox (2015) aborda os mais importantes aspectos e características que devem fazer parte da criação de indicadores urbanos, ao especificar a criação de um indicador educacional, ilustrado na Figura 2. Nesta, podem ser vistas as classes e propriedades que o compõem e de quais ontologias de alto nível foram importados os conceitos e relações presentes nesta conceitualização. Por exemplo, a ontologia *time.owl* está presente e identificada com o prefixo **ot**, assim como, a ontologia *prov-o.owl* com o prefixo **pr**. Outras duas ontologias que emprestaram suas classes nesta definição são: *om-1.6.owl* (*Units of Measure*) e *kp.owl* (*Knowledge Provenance*) identificadas respectivamente pelos prefixos: **om** e **kp**.

Figura 2 – Reusabilidade de componentes no indicador Primary_Student_teacher_ratio



Fonte: (Fox, 2015)

Em Ghahremanloo, Thom e Magee (2012), os autores apresentam uma abordagem que utiliza as ferramentas METHONTOLOGY (FERNÁNDEZ-LÓPEZ et al., 1997) e ROMEO (YU et al., 2009) para a modelagem e avaliação de uma ontologia que faz uso de indicadores em um projeto denominado OSIS – *Ontology for Sustainability Indicators Sets*. Embora esta abordagem não esteja relacionada ao domínio da poluição ambiental, vê-se a presença de uma classe superior que representa um indicador que está definida com propriedades de objetos e propriedades de dados. E, algumas destas propriedades são equivalentes às utilizadas em Fox (2015) porque abordam os aspectos temporais, das medições e da proveniência dos dados.

Para melhor entendimento e garantia de se referenciar corretamente o conteúdo empregado pelos autores, se apresenta no Quadro 1, como as classes indicador e conjunto de indicadores foram definidas no modelo genérico adotado na ontologia OSIS.

Quadro 1 – Classes Indicator e IndicatorSet do modelo OSIS – Generic

Domain (Subject)	Property (Predicate)	Rang (Object)
Superclass: Indicator+	dc:title dc:type dc:description dc:periodOfTime dc:publisher osis:instance-of osis:hasCategory osis:hasReference osis:hasUnitofMeasurement osis:hasIssue osis:belongsToIndicatorSystem	String String Superclass: Description String Superclass: IndicatorSet Superclass: Indicator Superclass: Category Superclass: Reference String SuperClass: Issue SuperClass: IndicatorSet
SuperClass: IndicatorSet+	dc:title dc:text osis:hasIndicator	String String Superclass: Indicator

Fonte: Ghahremanloo, Thom e Magee (2012).

A ontologia OSIS foi desenvolvida com o propósito de representar conhecimentos sobre indicadores de sustentabilidade no contexto de uma aplicação específica. Ela pode ser reutilizada para interpretação semântica, replicação e consultas sobre os indicadores, principalmente, porque foi criada em linguagem de representação RDF/OWL que garante sua reusabilidade. Além disto, este formato capacita agentes computacionais a interpretarem as conceitualizações e axiomas definidos no modelo, como também permite que aplicações possam consultar a sua base de conhecimento através da linguagem SPARQL⁵.

Sob a ótica dos indicadores e como são modelados, Hoornweg et al. (2007) propõem em seu artigo que estes elementos devem ser precisos, temporais e relevantes e, para que estas qualidades sejam avaliadas, deve-se observar se eles atendem os critérios a seguir:

- terem objetividade: bem definidos e não ambíguos, além de serem simples de se entender;
- serem mensuráveis e replicáveis;
- serem auditáveis: válidos e disponíveis para verificações por terceiros;
- serem comparáveis;
- estarem padronizados;
- serem efetivos: para tomadas de decisão, tanto no aspecto do planejamento quanto no gerenciamento;
- serem econômicos: tenham facilidade para coletar dados existentes.

Alguns dos critérios propostos por Hoornweg et al. (2007) para a modelagem de indicadores foram considerados nas alternativas desenvolvidas neste trabalho.

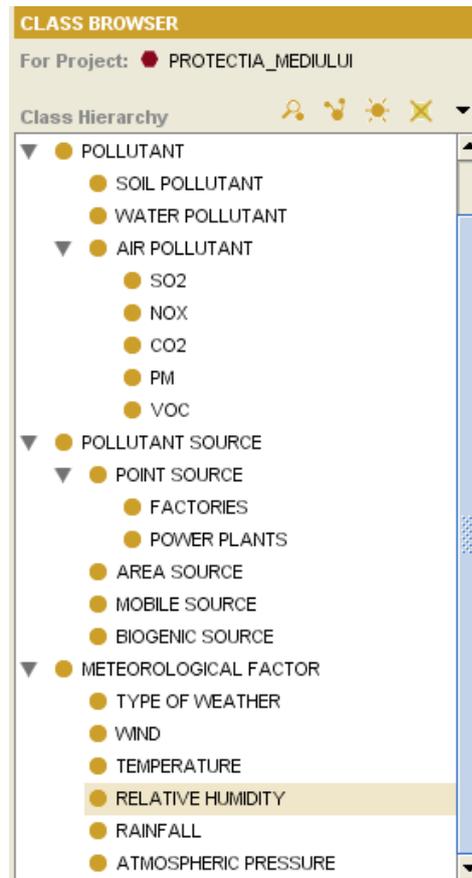
⁵SPARQL: Sparql Protocol and RDF Query Language.

2.2 ONTOLOGIAS DE POLUENTES COM INDICADORES

Segundo Oprea (2009), a prevenção é um passo importante para o controle da poluição do ar. Neste sentido, a autora propôs a adoção de sua ontologia, denominada de AIR_POLLUTION_Onto. Em seu projeto, a autora sugere que uma aplicação faça uso desta ontologia para prover uma base de conhecimento a ser utilizada para identificar o nível de qualidade do ar atrelado a cada poluente que for monitorado, comparando o valor obtido em seus indicadores com patamares, chamados por ela de valores máximos de concentração admissíveis (MAC).

Em síntese, dependendo de quanto o valor do indicador do poluente exceder ou não os valores máximos admitidos para cada poluente, um nível de poluição será atribuído dentre nove níveis possíveis. Os níveis foram ordenados e nomeados e, fazem parte do sistema criado em seu estudo de caso, chamado de DIAGNOZA_MEDIU. As comparações dos valores são realizadas por meio de regras implementadas na aplicação e os valores são obtidos de uma base de conhecimento compartilhável. Na Figura 3, pode se ver parte da hierarquia de classes, presente na ontologia que não está disponível para consulta ou uso, impossibilitando assim, que outras informações fossem obtidas para proveito deste trabalho.

Figura 3 – Hierarquia de classes da ontologia AIR_POLLUTION_Onto



Fonte: Oprea (2009)

3 FUNDAMENTAÇÃO TEÓRICA

3.1 ONTOLOGIAS

Segundo definição de Gruber (1995), uma ontologia é uma especificação explícita de uma conceitualização. O termo foi emprestado da filosofia, e neste campo entende-se que uma ontologia é uma descrição sistemática da existência.

Para melhor compreensão da definição de Gruber, deve-se entender melhor o que é conceitualizar, quando está relacionado às ontologias. De acordo com Guizzardi (2006), conceitualização é um conjunto de conceitos utilizados para estruturar abstrações do estado das coisas em um dado domínio (por exemplo, categorias e relações de hierarquia entre categorias). De outra forma, com quais conceitos um indivíduo organiza e estrutura a realidade percebida, independentemente da linguagem e do estado de coisas (*state of affairs*). Modelo, segundo o mesmo autor, é uma abstração de uma porção da realidade articulada segundo uma conceitualização de um domínio. Sendo, portanto, uma instanciação particular de uma conceitualização (por exemplo, classificação das espécies animais utilizando-se categorias e relações de hierarquia) dentre as várias instanciações possíveis. Logo, uma conceitualização permite ao indivíduo criar vários modelos.

Tanto conceitualização como modelo existem apenas na mente de um indivíduo. Isto significa que cada engenheiro de ontologias terá uma visão particular de um dado domínio, pois cada um tem uma forma própria de conceitualizar, podendo até serem similares, mas dificilmente iguais. Concretamente, em computação se trabalha com especificações de modelos mentais feitas em uma linguagem de modelagem que permite expressar (representar) conceitualizações.

A abordagem de Guizzardi (2006) encontra fundamentação nos trabalhos de Guarino e Giaretta (1995) que exploraram a definição proposta por Gruber. Para estes autores, é preciso entender o caráter intencional para que depois seja possível definir o que é uma ontologia. De fato, Guarino e Giaretta fazem uma crítica à Gruber dizendo que na definição de Gruber, o termo conceitualização faz referência a uma noção intuitiva de conceitualização que depende da extensão dos conceitos.

Guarino e Giaretta dizem que não se pode considerar uma teoria particular definida pela extensão dos conceitos como uma especificação de uma conceitualização, uma vez que esta representa apenas um estado de coisas dentre os muitos possíveis. E concluem que o que

se deve especificar é uma estrutura intencional capaz de representar um conjunto de estados de coisas.

Em Almeida e Bax (2003, p. 2) encontra-se uma explicação para a diferença entre os aspectos intencional e extensional:

uma definição intencional consiste de uma lista de características do conceito. Por exemplo, lâmpada incandescente é a lâmpada elétrica que emite luz a partir do aquecimento de um filamento pela corrente elétrica. A lâmpada incandescente é definida com o auxílio do gênero mais próximo (lâmpada elétrica) e de suas características. Uma definição extensional é uma enumeração de aspectos de todas as espécies que são do mesmo nível de abstração. Por exemplo, os planetas do sistema solar são Mercúrio, Vênus, Terra, Marte, Júpiter, Saturno, Urano, Netuno e Plutão (ISO-Standard 704).

A partir das críticas feitas à definição de Gruber, Guarino e Giaretta (1995, p. 25-32) a definem como: “a logical theory which gives an explicit, partial account of a conceptualization”. E, mais especificamente, no contexto da Inteligência Artificial, afirmam que:

an ontology refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words. This set of assumptions has usually the form of a first-order logical theory, where vocabulary words appear as unary or binary predicate names, respectively called concepts and relations.

3.2 LÓGICAS DE DESCRIÇÃO (DLS)

A família de DLs constitui-se em um fragmento da linguagem FOL (*First Order Logic*) ou Lógica de Primeira Ordem. A linguagem FOL é um formalismo lógico para representar relações ou propriedades entre objetos de um dado universo, e estas relações podem valer ou não, tomando-se por base um estado de coisas (SILVA, 2018). Além disso, FOL introduz os quantificadores universal e existencial na sua argumentação, conhecidos respectivamente pelas expressões: *For Every* (para todo) e *There Exists* (ao menos um).

As DLs são especialmente importantes para serem utilizadas na especificação de ontologias computacionais. Elas alteram a forma de tratamento dos elementos da FOL para que seja possível executar inferências em tempo razoável ao ser humano e para escapar da indecidibilidade da FOL.

Segundo Krotzsch et al. (2012), a maioria das DLs são fragmentos decidíveis da linguagem FOL. Entende-se por linguagem decidível, aquela que pode solucionar problemas de decisão por meio de seus algoritmos, isto é, ser capaz de provar ambos os casos: que uma

sentença é ou não é consequência lógica de uma teoria. A seguir, no Quadro 2, mostram-se as equivalências entre os construtos básicos destas duas linguagens.

Quadro 2 – Equivalência semântica entre construtos básicos de DL e FOL

DL	FOL	Significado
Conceitos	Predicados unários	Representam conjuntos de indivíduos
Papéis	Predicados binários	Representam relações binárias entre indivíduos
Nomes dos indivíduos	Constantes	Representam indivíduos

Fonte: Autoria própria.

Em DL existe uma separação na representação do conhecimento intencional e do estado das coisas. A caixa ou repositório terminológico (TBox) contém as definições axiomáticas de conceitos e papéis. E, a caixa de asserções (ABox) contém as asserções ou afirmações acerca de conceitos e indivíduos e entre papéis e indivíduos (instanciações). Logo, a TBox contém as definições em intenção (teoria) e a ABox, os modelos que satisfazem a teoria.

3.3 OWL – LINGUAGEM WEB PARA ONTOLOGIAS

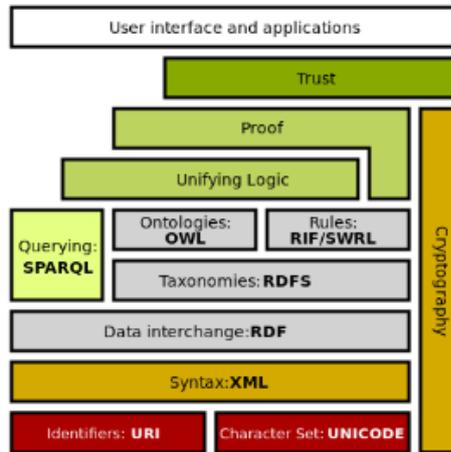
Em se tratando de ontologias, é primordial que se conheçam as suas linguagens de criação. As diversas linguagens para construção de ontologias fornecem diferentes funcionalidades e por isso devem ser avaliadas previamente ao seu uso. As linguagens precursoras são: OML – *Ontology Markup Language* (KENT, 2000), OIL – *Ontology Interchange Language* (FENSEL et al., 2001), DAML+OIL – *DARPA Agent Markup Language + OIL* (HORROCKS et al., 2002), entre outras.

Dentre elas, se destaca o padrão OWL de linguagens para ontologias, desenvolvidas no âmbito do W3C-*World Wide Web Consortium*⁶. As linguagens OWL se baseiam na linguagem predecessora DAML+OIL. OWL é o acrônimo para *Web Ontology Language* e de acordo com a W3C, é uma linguagem para *Web Semântica* interpretável por programas, projetada e recomendada para representar conhecimento declarativo sobre coisas, grupos de coisas e relações entre coisas. OWL foi desenvolvida sobre o modelo de dados definido pela RDF – *Resource Description Framework* que, por sua vez, se utiliza da sintaxe XML – *Extensible Markup Language*. A RDF se caracteriza por declarar informações sobre recursos

⁶Disponível em: <<http://www.das.ufsc.br>>.

por meio das triplas formadas por sujeito-predicado-objeto. A Figura 4 apresenta a pilha de camadas da *Web Semântica*.

Figura 4 – Pilha de camadas da Web Semântica.



Fonte: Open Geospatial Consortium (2015)

Quanto à linguagem OWL DL, pode-se afirmar que uma ontologia desenvolvida nesta linguagem não descreve apenas um estado de coisas particular, como fazem os bancos de dados. Graças a separação das definições intencionais que são armazenadas na TBox e das asserções sobre indivíduos e suas relações declaradas na ABox, ela pode descrever vários mundos possíveis. A linguagem OWL DL faz uso de recursos presentes na lógica de descrição (DL). Isto se dá nos casos listados no Quadro 3 a seguir.

Quadro 3 – Elementos da lógica de descrição (DL) empregados em OWL DL

Casos	Construtos
Restrições de Propriedade	<i>hasValue</i> (valor específico)
Restrições de Cardinalidade	<i>Cardinality</i> (intervalo)
Restrições sobre Classes	<i>disjointWith</i>
Construtores de Classes	<i>OneOf</i> (classes de enumeração) <i>unionOf</i> <i>complementOf</i> <i>intersectionOf</i>

Fonte: Autoria própria.

No Quadro 4, podem ser vistas as duas formas para se descrever uma classe de forma completa, isto é, utilizando-se a equivalência lógica da linguagem OWL DL na sintaxe Manchester⁷ proposta em Horridge et al. (2006) para definir um axioma de igualdade e, por outro lado, como seria a mesma expressão em linguagem DL.

⁷Sintaxe Manchester: formato que representa conectivos lógicos em linguagem natural (inglês) em vez de símbolos gráficos o que aumenta a legibilidade das sentenças.

Quadro 4 – Exemplo comparativo das representações para a definição de uma classe/conceito

OWL DL na sintaxe Manchester	DL
Classe: PizzaDeQueijo Equivalent To: Pizza AND (temRecheio SOME Queijo)	Conceito: PizzaDeQueijo $PizzaDeQueijo \equiv$ $Pizza \cap \exists temRecheio. Queijo$

Fonte: Autoria própria.

A linguagem OWL (versão 1) está subdividida em três linguagens projetadas para uso por diferentes tipos de desenvolvedores. Estas linguagens são:

- OWL Lite: de menor alcance e menos recursos, atende principalmente as necessidades básicas da representação de conhecimento como, a classificação hierárquica de termos e a definição de conceitos usando restrições simples;
- OWL DL: de grande alcance, mas com garantia computacional pelos *reasoners*, inclui todas as características de OWL Lite e leva o termo DL (Description Logic) no nome, por estar alinhada à lógica de descrição, exige a separação de tipos – classe, propriedade e indivíduo;
- OWL Full: de alcance bem extenso, inclui todas as funcionalidades de OWL DL, atende aqueles usuários que querem o máximo de recursos para expressar o conhecimento e também liberdade sintática para uso de RDF, mas sem garantias computacionais, tanto que, é improvável que os *reasoners*, sejam capazes de suportar todas as funcionalidades desta versão.

A OWL 2 introduziu três sublinguagens de OWL DL denominadas de perfis (*profiles*) projetadas para diferentes usos: OWL 2 EL, QL e RL (KRÖTZSCH, 2012).

O uso do perfil EL é adequado para ontologias grandes com muitas definições de conceitos. O perfil QL é destinado primordialmente a responder queries de modo eficiente, portanto útil na integração de *schemas* de bancos de dados relacionais. Este perfil permite que o armazenamento de dados e a execução das queries sejam delegadas a sistemas gerenciadores de Bancos de Dados Relacionais comuns (BDRs). OWL RL é adequada à representação de regras, portanto indicado para aplicações que necessitam de raciocínio baseado em regras sem perder expressividade. Pode ser implementada utilizando tecnologias baseadas em regras como Prolog e Jess. Os três perfis se diferem tanto em expressividade quanto em eficiência de raciocínio.

3.4 PADRÕES DE MODELAGEM

Nas ontologias são utilizados padrões de modelagem para prover clareza e consistência anteriormente verificada. Os padrões são utilizados em muitas áreas como modelos ou descrições abstratas que contém as melhores práticas em algum campo de estudo (GANGEMI, 2010).

Foi no final dos anos 70 que o termo *design pattern* foi introduzido pelo arquiteto e matemático Christopher Alexander para propor o uso de um conjunto de regras empacotadas em forma de um padrão, com o propósito de ajudar a resolver problemas recorrentes em projetos (GANGEMI, 2005).

No campo das ontologias, os padrões são genericamente chamados na língua inglesa pelo termo “Ontology Design Patterns (ODP)”, expressão que pode ser traduzida para “Padrões de Modelagem de Ontologias”. Segundo Gangemi e Presutti (2009), um padrão de modelagem para ontologias é uma solução de modelagem para resolver um problema recorrente na criação de ontologias.

Estes ODPs se subdividem em categorias conforme as suas características. Os autores propõem seis famílias de padrões: estruturais, correspondentes, conteúdos, *reasoning*, apresentação e léxico-sintáticos (vide Quadro 5). A família, na qual se enquadram os padrões a serem elaborados neste trabalho, está ligada aos padrões de conteúdo porque procuram solucionar problemas das classes e relações que integram uma ontologia.

Quadro 5 – Famílias de padrões de modelagem (ODP)

Tipo de Padrão	Descrição
Estrutural	Composições de construtos lógicos para resolver um problema de expressividade
Correspondente	Construtos para solucionar o problema de transformar um recurso não ontológico em uma nova ontologia
Conteúdo	Construtos para resolver problemas de modelagem dos conceitos e propriedades que integram uma ontologia
<i>Reasoning</i>	Construtos que servem para informar ao sistema, qual <i>reasoning</i> deve ser executado na ontologia para atender consultas e avaliações
Apresentação	Construtos que visam suportar a reusabilidade das ontologias sob a ótica do usuário
Léxico-sintático	São estruturas linguísticas ou esquemas formados por certos tipos de termos que seguem uma ordem específica

Fonte: Adaptado de Gangemi e Presutti (2009)

Os autores apresentam uma abordagem técnica para a descrição de padrões de conteúdo, nomeados na língua inglesa por *Content Patterns* (CP). Na descrição, deve ser provido um conjunto de preceitos, os quais atuam como um registro de catálogo.

A seguir, propõe-se o formato dos padrões, adotado neste trabalho. Este formato está baseado nas orientações descritas pelos autores acima referidos, com a finalidade de representar padrões de modelagem para ontologias. Dos preceitos propostos, são adotados aqueles que trazem os maiores benefícios para a representação de conteúdo, de modo que as informações essenciais são descritas para uso pelo modelador. Logo, os parâmetros escolhidos para a composição dos padrões neste trabalho são:

- Nome do padrão;
- Questões de competência ou cenário – contém pelo menos uma questão que deve ser respondida pela base de conhecimento da ontologia em conjunto com o próprio padrão de modelagem;
- Elementos – classes e relações na sintaxe Manchester;
- Diagrama – diagrama de classes em UML para facilitar a compreensão do padrão.

3.5 PADRÕES ELEMENTARES

A partição de valores, por exemplo, é um recurso de modelagem adequado quando se quer descrever várias características afins relacionadas a um determinado conceito, e somente um deles será selecionado para representar o conceito, quando são aplicados os critérios das propriedades relacionadas.

Por exemplo, podemos ter que descrever que uma classe pode assumir um dos valores: quente, morno ou frio. A W3C propõe a adoção de um padrão para se declarar uma partição de valores, apresentado na continuidade.

Tome-se uma classe que deverá conter subclasses que a representam, mas de maneira mutuamente exclusiva. Por exemplo, a classe ‘temperatura’ que deve ser particionada em três subclasses: ‘quente’, ‘morna’ e ‘fria’. Estas subclasses são descritas por duas propriedades que servirão para delimitar seus valores mínimo e máximo em intervalos diferentes, porém complementares. Então, pode-se dizer que estas três subclasses completam totalmente a classe superior, ou seja, a união destas subclasses é igual à superclasse.

Para que isso seja de fato verdadeiro, deve-se defini-las como subclasses disjuntas. Pode-se ver a seguir, uma declaração simplificada das subclasses que compõem esta classe, na

representação em sintaxe Manchester: **DisjointClasses: Fria, Morna, Quente**. E, na Figura 5, o padrão para definição da partição de valores também na sintaxe Manchester.

Figura 5 – Padrão: Partição de valores

```
'ValuePartition:' className
objectPropertyID '['classID classID {classID}']'
```

Fonte: W3C (2018)

A disjunção é outro recurso importante para a modelagem de ontologias. A disjunção tem a função de garantir que não exista sobreposição de papéis entre membros de uma classe superior. Ou seja, um indivíduo membro de uma das classes não poderá ser simultaneamente, membro de outra classe existente na partição.

Para exemplificar o uso do formato proposto previamente, toma-se o seguinte cenário para representar o padrão referente à classe 'temperatura':

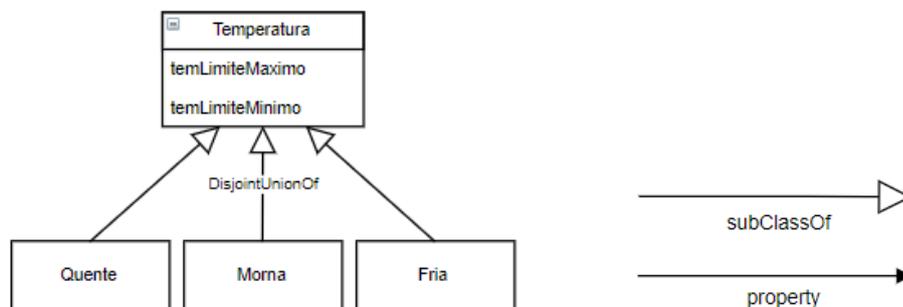
- Nome do padrão: temperatura;
- Cenário: a temperatura de uma substância pode assumir somente um dos níveis possíveis, isto é: quente, morno ou frio, e os valores que delimitam estes níveis são descritos pelas propriedades **temLimiteMaximo** e **temLimiteMinimo**;
- Elementos (em sintaxe Manchester):

```
Class: <tp:Temperatura>

SubClassOf:
  <tp:temLimiteMaximo> exactly 1 xsd:decimal,
  <tp:temLimiteMinimo> exactly 1 xsd:decimal

DisjointUnionOf:
  <tp:Fria>, <tp:Morna>, <tp:Quente>
```

- Diagrama (de classes em UML):



3.6 CRITÉRIOS PARA AVALIAÇÃO DE MODELAGENS

Com relação às ontologias, existem várias abordagens para fazer uma avaliação de seus conteúdos. Ghahremanloo, Thom e Magee (2012) descrevem três maneiras para se fazer isso, cada qual se baseia em um aspecto diferente, são elas: baseada em tarefas, baseada em critérios ou baseada nos aspectos conceituais e léxicos. Por léxico devemos entender o seu aspecto na questão da uniformidade na composição dos termos referentes ao tema referenciado pela ontologia e, da área de conhecimento presentes na ontologia.

Na abordagem baseada em tarefas, a ontologia é avaliada pela sua capacidade em completar as tarefas à que se propõe com um desempenho adequado (GHAHREMANLOO; THOM; MAGEE, 2012). Por outro lado, a abordagem baseada em critérios faz a avaliação no sentido da consistência, da completude, da concisão, da expansibilidade e da sensibilidade como explica Gómez-Peréz (1996).

Segundo a autora, a consistência trata da possibilidade de se obter conclusões contraditórias a partir de dados de entrada válidos. A concisão está na utilidade e precisão das informações presentes na ontologia sem implicar na ausência de redundâncias. A expansibilidade pode ser medida pela facilidade ou não em se adicionar novas definições ou novas informações nas definições existentes sem prejuízo da integridade da ontologia. E a sensibilidade está relacionada com o efeito produzido por pequenas alterações nas definições do modelo ontológico, o qual não deve afetar sua qualidade.

A avaliação baseada em tarefas está atrelada ao desempenho da ontologia em executar operações sobre o modelo para atender uma aplicação usuária. Este modo de avaliação foi explorado por Yu et al. (2009), os quais propõem uma metodologia de avaliação orientada por requisitos chamada de ROMEO⁸. Nesta técnica comparativa são utilizados requisitos genéricos como: competência, funcionalidade, capacidade e padronização. E medições são tomadas na obtenção de respostas a algumas questões de competência para serem comparadas em função destes requisitos.

Outras abordagens foram sugeridas em trabalhos precusores como em Maedche e Staab (2002), onde os autores realizam um experimento no qual ontologias produzidas por estudantes não especialistas sejam comparadas com uma ontologia modelada por um engenheiro especialista denominada de “*gold standard*”. Porzel e Malaka (2004) propõem uma avaliação dos resultados pelo uso da ontologia por uma aplicação. Por outro lado,

⁸ROMEO: *Requirements Oriented Methodology for Evaluating Ontologies*.

Brewster et al. (2004) sugerem a comparação do domínio coberto pela ontologia com material de uma fonte relevante que trata do assunto abordado.

Assim, quanto mais conceitos presentes na ontologia estiverem também na documentação, melhor será o seu conteúdo.

Em síntese, há várias formas e critérios para se avaliar ontologias. Os autores que exploraram esta temática têm opiniões razoavelmente convergentes em termos dos critérios a usar nas avaliações. Em outro aspecto, relativo à ordem de avaliação, Gómez-Pérez (1996) defende uma avaliação baseada em critérios antes de disponibilizá-las para o uso de uma aplicação, diferentemente de Porzel e Malaka (2004) que sugerem que a ontologia seja diretamente avaliada por meio de uma aplicação. O Quadro 6 sumariza os tipos de abordagens principais encontrados na literatura, associadas aos critérios que são aplicados em cada técnica proposta e seus autores. Os critérios sublinhados foram os escolhidos para a avaliação das ontologias desenvolvidas com as modelagens propostas no trabalho.

Quadro 6 – Comparativo das abordagens e critérios de avaliação de ontologias.

Autores	Abordagem	Critério(s)
Ghahremanloo, Thom e Magee 2012	Baseada em tarefas	Desempenho no cumprimento de cada tarefa para a aplicação
Yu et al. 2009	ROMEO – baseada em tarefas	Competência, funcionalidade, capacidade e padronização
Gómez-Pérez 1996	Baseada em critérios	Consistência, <u>completude</u> , <u>concisão</u> , <u>expansibilidade</u> e sensibilidade
Maedche e Staab 2002	Comparação com uma ontologia de padrão ouro: “Gold Standard”	Uniformidade dos conceitos utilizados, aspectos léxicos e sintáticos da linguagem empregada
Porzel e Malaka 2004	Baseada pelo uso por uma aplicação	<u>Qualidade dos resultados produzidos para a aplicação</u> e o desempenho alcançado na execução de tarefas
Brewster et al. 2004	Baseada em dados	Quantidade de termos que aparecem na ontologia que estão em documentos relevantes sobre o domínio tratado.

Fonte: Autoria própria.

Em relação aos critérios, é importante buscar um entendimento dos termos quanto ao desenvolvimento das ontologias. Para Gómez-Pérez (1996), consistência é a capacidade de

uma ontologia em permitir somente conclusões não contraditórias quando se utilizam dados de entrada válidos. Isto é, as conclusões obtidas a partir da interpretação de seus conceitos e relações devem permanecer as mesmas, ou seja, existe estabilidade nos resultados que são gerados. Ainda, segundo a autora, outro critério importante é a completude, para indicar se em uma ontologia todas as definições necessárias para representar um domínio estejam declaradas ou possam ser inferidas por meio das outras definições presentes no seu modelo.

Para a autora, o processo para avaliar este critério é procurar por falhas no conteúdo do modelo, testando as três características negativas que devem ser evitadas na modelagem da ontologia, isto é:

- a. escopo limitado – que não possibilita o reuso e o compartilhamento das definições por outras aplicações;
- b. definições incompletas – definições que não estão exhaustivamente ou profundamente definidas;
- c. granularidade deficitária – nível de detalhe insuficiente nas definições.

Seguem as explicações da mesma autora para os termos concisão, expansibilidade e sensibilidade no contexto da qualidade de uma ontologia. Concisão é a propriedade apresentada por uma ontologia quando esta não possui redundâncias de definições e nem definições não utilizadas. Expansibilidade trata do esforço necessário para adicionar novas definições sem afetar o conjunto de definições existentes que já foram validadas previamente. Em relação à sensibilidade, se deseja que uma ontologia esteja livre de erros mesmo depois de serem aplicadas pequenas alterações nas definições, mantendo-se correta léxica e sintaticamente.

O termo qualidade, neste contexto, se refere à propriedade de produzir respostas corretas quando consultada, com o mínimo de ruídos – *precision* e, o máximo de cobertura - *recall*. Usualmente, isto é realizado em relação às respostas trazidas para as questões de competência.

De posse dos critérios de avaliação, prossegue-se com a busca por trabalhos correlatos que fazem uso de declarações de indicadores em suas modelagens. Deste modo, exploram-se outras formas de modelagem passíveis de serem utilizadas na construção de uma ontologia no domínio do monitoramento da poluição do ar.

3.7 ÍNDICES SOBRE POLUIÇÃO DO AR

Índices ambientais foram elaborados em bases científicas por entidades reguladoras, para uso no monitoramento da poluição ambiental. Estes índices foram estudados e serviram como base para as modelagens apresentadas neste trabalho.

Dentre os diversos índices existentes, como dos Estados Unidos, México, Canadá e Reino Unido, além do próprio índice adotado pelo estado do Paraná no Brasil, gerenciado pelo Instituto Ambiental do Paraná (IAP), foram pesquisados com mais detalhes aqueles índices que são produzidos por métodos mais complexos e que são utilizados como referências por seus pares (vide Quadro 7).

Quadro 7 – Relação dos índices pesquisados

Nome (sigla)	Responsável	URL
Air quality index (AQI)	EPA/USA	< https://airnow.gov/aqi/ >
Air quality and health index (AQHI)	Government of Canada/Canada	< https://www.canada.ca/en/environment-climate-change/services/air-quality-health-index/use.html ou https://www.canada.ca/en/environment-climate-change/services/air-quality-health-index/understanding-messages.html >
Índice Metropolitano de la Calidad del Aire (IMECA)	Gobierno del Distrito Federal/México	< http://www.aire.cdmx.gob.mx/descargas/monitoreo/normatividad/NADF-009-AIRE-2006.pdf >
Índice de qualidade do Ar (IQA)	IAP/Paraná	< http://www.iap.pr.gov.br/pagina-1413.html >
Daily air quality index (DAQI)	DEFRA/United Kingdom	< https://uk-air.defra.gov.uk/air-pollution/daq?view=more-info >

Fonte: Autoria própria.

É importante lembrar que os órgãos ambientais locais e internacionais utilizam periodicidades similares que dependem do poluente sendo monitorado. Tais índices têm semelhanças e diferenças bem marcantes. Por exemplo, o índice do México (IMECA) está vinculado a cada poluente. Em seu cálculo, se aplicam fórmulas distintas por poluente e para cada nível de concentração obtido para os indicadores.

Por outro lado, o índice do Reino Unido (DAQI-UK) não faz uso de fórmulas, mas sim, de uma função de agregação, isto é, pelo pior valor de concentração dentre cinco poluentes. Neste caso, o índice é derivado de um conjunto de indicadores associados aos

cinco poluentes sendo que as medições de cada poluente são processadas em diferentes periodicidades para a geração dos indicadores.

Outra abordagem é seguida pelos índices dos Estados Unidos (AQI) e do Canadá (AQHI). Ambos têm somente uma fórmula para ser executada, diferentes entre si. E o valor final obtido será enquadrado em uma de suas faixas dos níveis de poluição ambiental. O índice AQHI se baseia somente em indicadores de três poluentes, diferentemente do índice DAQI que utiliza cinco poluentes, enquanto que outros índices eventualmente monitoram mais poluentes. Isto vai depender do equipamento utilizado e da presença significativa do poluente no local onde ocorre a coleta, tornando necessário seu monitoramento.

3.7.1 Composição do Índice IMECA – Índice Metropolitano de la Calidad del Aire

O índice mantido pela Secretaria de Meio Ambiente do Distrito Federal (Cidade do México) teve seus requisitos estabelecidos pela Norma Ambiental de 2006. Este índice é composto da seguinte forma:

- os poluentes do ar monitorizados periodicamente, são: O₃, NO₂, SO₂, CO, PM₁₀ e PM_{2.5};
- os períodos de monitoramento variam de acordo com o poluente, por ex. SO₂, PM_{2.5} e PM₁₀ são mapeados a cada 24h, o dióxido de nitrogênio é mapeado a cada 1h, ...;
- a densidade de massa dos poluentes é medida em µg/m³ ou ppm, por exemplo, as partículas suspensas são medidas em µg/m³ e os demais poluentes, medidos em ppm;
- o valor médio obtido das medições para um dado período, para cada poluente se torna um indicador;
- os indicadores estão categorizados em faixas de concentração;
- as faixas de categorização são representadas por um valor mínimo e um valor máximo, por exemplo: O₃ medido a cada 1h, tem os valores na faixa *Buena* [0.000,0.055], na faixa *Regular* [0.056,0.110], ...;
- estes valores são utilizados nas fórmulas de cálculo do índice deste poluente;
- o índice IMECA está categorizado em faixas de concentração válidas para todos os poluentes (vide Figura 6);
- estas faixas também são representadas por um intervalo de valores, mínimo e máximo e, por uma descrição, por exemplo: a faixa *Buena* [0,50], *Regular* [51,100], ..., *Extremadamente Mala* [201, +∞];
- o índice IMECA está vinculado a um indicador e, portanto, está atrelado a um dos poluentes;

- o índice IMECA tem várias fórmulas que estão vinculadas às categorias (faixas de concentração dos indicadores) e às suas próprias categorias, de modo que é preciso enquadrar o valor obtido para o seu indicador em uma categoria de indicador e uma categoria de índice, para saber qual a fórmula correta a ser empregada no cálculo final do índice.

Figura 6 – Categorias e Faixas de Valores do Índice IMECA



Fonte: Secretaría del Medio Ambiente, Mexico (2006)

A Figura 7 ilustra as faixas de valores do índice IMECA e do indicador referente ao poluente PM_{10} , com as fórmulas associadas às suas categorias. Assim como para este poluente, existem categorias e fórmulas para cada um dos demais poluentes que são tratados neste índice de poluição do ar que não foram aqui copiados, mas estão disponíveis na norma.

Figura 7 – Categorias e Fórmulas para o Índice IMECA referente ao Poluente PM_{10}

PARTÍCULAS MENORES A 10 MICROMETROS (PM_{10})			
Intervalo del IME CA	Intervalo de concentraciones [$\mu\text{g}/\text{m}^3$]	Ecuaciones de transformación	Ecuaciones de transformación simplificadas
0 – 50	0 – 60	$I[PM_{10}] = C[PM_{10}] * 50/60$	$I[PM_{10}] = C[PM_{10}] * 5/6$
51 – 100	61 – 120	$I[PM_{10}] = 50-60 * 50/60 + C[PM_{10}] * 50/60$ $I[PM_{10}] = C[PM_{10}] * 50/60$	
101 – 150	121 – 220	$I[PM_{10}] = 100-120 * 50/100 + C[PM_{10}] * 50/100$ $I[PM_{10}] = 40 + C[PM_{10}] * 50/100$	$I[PM_{10}] = 40 + C[PM_{10}] * 0.5$
151 – 200	221 – 320	$I[PM_{10}] = 150-220 * 50/100 + C[PM_{10}] * 50/100$ $I[PM_{10}] = 40 + C[PM_{10}] * 50/100$	
> 200	> 320	$I[PM_{10}] = C[PM_{10}] * 200/320$	$I[PM_{10}] = C[PM_{10}] * 5/8$

Fonte: Secretaría del Medio Ambiente, Mexico (2006)

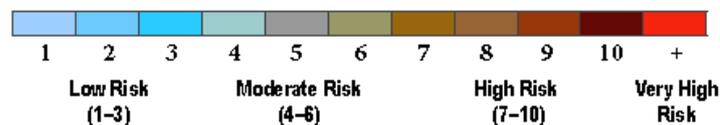
3.7.2 Composição do Índice AQHI – Air Quality Health Index

O índice AQHI que se pode traduzir como: Índice de Qualidade do Ar e Saúde, foi criado pelos órgãos federais do governo do Canadá, isto é, *Environment Canada* (EC) e *Health Canada* (HC), com o intuito de auxiliar o povo canadense a tomar decisões a respeito dos riscos da poluição do ar. Este índice está composto da seguinte forma:

- uma escala para indicar o risco de exposição à poluição do ar (vide Figura 8);

- o índice está categorizado em faixas de valores para identificação do nível de risco de exposição;
- as faixas de categorização do índice são representadas por um valor mínimo e um valor máximo, por exemplo: na faixa de baixo risco: *LowRisk* [1,3]; na faixa de risco moderado: *ModerateRisk* [4,6], ...;
- monitoramento para produção do índice ocorre para os poluentes O₃, NO₂ e PM_{2.5};
- indicadores destes poluentes são gerados em períodos de 3 horas;
- aplicação de uma fórmula única com uso dos três indicadores para cálculo do índice (vide Figura 9).

Figura 8 – Categorias e Faixas de Valores do Índice AQHI



Fonte: <http://www.airhealth.ca> (2018)

Figura 9 – Fórmula de Cálculo do Índice AQHI

$$AQHI = \left(\frac{1000}{10.4}\right) \times [(e^{0.000537 \times O_3} - 1) + (e^{0.000871 \times NO_2} - 1) + (e^{0.000487 \times PM_{2.5}} - 1)]$$

Fonte: Stieb, D. et al. (2008)

3.8 CARACTERÍSTICAS DOS POLUENTES DO AR

Os agentes poluentes do ar se dividem em duas grandes categorias, que se baseiam na forma como se apresentam na atmosfera. Quando são produzidos por atividades humanas, são chamados de antropogênicos. Quando estão presentes na natureza independentemente de ação humana, são chamados de naturais.

No Quadro 8 são apresentadas informações referentes aos poluentes, incluindo a quantidade de horas padrão para geração de um indicador. A densidade de massa⁹ destes poluentes é normalmente medida em µg/m³, em ppb ou em ppm.

⁹Unidades de medida de massa: ppb/ppm – partes por bilhão/milhão, µg/m³ – microgramas por metro cúbico.

Quadro 8 – Características dos poluentes do Ar

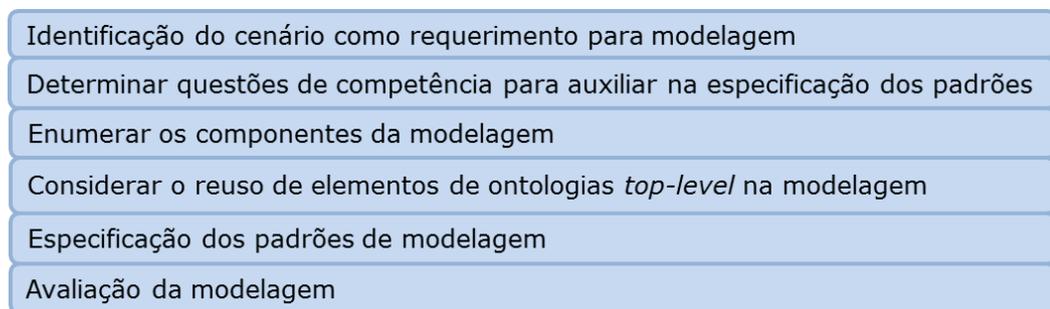
Poluente	Símbolo	Categoria	Periodicidade*
Ozônio	O ₃	Antropogênico	1h
Dióxido de Nitrogênio	NO ₂	Antropogênico	1h
Monóxido de Carbono	CO	Antropogênico	8h
Dióxido de Enxofre	SO ₂	Antropogênico	24h
Partículas totais em suspensão ou Partículas inaláveis	PTS ou PI	Natural	24h
Partículas suspensas menores do que 10µm	PM ₁₀	Natural	24h
Partículas suspensas menores do que 2.5µm	PM _{2,5}	Natural	24h
Fumaça	FMC	Natural	24h
(*) A periodicidade pode variar de um país para outro. Os períodos informados estão baseados nas normas do IAP/Pr.			

Fonte: Autoria própria.

4 METODOLOGIA

Nesta seção, descreve-se a metodologia adotada nesta dissertação para desenvolvimento dos padrões de modelagem. As etapas relacionadas a seguir, são utilizadas nas duas modelagens apresentadas nas seções 5 e 6. Sua ordem, em parte se baseou em Gruninger e Fox (1995) e foi dividida em seis etapas mostradas na Figura 10 e detalhadas na sequência.

Figura 10 – Etapas da metodologia



Fonte: Autoria própria.

4.1 IDENTIFICAÇÃO DO CENÁRIO COMO REQUERIMENTO PARA MODELAGEM

Gruninger e Fox (1995) propõem como primeiro passo de desenvolvimento de uma ontologia, a identificação do cenário no qual os produtos de uma modelagem serão empregados. Esta etapa é necessária principalmente para:

- identificar os conceitos principais que fazem parte do domínio proposto e como se relacionam;
- auxiliar a elaboração das questões de competência que devem ser respondidas pela ontologia depois de construída com os padrões propostos;
- compreender como o modelo, por meio de uma aplicação, deverá subsidiar a interação com seus usuários;
- determinar seu escopo de atuação.

Esta etapa iniciou-se com o levantamento do referencial teórico sobre índices e indicadores (conhecimento do domínio). Este levantamento é útil para compreender e identificar os pontos em comum dos vários índices e indicadores e igualmente, suas diferenças. Esta etapa também contribui para se encontrar os termos mais comuns relacionados ao domínio que está sendo tratado e as suas inter-relações.

4.2 DETERMINAR QUESTÕES DE COMPETÊNCIA PARA AUXILIAR NA ESPECIFICAÇÃO DOS PADRÕES

Os autores Gruninger e Fox (1995) propõem que no segundo passo se deve enunciar questões de competência informais para as quais se deseja obter respostas com cada um dos modelos a desenvolver. Nesta abordagem, optou-se por descrever as questões em modo formal, antecipando a etapa 4 proposta pelos referidos autores. As questões de competência são escritas em linguagem natural e tem por objetivo guiar a construção do padrão e servir de referência aos potenciais usuários de um sistema fictício que fará uso da ontologia a ser construída com os padrões desenvolvidos.

Ren et al. (2014) descrevem a identificação dos padrões das questões de competência pelo método de modelagem por características apresentados em Palmer e Felsing (2001) para se obter os arquétipos mais frequentes nos quais as questões de competência se enquadram.

As características principais a serem empregadas pelas questões de competência são: a aridade do predicado, o tipo de relação entre elementos, o modificador empregado para impor restrições e a independência de um domínio. Além disso, relatam que questões de competência são normalmente compostas de classes, propriedades de objetos ou propriedades de dados.

4.3 ENUMERAR OS COMPONENTES DA MODELAGEM

O passo seguinte é a complementação da primeira etapa. Uma vez que o cenário está delineado, é possível conhecer quais são os elementos relevantes para o modelo. Em outras palavras, quais são os conceitos e relações entre os objetos que farão parte do modelo.

Uma das formas utilizadas para realizar esta etapa é a de relacionar primeiramente os conceitos conforme o Quadro 9, representados geralmente por termos substantivos. A partir deste quadro, é possível criar uma hierarquia entre os conceitos definindo relações de generalização-especialização entre eles (ou classe-subclasse em OWL DL). A hierarquia é usualmente representada em um diagrama.

Quadro 9 – Modelo para enumeração dos termos a serem modelados como classes.

Termo	Nome da Classe	Descrição	Superclasse
<em linguagem natural>	<p:nome adotado no modelo; p designa o prefixo da ontologia de origem em caso de reuso>	<descrição em linguagem natural>	<nome da(s) classe(s) diretamente superior(es)>

...
-----	-----	-----	-----

Fonte: adaptado de Fernández-López e Juristo (1997).

Em seguida, identificam-se as relações entre conceitos conforme o Quadro 10. Nem sempre é possível definir todos os elementos das relações em uma única iteração. Nas primeiras iterações, normalmente as definições são mais relaxadas (sem definição de domínio e imagem ou tipo) sendo refinadas com a sucessão de iterações.

Quadro 10 – Modelo para enumeração dos termos a serem modelados como propriedades.

Termo	Nome da propriedade	Descrição	Domínio x Imagem	Tipo
<em linguagem natural>	<p:nome adotado no modelo; p designa o prefixo da ontologia em caso de reuso>	<descrição em linguagem natural>	<definir domínio e imagem da relação>	<funcional, inversa de <X>, simétrica, reflexiva, transitiva>
...				

Fonte: Adaptado de Fernández-López e Juristo (1997).

4.4 CONSIDERAR O REUSO DE ELEMENTOS DE ONTOLOGIAS *TOP-LEVEL* NA MODELAGEM

Esta etapa é tradicionalmente indicada para fazer o reaproveitamento de conceitualizações já desenvolvidas e presentes em ontologias de alto nível. O reuso de elementos de outras ontologias, sejam classes ou propriedades, é frequentemente indicado por autores especialistas em modelagem de ontologias, como em Villazón-Terrazas et al. (2011), Noy e McGuinness (2001) e Panetto (2016). Além disso, Fernández-López et al. (2013) desenvolveram uma metodologia para auxiliar os engenheiros a aplicar a reusabilidade quando forem criar o seu próprio modelo.

Entende-se, portanto, que seja relevante para a construção dos modelos, a busca pela reusabilidade dos componentes que irão compor a ontologia. O que se vislumbra aqui é a economia de esforço de trabalho, o aproveitamento de conceitos que já foram validados por outras aplicações trazendo mais consistência aos padrões que serão criados, ganhos estes corroborados por Keet (2011) no contexto de reuso de construtos de ontologias fundacionais. É importante mencionar também que a adoção do reuso viabiliza a construção da Web Semântica – permitindo a semantificação de objetos na rede.

Para localizar ontologias similares, efetuam-se pesquisas por meio de buscadores de páginas *web* utilizando-se termos identificados como importantes na etapa anterior e, também,

por exploração das referências e citações encontradas em artigos científicos. Além das ontologias, buscam-se também manuais de referências e artigos que explicam as intenções das definições encontradas nestas ontologias. Tais ontologias são normalmente publicadas por organizações (ex. Dublin Core, W3C, ISO) e grupos de pesquisa.

Localizada uma ontologia, seu potencial de reuso é avaliado com base na experiência e conhecimento do assunto pelo avaliador. O avaliador realiza uma análise de similaridade das classes e propriedades com os termos relacionados às questões de competência identificadas e aos termos, todos já identificados anteriormente.

Uma vez identificadas às ontologias com potencial de reuso, correlaciona-se os componentes principais definidos na etapa anterior com os existentes nestas ontologias. Portanto, ao final desta atividade, haverá um conjunto revisado de classes e propriedades em relação ao que foi produzido na etapa anterior devido às importações realizadas das ontologias de reuso. Também se obtém uma lista de ontologias cujos elementos foram reusados com as respectivas referências textuais (ex. artigos, manuais).

4.5 ESPECIFICAÇÃO DOS PADRÕES DE MODELAGEM

Nesta etapa, aplica-se o formato definido na Seção 3.4, para a representação dos padrões de modelagem das principais classes que compõe a hierarquia da ontologia.

A especificação se dá: (i) pela escolha de um nome para o padrão inicialmente; (ii) elaboração de uma questão de competência por meio do uso dos termos e relações enumerados na etapa anterior; (iii) descrição em sintaxe Manchester dos elementos que compõem a definição do padrão, conforme apresentado na seção anterior; (iv) adicionalmente, um diagrama representativo em UML pode ser utilizado para ilustrar os elementos que fazem parte do padrão em modo gráfico.

4.6 AVALIAÇÃO DA MODELAGEM

Nesta etapa, avalia-se a modelagem por meio de dois estudos de caso que fazem uso dos padrões propostos. Além disso, é realizada uma confrontação dos padrões de modelagem em relação às questões de competências elaboradas. Procura-se responder se a modelagem serve ao seu propósito e se o modelo construído consegue responder às questões de competência sugeridas.

Na fundamentação teórica foi visto que existem maneiras diversas para se avaliar ontologias, por exemplo: por tarefas, desempenho, critérios, etc. Este trabalho se baseou em critérios, os quais, procuram avaliar aspectos referentes ao aproveitamento dos objetos, à quantidade, à redundância, à possibilidade de expansão dos axiomas e à qualidade dos resultados obtidos pelo uso das ontologias. Todos são importantes, e para efeito deste trabalho, são utilizados os critérios de: expansibilidade, completude, concisão e qualidade nas saídas geradas. A escolha destes critérios se deu pelas seguintes razões:

- expansibilidade: indicar a facilidade para acrescentar novas definições ou aprimorar as definições existentes;
- completude: verificar se tudo que se supõe estar descrito na ontologia ou pode ser inferido dela, de fato está apresentado e, para confirmar que ela pode responder as questões de competência (GRUNINGER; FOX, 1995);
- concisão: avaliar a existência de definições não necessárias ou sem utilidade;
- qualidade nas saídas geradas: confrontar saídas produzidas do modelo com os resultados esperados de casos de testes (questões de competência).

Para poder demonstrar a avaliação das modelagens de forma mais objetiva e sucinta, propõe-se uma avaliação qualitativa baseada em uma análise de desempenho da ontologia em relação a cada critério proposto. Adicionalmente, se propõe pontuar o desempenho pelo uso de uma escala numérica de 1 a 3 (inteiros) que significam, respectivamente: não atendido, parcialmente atendido e atendido. Esta avaliação pode ser demonstrada pelo uso do Quadro 11 que contém os itens de avaliação.

Quadro 11 – Quadro para apresentação da avaliação de uma ontologia.

Ontologia <X>: <nome da ontologia> - Modelagem: <Y>			
Questões s/ critérios	Qtde	Observação	Pontuação
<critério 1>	<null; inteiro>	<texto para definir a pontuação>	{1,2,3}
<critério 2>	<null; inteiro>	<texto para definir a pontuação>	{1,2,3}
...
		Média Geral	<real>

Fonte: Autoria própria.

4.6.1 Recursos Utilizados

A realização dos estudos de caso compreende: o desenvolvimento dos padrões das classes essenciais relacionados com a representação de índices e indicadores; a definição das

classes e propriedades; a instanciação dos indivíduos nas ontologias e; o uso destas por uma aplicação. Para este trabalho foi utilizada a ferramenta Protégé¹⁰ versões 5.0 e 5.2, com auxílio de um de seus *reasoners*, no caso, o Hermit versão 1.3.8.413 e, uma aplicação foi criada pela API¹¹, em linguagem Java, fornecida pelo framework Apache Jena¹².

¹⁰Protégé: editor de ontologias sob licença aberta para uso na construção de sistemas inteligentes (MUSEN, 2015).

¹¹API: Application Programming Interface.

¹²Apache Jena: Framework Java aberto para construir aplicações Linked Data.

5 PADRÕES PROPOSTOS – MODELAGEM A

O desenvolvimento da primeira modelagem, dentre as duas, é apresentado a seguir de acordo com as etapas descritas no Capítulo 4 – Metodologia. Cada modelagem possui características diferentes resultantes das escolhas para solucionar problemas encontrados na descrição de índices, indicadores e medições. As modelagens se basearam na compreensão do autor em relação aos cenários propostos. A primeira abordagem se chamará modelagem “A” e seu início de desenvolvimento se dá com a identificação do cenário proposto.

5.1 IDENTIFICAÇÃO DO CENÁRIO COMO REQUERIMENTO PARA MODELAGEM

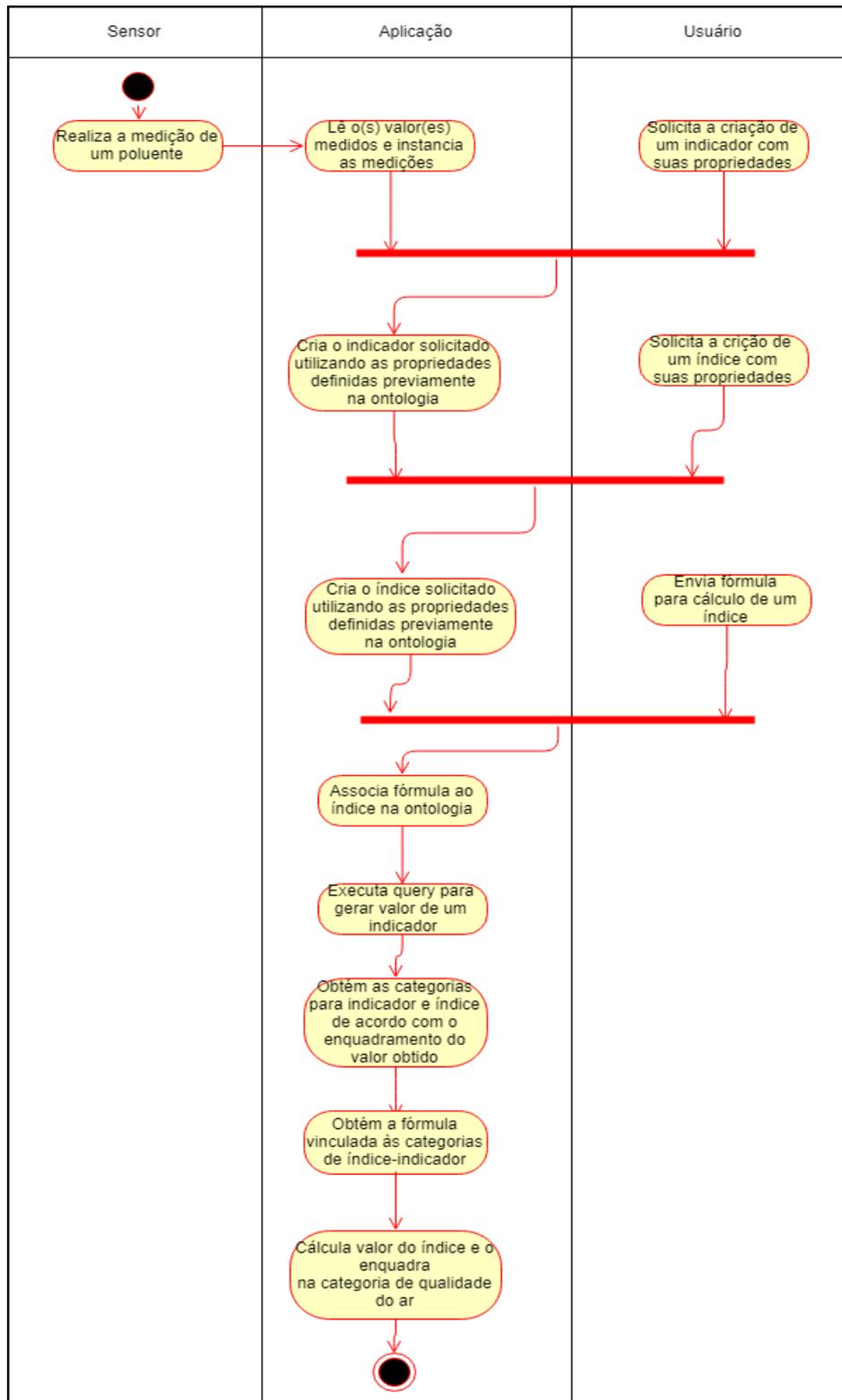
Em relação ao cenário, foram simulados alguns atores executando ações para relacionar-se com um sistema fictício de monitoramento do ar. Tal sistema permite aos usuários utilizarem índices e indicadores existentes (ex. IMECA) e utilizados por organizações que realizam monitoramento do ar, mas também deve permitir que usuários possam criar novos índices e indicadores. Nesta dissertação, o foco está na representação dos índices e indicadores e não no desempenho de um sistema de armazenamento de medições para posterior recuperação.

Um diagrama de atividades foi elaborado para demonstrar as ações dos atores envolvidos com este sistema de monitoramento (ver Figura 11). É importante ressaltar que o modelador deve desenvolver a ontologia fazendo uso dos padrões desenvolvidos para em seguida criar as classes com as suas propriedades e que estas atividades antecedem as interações dos atores com este sistema.

A partir da observação de vários índices como, por exemplo: AQI; IMECA; AQHI; DAQI e IAP, foram identificados conceitos e relações estruturais. Neste trabalho, pressupõe-se que novos índices e indicadores criados pelos usuários podem ser construídos a partir dos requerimentos existentes nos índices pesquisados que estão descritos a seguir:

- um índice é categorizado em pelo menos uma categoria;
- um índice é calculado ao menos por uma fórmula ou por uma forma de derivação;
- um indicador está associado a pelo menos um poluente;
- um indicador está associado ao menos à uma categoria;
- um indicador é calculado por uma forma de agregação sobre as medições;
- um poluente tem uma concentração mensurada por uma unidade de medida;
- um poluente tem sua concentração medida por um certo período de tempo.

Figura 11 – Diagrama de atividades do sistema fictício de monitoramento da poluição do ar.



5.1.1 Características do Cenário para a Modelagem A

Este cenário traz algumas peculiaridades que precisam ser consideradas durante a modelagem, descritas a seguir:

1. categorização de um índice nas categorias (ex. Bom, Regular, Ruim, Muito Ruim, Extremamente Ruim) em faixas de valores (ex. Bom [0, 50], Regular [51, 11], Ruim [101, 150], etc.);
2. categorização de um indicador para um poluente nas categorias (ex. monóxido de carbono (CO) nas categorias Bom [0.00, 5.50], Regular [5.51, 11.00], etc.);
3. uma ou mais fórmulas de cálculo para cada poluente e categoria. Por exemplo, o poluente monóxido de carbono (CO) na categoria Bom tem a fórmula $I[\text{CO}] = C[\text{CO}] * 50/5.6$, na categoria Extremamente Ruim tem a fórmula $I[\text{PM}_{10}] = C[\text{PM}_{10}] * 200/320$ e assim por diante;
4. a relação índice-indicador para os poluentes tem uma fórmula para todas as categorias e, em alguns casos, há uma fórmula para cada uma das suas categorias;
5. a unidade de medida adotada para os poluentes usualmente varia em função do poluente, mas também do país e do órgão regulador, neste cenário foi assumido que PM_{10} e $\text{PM}_{2.5}$ estão medidos em $\mu\text{g}/\text{m}^3$ e que os demais poluentes são medidos em ppm (partes por milhão);
6. a periodicidade de geração do indicador de um poluente é variável, neste cenário considerou-se que CO é obtido em intervalos de 8h, o poluente NO_2 é obtido a cada 1h e os demais poluentes, em períodos de 24h.

5.2 DETERMINAR AS QUESTÕES DE COMPETÊNCIA PARA AVALIAÇÃO DA MODELAGEM

Passa-se então, à etapa seguinte, onde se estabelecem as questões de competência para posterior avaliação da ontologia. As questões a seguir foram elaboradas pelo autor, considerando o cenário proposto e os componentes mais importantes que foram mencionados no passo anterior. Nas questões que contêm elementos variáveis utiliza-se a notação $\langle l \rangle$ para representa-los, tal que l é uma letra do alfabeto. Por exemplo, $\langle P \rangle$ denota que naquela posição deve haver um nome de um poluente. As questões identificadas são:

- (Q1) Quais são as medições do poluente $\langle P \rangle$ coletadas pela estação $\langle S \rangle$ de responsabilidade do agente $\langle A \rangle$ que ocorreram no período $\langle T \rangle$?
- (Q2) Quais indicadores fazem referência ao poluente $\langle P \rangle$?
- (Q3) Como o indicador $\langle R \rangle$ é calculado a partir das medições?
- (Q4) Como o indicador $\langle R \rangle$ é categorizado?
- (Q5) Qual a categorização para o indicador $\langle R \rangle$ na periodicidade $\langle H \rangle$?

- (Q6) Qual o valor para o indicador <R> no período <T> com a forma de derivação <D>?
- (Q7) Qual é a categoria para o valor <V> do indicador <R>?
- (Q8) Qual é a categoria para o valor <V> do índice <I>?
- (Q9) Quais são as categorias definidas para o índice <I>?
- (Q10) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI> e uma categoria do indicador <CR>?
- (Q11) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI>?
- (Q12) Qual a unidade de medida <U> utilizada na medição do poluente <P>?
- (Q13) Qual é a forma de derivação <D> do índice <I>?
- (Q14) Qual a periodicidade <H> e o poluente <P> calculados pelo indicador <R>?

5.3 ENUMERAR OS COMPONENTES DA MODELAGEM

O terceiro passo trata dos componentes que farão parte da ontologia a ser construída com os padrões desta modelagem. De acordo com a análise do cenário proposto e pelo conhecimento adquirido do referencial teórico extraíram-se conceitos como: medições, poluentes, indicadores e índices, além das relações entre eles. Também foi possível constatar a presença de faixas de valores que representam as categorizações para índices e indicadores, que também devem ser definidas.

Neste momento, listam-se os termos selecionados pelo uso do Quadro 9, proposto na Seção 4.3, para figurarem como classes nesta ontologia que será denominada de AirQualityOnto (AQ) que podem ser vistos no quadro 12.

Quadro 12 – Enumeração dos termos a serem modelados como classes.

Termo	Nome da Classe	Descrição	Superclasse
Poluente	Pollutant	Poluente em geral	-
Poluente do Ar	Air_pollutant	Poluente do ar	Pollutant
Antropogênico	Antropogenic	Tipo de poluente gerado por atividade humana	Air_Pollutant
Natural	Natural	Tipo de poluente presente na natureza sem a interferência humana	Air_Pollutant
Dióxido de Enxofre	SO2	Poluente do ar antropogênico	Antropogenic
Monóxido de Carbono	CO	Poluente do ar antropogênico	Antropogenic
Ozônio	O3	Poluente do ar antropogênico	Antropogenic
Dióxido de Nitrogênio	NO2	Poluente do ar antropogênico	Antropogenic
Partículas Suspensas menores que 10µm	PM10	Poluente do ar natural	Natural
Partículas Suspensas menores que 2.5µm	PM25	Poluente do ar natural	Natural

Estação de coleta	Station	Utilizada para coletar dados de poluentes.	-
Estação automática	Automatic	Estação que opera em modo automático.	Station
Estação manual	Manual	Estação que opera em modo manual.	Station
Medição	Measurement	Medição coletada por uma estação de coleta	-
Categorização	Categorizations	Categorizações para escala de valores	-
Categoria de Indicador	IndicatorCategorization	Categorização para indicadores	Categorizations
Categoria de índice	IndexCategorization	Categorização para índices	Categorizations
Periodicidade	Periodicity	Intervalo de tempo	-
Indicador	UserIndicator	Indicador de referência para o usuário, a ser obtido de um conjunto de medições	-
Índice	Index	Índice de referência a ser obtido de um indicador	-
Índice derivável	Index_with_DF	Índice de referência a ser obtido por meio de uma forma de derivação sobre indicadores	Index
Fórmula	Formula	Expressão matemática para gerar o valor resultante de um índice	-
Forma de derivação	DerivationForms	Forma de derivação ou função de agregação sobre conjuntos (dados)	-
Média	Average	Forma de derivação	DerivationForms
Máxima	Maximum	Forma de derivação	DerivationForms
Mínima	Minimum	Forma de derivação	DerivationForms

Fonte: Autoria própria.

Em seguida, identificam-se as relações entre as classes obtidas acima, em forma de propriedades. Seguindo o Quadro 10, proposto na Seção 4.3 para relacionar as propriedades com suas características, elaborou-se o Quadro 13 que contém as propriedades relativas ao cenário tratado nesta modelagem.

Quadro 13 – Enumeração dos termos a serem modelados como propriedades.

Termo	Nome da propriedade	Descrição	Domínio x Imagem*	Tipo
Momento da medição	hasHappenedOn	Momento no tempo que ocorre a coleta (data/hora)	Measurement, xsd:dateTime	Funcional
Limite inferior do intervalo	hasLowerLimit	Valor mínimo de um intervalo	[Categorizations], xsd:decimal	Funcional
Limite superior do intervalo	hasUpperLimit	Valor máximo de um intervalo	[Categorizations], xsd:decimal	Funcional
Medição	hasMeasurementOf		Measurement, Air_pollutant	Funcional
Periodicidade	hasPeriodicity		[UserIndicator], [Periodicity]	Funcional
Categorização	isCategorizedBy		[UserIndicator; Index], [Categorizations]	Não Funcional
Derivação ou agregação	isDerivedBy		[UserIndicator; Index], DerivationForms	Funcional

Indicador	isIndicatorFor		[UserIndicator], Air_pollutant	Funcional
Aparelho físico	takesPlaceAt		[Measurement], Station	Funcional
(*) Objetos entre [] são classes implícitas na definição				

Fonte: Autoria própria.

Até este passo, está identificado o cenário para modelagem, elaboradas as questões de competência e listados os termos que devem estar presentes na modelagem da ontologia.

5.4 CONSIDERAR O REUSO DE ELEMENTOS DE ONTOLOGIAS TOP-LEVEL NA MODELAGEM

Nesta etapa, são importados os elementos de ontologias *top-level* que possuem termos e propriedades em afinidade com os termos da ontologia proposta. Conforme indica a etapa, foram feitas consultas por meio dos buscadores, em bibliotecas de *sites* e de universidades que mantém pesquisas sobre o assunto (QUADRO 14), nas ontologias que detinham conteúdo em similaridade com o domínio aqui tratado.

Quadro 14 – Fontes de organizações e mecanismos para busca de ontologias candidatas a reuso

Proprietário	URL
DAML	http://www.daml.org/ontologies/
W3C	https://www.w3.org/wiki/Lists_of_ontologies
Gene Ontology Consortium	http://www.geneontology.org/
Envo	http://environmentontology.org/
Swoogle	http://swoogle.umbc.edu/
Open Ontology Repository	http://www.oor.net/
Protégé	https://protegewiki.stanford.edu/wiki/Protege_Ontology_Library
Enterprise Integration Lab University of Toronto	http://ontology.eil.utoronto.ca/

Fonte: Autoria própria.

Desta pesquisa, destacam-se as ontologias desenvolvidas no âmbito do projeto GCI da Universidade de Toronto (CA) além da conhecida ontologia *Time.owl* da W3C.

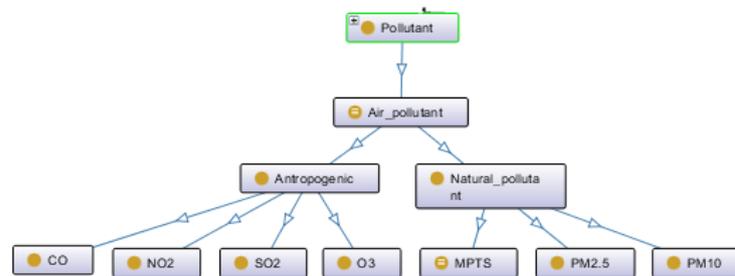
Em síntese, os passos desta busca, foram:

- identificação das classes relacionadas aos poluentes e de suas relações com outras classes;
- identificação das classes relacionadas com indicadores e de suas relações com outras classes ligadas com as medições e com o tempo;
- identificação das classes de sensores e suas relações;
- identificação das classes relacionadas com a proveniência dos dados.

5.4.1 Identificação das Classes e Relações Associadas aos Poluentes

A classe dos poluentes (*pollutant* em inglês) foi encontrada nas ontologias *Environment.owl* e *Pollution.owl* (Projeto GCI) com as denominações: **Pollution** e **Air_pollution**. Neste modelo, optou-se pelo uso dos nomes **Pollutant** e **Air_pollutant** embora tenham o mesmo propósito, isto é, o de representar os poluentes existentes (ex. O₃, CO e NO₂). Em seguida, dividiu-se a classe dos poluentes do ar em outras duas especializações, a classe **Antropogenic** e a classe **Natural_pollutant** como subclasses de **Air_pollutant** ilustradas na Figura 12.

Figura 12 – Hierarquia das classes de poluentes do ar.

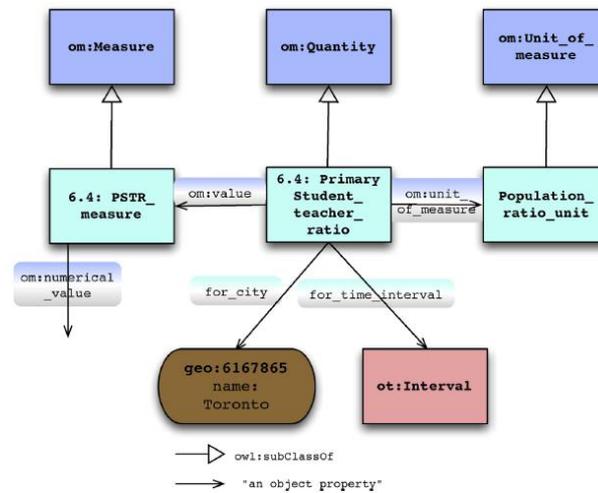


Fonte: O autor (2018)

5.4.2 Identificação das Classes e Relações Associadas aos Indicadores

Outro conceito principal desta modelagem refere-se ao indicador. Para a criação do padrão para o conceito indicador nesta modelagem, foi definida uma classe para representá-lo com o nome de **UserIndicator**. Pode se ver pela Figura 13, uma classe do tipo indicador fazendo uso de outras classes e propriedades das ontologias *om-1.6.owl* (OM) que trata de medições e *time.owl* (OT) que trata dos aspectos temporais. Este padrão foi adotado por (FOX, 2015) em seu projeto para uma ontologia sobre indicadores educacionais.

Figura 13 – Indicador PSTR com classes das ontologias Units of Measure (OM) e Time (OT)



Fonte: Mark S. Fox (2015)

Em sua definição, pode-se ver a necessidade de se ter uma relação temporal associada com o indicador identificada aqui pela propriedade **for_time_interval** e a classe **Interval** (OT). Além disso, também é preciso existir uma relação de valor com uma classe que represente as medições coletadas ou carregadas na ontologia.

A classe de medições **om:measure** foi reutilizada da ontologia *om-1.6.owl* (OM) da plataforma *wurvoc.org* entretanto, é importante esclarecer que mais recentemente seus idealizadores lançaram a versão *Om-2.0.rdf* disponível na plataforma **GitHub**¹³, com consideráveis alterações nas propriedades que definem a formação de uma grandeza, de modo que por motivos de compatibilidade com o que já havia sido desenvolvido para a ontologia AQ, optou-se por se utilizar as definições da versão anterior.

De volta ao contexto proposto nesta modelagem, deve-se ter em conta que o indicador de poluição do ar terá um valor a ser derivado de medições para um determinado período relativo a um poluente que está sendo monitorado. Indicadores nesta abordagem são valores agregados relacionados a um poluente para um intervalo de tempo. Explorando este aspecto temporal, buscou-se uma classe para representar a periodicidade para geração dos indicadores, no caso a que mais se encaixa com a necessidade da modelagem, é a classe **DurationDescription** da ontologia *Time.owl* (OT).

As representações em sintaxe Manchester e em um diagrama UML presentes nas Figuras 14 e 15 respectivamente, demonstram como está definida esta classe, cuja função é

¹³GitHub: plataforma de hospedagem para códigos fonte privados ou abertos.

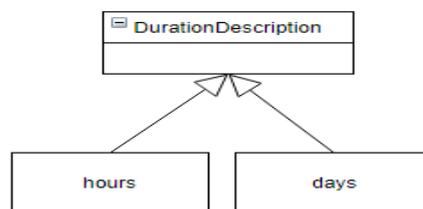
representar um período de tempo a ser associado a um evento ou a um conceito que seja dependente de uma referência temporal, como são os indicadores de poluição ambiental.

Figura 14 – Classe DurationDescription em sintaxe Manchester.

```
Class: DurationDescription
SubClassOf:
  ot:days max 1 rdfs:Literal
  ot:hours max 1 rdfs:Literal
```

Fonte: O autor (2018)

Figura 15 – Diagrama UML da Classe DurationDescription importada de Time.owl.



Fonte: O autor (2018)

Nota-se que suas propriedades foram customizadas para trabalhar com dias e horas que são os períodos utilizados em todos os indicadores pesquisados. Entretanto, sua definição é expansível para outras unidades de tempo como, minutos, segundos, etc.

5.4.3 Identificação das Classes e Relações Associadas aos Sensores

Como proposto pelo cenário de monitoramento, uma coleta ocorre por meio de um dispositivo físico que faz as leituras ou medições das concentrações emitidas por poluentes do ar. Como estes dispositivos também chamados de estações, utilizam sensores, utilizou-se o conceito definido na ontologia *ssn.owl* (SE) da W3C.

Optou-se então por se definir uma classe interna de nome **Station** equivalente à classe **se:Sensor** (vide Figura 16). Estas classes representam o dispositivo de coleta, ou seja, o componente físico que faz a leitura do dado que será posteriormente convertido em uma medição a qual servirá de insumo para um indicador.

Figura 16 – Classe Station em sintaxe Manchester

```

Class: aq:Station

EquivalentTo:
  <http://purl.oclc.org/NET/ssnx/ssn#Sensor>,
  aq:hasOperationMode exactly 1 aq:Operation_Mode

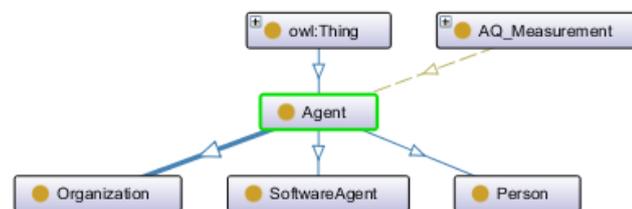
```

Fonte: O autor (2018)

5.4.4 Identificação das Classes e Relações Associadas à Proveniência dos Dados

Quando se fala em propriedade dos dados, deseja-se saber a proveniência dos mesmos, isto é, quem é o agente, detentor e provedor das informações coletadas. Neste caso, a classe a ser reutilizada é **Agent** (FIGURA 17), definida na ontologia *Prov-o.owl* (PR) da W3C. A função desta ontologia é conceitualizar aspectos de proveniência de dados. Com o seu uso agrega-se informação sobre a autoria, isto é, quem foi o agente gerador do dado, aumentando assim a confiabilidade da informação presente na base de conhecimento do modelo.

Figura 17 – Classe Agent importada de Prov-o.owl.



Fonte: O autor (2018)

De acordo com a definição dada pela W3C, são subclasses de **Agent**:

- **Software Agent** quando o agente é um software.
- **Person** quando um agente é uma pessoa (ou várias).
- **Organization** quando o agente é uma instituição legal como uma companhia, sociedade, associação ou qualquer outro tipo de organização.

Os reusos dos elementos (em cor cinza) causaram modificações nas classes e propriedades inicialmente identificadas, conforme mostram respectivamente, os Quadros 15 e 16.

Quadro 15 – Enumeração dos termos a serem modelados como classes.

Termo	Nome da Classe	Descrição	Superclasse
Poluente	PO:Pollution	Poluente em geral	-
Poluente	Pollutant	Poluente em geral	-
Poluente do Ar	PO:Air_pollution	Poluente do ar	PO:Pollution

Poluente do Ar	Air_pollutant	Poluente do ar	Pollutant
Antropogênico	Antropogenic	Tipo de poluente gerado por atividade humana	Air_Pollutant
Natural	Natural	Tipo de poluente presente na natureza sem interferência humana	Air_Pollutant
Dióxido de Enxofre	SO2	Poluente do ar antropogênico	Antropogenic
Monóxido de Carbono	CO	Poluente do ar antropogênico	Antropogenic
Ozônio	O3	Poluente do ar antropogênico	Antropogenic
Dióxido de Nitrogênio	NO2	Poluente do ar antropogênico	Antropogenic
Partículas Suspensas menores que 10µm	PM10	Poluente do ar natural	Natural
Partículas Suspensas menores que 2.5µm	PM25	Poluente do ar natural	Natural
Sensor	SE:Sensor	Sensor ou dispositivo para coletar dados	DUL:PhysicalObject
Estação de coleta	Station	Utilizada para coletar dados	SE:Sensor
Estação automática	Automatic	Estação que opera em modo automático.	Station
Estação manual	Manual	Estação que opera em modo manual	Station
Medição	Measurement	Medição coletada por uma estação de coleta	OM:Measure
Unidade de medida	OM:Unit_of_measure	Grandeza que caracteriza um valor medido	-
Categorização	Categorizations	Categorizações para uma escala de valores	-
Categoria de Indicador	IndicatorCategorization	Categorização para indicadores	Categorizations
Categoria de índice	IndexCategorization	Categorização para índices	Categorizations
Duração	OT:DurationDescription	Duração de um evento	-
Periodicidade	Periodicity	Intervalo de tempo	OT:DurationDescription
Hora	OT:Hour	Duração em quantidade de horas	OT:DurationDescription
Dia	OT:Days	Duração em quantidade de dias	OT:DurationDescription
Indicador	UserIndicator	Indicador de referência para o usuário, a ser obtido de um conjunto de medições	-
Índice	Index	Índice de referência a ser obtido de um indicador	-
Índice derivável	Index_with_DF	Índice de referência a ser obtido por meio de uma forma de derivação sobre indicadores	Index
Fórmula	Formula	Expressão matemática para gerar o valor resultante de um índice	-
Forma de derivação	DerivationForms	Forma de derivação ou função de agregação sobre um conjunto de dados	-
Média	Average	Forma de derivação	DerivationForms
Máxima	Maximum	Forma de derivação	DerivationForms
Mínima	Minimum	Forma de derivação	DerivationForms
Agente	PR:Agent	Agente responsável pela coleta e publicação da medição	-

Fonte: Autoria própria.

Em seguida, pode ser vista uma relação das propriedades no Quadro 16, no qual as propriedades oriundas do reuso estão na cor cinza.

Quadro 16 – Enumeração dos termos a serem modelados como propriedades.

Termo	Nome da propriedade	Descrição	Domínio x Imagem*	Tipo
Momento da medição	hasHappenedOn	Momento no tempo que ocorre a coleta data/hora	Measurement, xsd:dateTime	Funcional
Limite inferior do intervalo	hasLowerLimit	Valor mínimo de um intervalo	[Categorizations], xsd:decimal	Funcional
Limite superior do intervalo	hasUpperLimit	Valor máximo de um intervalo	[Categorizations], xsd:decimal	Funcional
Medição	hasMeasurementOf	Tem a medição de um poluente	Measurement, Air_pollutant	Funcional
Valor da medição	OM:numerical_value	Valor real de uma medição	Measure, xsd:decimal	Funcional
Unidade de medida	OM:unit_of_measure_or_measurement_scale	Unidade de medida ou escala de medida	measure, Unit_of_measure_or_measurement_scale	Funcional
Periodicidade	hasPeriodicity	Tem a periodicidade	[UserIndicator], [Periodicity]	Funcional
Duração ou periodicidade	OT:days	Dias	ot:DurationDescription, xsd:decimal	Funcional
Duração ou periodicidade	OT:hours	Horas	ot:DurationDescription, xsd:decimal	Funcional
Categorização	isCategorizedBy	É categorizado por	[UserIndicator; Index], [Categorizations]	Não Funcional
Derivação ou agregação	isDerivedBy	É derivado de	[UserIndicator; Index_with_DF], DerivationForms	Funcional
Indicador	isIndicatorFor	É indicador para o poluente	[UserIndicator], Air_pollutant	Funcional
Aparelho responsável pela medição	takesPlaceAt	Ocorre em	[Measurement], Station	Funcional
(*) Objetos entre [] são classes implícitas na definição				

Fonte: Autoria própria.

Recorda-se a seguir, os prefixos de ontologias utilizados nas classes ou propriedades importadas: PO-Pollution.owl, SE-Sensor.owl, OM-OM1.6.owl, OT:Time.owl, PR:Provo.owl e DUL:DUL.owl.

5.5 ESPECIFICAÇÃO DOS PADRÕES DE MODELAGEM

Após a identificação dos componentes que precisam ser modelados e dos conceitos e relações que podem ser importadas das ontologias de alto nível, partiu-se para o desenvolvimento dos padrões. Considerando a ordem com que os principais elementos do domínio escolhido ocorrem em um processo de monitoramento, isto é: 1) medições de

poluentes são coletadas; 2) indicadores são derivados periodicamente e 3) índices são calculados, propõe-se iniciar a definição dos padrões pela classe das medições.

O padrão para esta classe será chamado de **Measurement** e suas instâncias vão representar as medições resultantes das leituras feitas pelos sensores das estações de coleta. Mais precisamente, uma instância de **Measurement** representa uma única medição de um poluente realizada por uma estação em um dado momento sendo responsabilidade de um agente. Uma mesma medição pode ser dada em várias unidades de medida, sendo então representada por mais de um par, composto por um valor numérico e uma unidade de medida. Assim, o padrão de modelagem desta classe seguindo o formato proposto, ficaria definido da seguinte maneira.

Nome do padrão: **Measurement**

Questões de competência:

(Q1) Quais são as medições do poluente <P> coletados pela estação <S> de responsabilidade do agente <A> que ocorreram no período <T>?

(Q12) Qual a unidade de medida <U> utilizada na medição do poluente <P>?

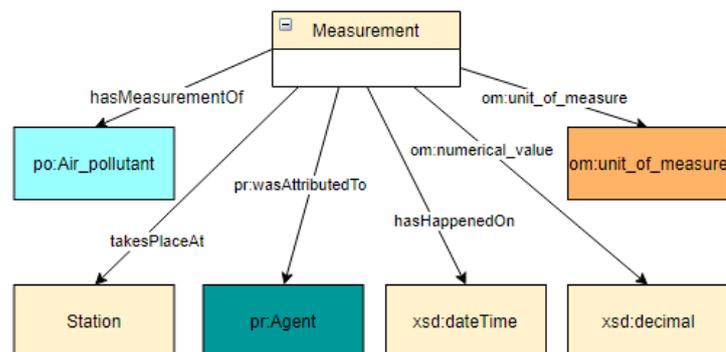
Elementos:

```

Class: Measurement

SubClassOf:
    aq:hasMeasurementOf exactly 1 po:Air_pollutant
    aq:takesPlaceAt exactly 1 aq:Station
    pr:wasAttributedTo exactly 1 pr:Agent
    aq:hasHappenedOn exactly 1 xsd:dateTime
    om:numerical_value some xsd:decimal
    om:unit_of_measure some om:Unit_of_measure
  
```

Diagrama:



A próxima classe a ter um padrão de modelagem é a dos indicadores, que será tratada pelo termo **UserIndicator**. A classe **UserIndicator** deve permitir descrever o valor de concentração de um poluente cujo valor é gerado a partir de várias medições por meio de uma forma de cálculo dentro de uma periodicidade escolhida sendo categorizado por uma ou mais

categorias de emissão ou de concentração e, que o valor atribuído a cada indicador pode ser enquadrado em categorias particionadas. Considerando estas necessidades, o padrão classe no formato proposto fica assim descrito. Nome do padrão: **UserIndicator**

Questões de competência:

(Q2) Quais indicadores fazem referência ao poluente <P>?

(Q3) Como o indicador <R> é calculado a partir das medições?

(Q4) Como o indicador <R> é categorizado?

(Q5) Qual a categorização para o indicador <R> no período <T>?

(Q14) Qual a periodicidade <H> e o poluente <P> calculados pelo indicador <R>?

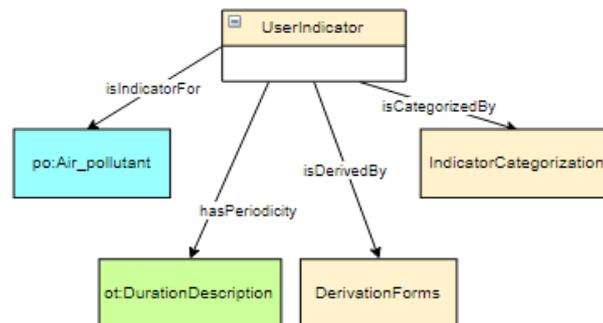
Elementos:

```

Class: UserIndicator

SubClassOf:
  aq:isIndicatorFor exactly 1 po:Air_pollutant
  aq:hasPeriodicity exactly 1 aq:DurationDescription
  aq:isDerivedBy exactly 1 aq:DerivationForms
  aq:isCategorizedBy some aq:AQ_Categorizations
  
```

Diagrama:



Como explicado anteriormente, as medições uma vez instanciadas devem ser processadas para que outras informações sejam derivadas e carregadas na classe de indicadores. Isto pode ocorrer pelo uso de consultas em SPARQL sobre triplas RDF/OWL, que uma vez executadas pela aplicação, irão viabilizar o carregamento das informações no indicador que for construído. Para fazer uso deste recurso, propõe-se um padrão de consulta em SPARQL para a classe **UserIndicator** que ficou assim definido.

Nome do padrão: **SPARQL_for_UserIndicator**

Questão de competência:

(Q6) Qual o valor para o indicador <R> no período <T> com a forma de derivação <D>?

Elementos:

```

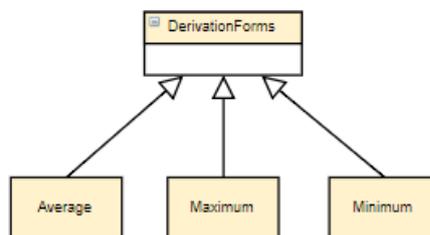
SELECT ?ui (?<df> (?valor) AS ?'<df>') ?df ?pol
WHERE {
  ?ui aq:isIndicatorFor ?pol .
  ?ui aq:isDerivedBy ?df .
  ?m aq:hasMeasurementOf ?pol .
  ?m aq:hasHappenedOn ?datahora .
  ?m om:numerical_value ?valor .

FILTER ( ?ui = <userIndicator> &&
  ?datahora >= 'YYYY-MM-DDThh:mm:ss'^^xsd:dateTime &&
  ?datahora <= 'YYYY-MM-DDThh:mm:ss'^^xsd:dateTime )
}
GROUP BY ?ui ?df ?pol

```

No padrão **UserIndicator**, observa-se a introdução de duas outras classes necessárias para a representação dos indicadores, que são: **DerivationForms** e **IndicatorCategorization**. A classe **DerivationForms** não é um padrão de modelagem, mas é necessária ao padrão **UserIndicator**. Ela tem como objetivo indicar a forma de derivação da concentração do poluente a partir das medições coletadas. Então, quando um indicador for criado por um usuário da ontologia, este deve definir sua forma de derivação. As instâncias desta classe são operações sobre conjuntos de dados, como a média, o valor máximo ou o valor mínimo dos valores medidos. Uma maneira de representar esta classe pode ser vista na Figura 18 a seguir.

Figura 18 – Classe DerivationForms com as suas subclasses.



Fonte: O autor (2018)

Entretanto, esta classe poderia ser definida com um dos recursos apresentados na subseção 3.3, isto é, o construto do tipo classes enumeradas *OneOf*. Neste caso a definição da classe se dá pela equivalência de seus membros. Assim, a classe em sintaxe Manchester ficaria definida como na Figura 19.

Figura 19 – Declaração em sintaxe Manchester utilizando enumeração de classes

```

Class: DerivationForms

EquivalentTo:
  {aq:Average , aq:Maximum , aq:Minimum}

```

Fonte: O autor (2018)

A classe **IndicatorCategorization** é uma especialização da classe **Categorizations** cuja conceitualização responde uma das questões de competência para construção deste modelo. Logo, **Categorizations** servirá como um padrão de modelagem.

Como foi visto nos índices estudados na referência teórica, índices e indicadores são normalmente categorizados em faixas para permitir um fácil entendimento por seus usuários. As faixas são descritas em categorias como, por exemplo: crítico, ruim, moderado, regular, bom. No caso, uma categorização é definida pela união disjunta de suas subclasses, isto é, não é possível que um indicador ou índice pertença a duas ou mais categorias de um mesmo conjunto de categorias.

Cada $\langle \text{categoria} \rangle_1$ é uma subclasse de $\langle \text{categorização} \rangle$ e possui uma faixa de valores dada pelos seus limites inferior e superior. Por exemplo, um indicador do poluente Ozônio (O_3) poderia ter as faixas: **good**[0,62.5], **regular**[62.6,125], etc. Neste caso, se descrevem as classes de $\langle \text{categorização} \rangle$ como: $O_3_categorization$ e suas categorias como: $\langle \text{categoria} \rangle_1 = \mathbf{good}$, $\langle \text{categoria} \rangle_2 = \mathbf{regular}$, e assim para as demais categorias. Uma instância de cada categoria deve ser criada para serem concretamente definidos os limites inferior e superior de cada faixa que fará parte de uma categorização. O padrão **Categorizations** de acordo com o formato proposto para modelagem, ficou assim definido.

Nota. Estas propriedades serão herdadas pelas subclasses: **IndexCategorization** e **IndicatorCategorization**.

Nome do padrão: **Categorizations**

Questões de competência:

(Q7) Qual é a categoria para o valor $\langle V \rangle$ do indicador $\langle R \rangle$?

(Q8) Qual é a categoria para o valor $\langle V \rangle$ do índice $\langle I \rangle$?

Elementos:

```

Class: Categorizations
SubClassOf:
    hasLowerLimit exactly 1 xsd:decimal,
    hasUpperLimit exactly 1 xsd:decimal

DisjointUnionOf:
    IndexCategorization, IndicatorCategorization

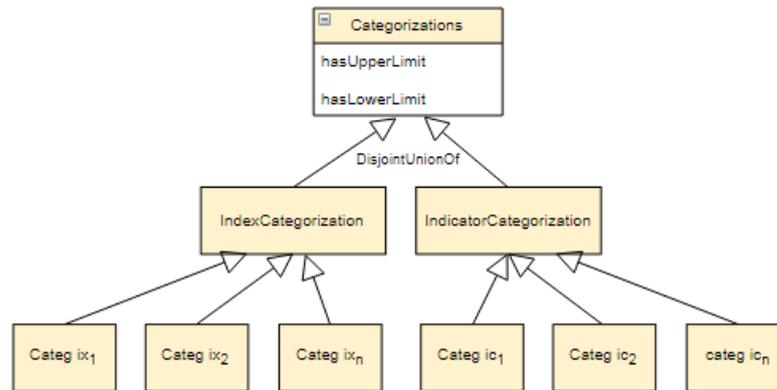
Class: <categorização>
SubClassOf:
    IndexCategorization ..or.. IndicatorCategorization

DisjointUnionOf:
    <categoria>1, <categoria>2 .. <categoria>n

Class: <categoria>i
SubClassOf:
    <categorização>

```

Diagrama:



Ressalta-se que as subclasses de categorias com as faixas de valores para índices e indicadores precisam também ser declaradas. Tome-se como exemplo, o índice do México de qualidade do ar (**IMECA**) e as faixas para categorizar seus indicadores relacionados ao poluente de partículas suspensas menores que 10 micrômetros (PM₁₀).

Para este indicador, há cinco categorias: **good**, **regular**, **poor**, **bad** e **critical**. Estas categorias devem então ser subclasses disjuntas da classe **IndicatorCategorization**. É preciso instanciar para cada faixa, os valores mínimo e máximo, a fim de concretizar a definição destas categorias. Assim, o valor derivado pela utilização da classe **UserIndicator** pode ser enquadrado em uma das faixas declaradas nas categorias de **IndicatorCategorization**.

O próximo padrão modelado trata da classe principal de ontologias para realizar o monitoramento ambiental, ou seja, o índice de qualidade do ar. O padrão para esta classe é referenciado por **Index**. Este padrão permite descrever um índice que é definido por uma ou mais categorizações. Um subtipo, denominado **Index_with_DF**, permite adicionalmente acrescentar uma forma de derivação para calcular um único valor de índice a partir dos valores dos seus indicadores. Sua definição ficou conforme segue.

Nome do padrão: **Index**

Questões de competência:

(Q9) Quais são as categorias definidas para o índice <I>?

(Q13) Qual é a forma de derivação <D> do índice <I>?

Elementos:

```

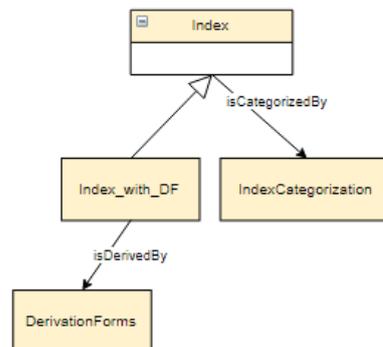
Class: Index
SubClassOf:
  aq:isCategorizedBy some aq:IndexCategorization

Class: aq:AQ_Index_with_DF
EquivalentTo:
  aq:isDerivedBy exactly 1 aq:DerivationForms

SubClassOf:
  aq:Index

```

Diagrama:



Seguindo nesta abordagem, ainda é preciso tratar as questões sobre a relação índice-indicador, e como representar uma fórmula ou função de cálculo. A questão dez foi elaborada para estimular a criação de um padrão de modelagem que faça este tratamento. A proposta que surgiu foi de se utilizar uma classe de associação entre índices e indicadores que devem referenciar suas categorizações com a adição de uma propriedade para indicar a forma de cálculo vinculada a esta relação.

A declaração da fórmula de cálculo do índice é outro passo importante a ser coberto nesta alternativa de modelagem. A opção escolhida para expressar a fórmula, de forma indicativa, foi por meio de um tipo de dado da linguagem OWL DL, o tipo **string**. Neste caso a interpretação da fórmula deve ser feita na implementação da aplicação que fará uso da ontologia a ser criada com o uso dos padrões propostos.

O padrão de modelagem desenvolvido para associar a fórmula de cálculo do índice em conjunto com a relação entre um índice e seu indicador é uma classe de ligação entre estes conceitos. Este padrão se baseou no requerimento existente no índice IMECA. Neste índice, existem de uma a muitas fórmulas dependentes de cada poluente. Como o valor atribuído ao indicador se refere a uma de suas faixas de concentração, a fórmula que está associada a faixa enquadrada deverá ser associada também ao seu índice. Logo, uma classe que relaciona o índice final ao indicador que o suporta atende esta necessidade.

Este padrão de modelagem foi chamado de **IndicatorToIndexFormula** e sua definição é dada pelas seguintes propriedades:

- **makesRefTo**: propriedade da relação índice-indicador que referencia uma das categorias do índice **IndexCategorization**;
- **makesRefTo**: propriedade da relação índice-indicador que referencia uma das categorias do indicador **IndicatorCategorization**;
- **hasFormula**: propriedade da relação índice-indicador que indica uma fórmula de cálculo para o índice, a qual será representada na forma indicativa pelo tipo de dado **string**. Este padrão ficou assim definido.

Nome do padrão: **IndicatorToIndexFormula**

Questão de competência:

(Q10) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI> e para uma categoria do indicador <CR>?

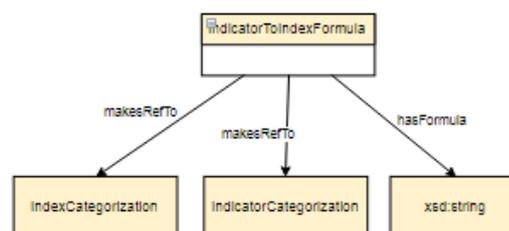
Elementos:

```

Class: IndicatorToIndexFormula

SubClassOf:
  aq:makesRefTo exactly 1 aq:IndexCategorization
  aq:makesRefTo exactly 1 aq:IndicatorCategorization
  aq:hasFormula exactly 1 xsd:string
  
```

Diagrama:



6 PADRÕES PROPOSTOS – MODELAGEM B

A segunda abordagem será denominada de modelagem “B” e seu desenvolvimento também se dá primeiramente com a identificação do cenário proposto. Na modelagem “A”, os padrões para as classes de índices e indicadores e suas categorizações tiveram uma abordagem adequada ao cenário que foi proposto. Nesta outra modelagem, modificações no cenário utilizado podem eventualmente, implicar em alterações nas questões de competência e em seus padrões.

6.1 IDENTIFICAÇÃO DO CENÁRIO COMO REQUERIMENTO PARA MODELAGEM

Algumas modificações nos requerimentos foram consideradas no cenário para o desenvolvimento dos padrões da modelagem “B”. Modificaram-se os requisitos em relação: à categorização de indicadores, às fórmulas de cálculo do índice, aos poluentes que são monitorados, às unidades de medida e aos termos que descrevem as categorias do índice. Estas alterações se basearam na especificação do índice do Canadá (AQHI).

Desta forma, estas mudanças vão implicar em algumas alterações nos padrões desta modelagem que serão explicadas adiante. No tocante ao sistema fictício do qual fazem parte a ontologia, a aplicação e os atores, não há mudanças e, a dinâmica entre eles pode ser vista na Figura 11, apresentada na Seção 5.1 deste documento.

6.1.1 Características do Cenário para a Modelagem B

A mudança de cenário implicará em algumas modificações nesta modelagem “B” que não foram tratadas anteriormente, tais como:

- na classe **UserIndicator**, deverá existir um tratamento para a ausência da necessidade de se categorizar indicadores;
- a classe **IndicatorToIndexFormula**, precisa tratar a ausência de categorizações para indicadores e a existência de uma única fórmula de cálculo para o índice;
- na classe **IndexCategorization**, para enumerar termos condizentes com o grau de risco de poluição;
- a classe de poluentes poderá, opcionalmente, omitir das especializações, os poluentes do ar que não fazem parte da fórmula do índice proposto neste cenário.

De acordo com o exposto, as características específicas deste cenário que devem ser consideradas estão descritas a seguir:

1. A categorização do índice deve considerar níveis relativos ao grau de risco de poluição, invertendo portanto, a interpretação dos usuários em relação à informação recebida (ex. Baixo, Moderado, Alto e Muito Alto) em faixas com intervalos de valores, como: Baixo [1, 10], Moderado [11, 20], .. Muito Alto [70, +∞];
2. A fórmula é única, válida para a concentração de qualquer um dos poluentes $\langle P_1 \rangle$, $\langle P_2 \rangle$ e $\langle P_3 \rangle$ e é descrita pela expressão: $I = \{(1000/10.4) * \{(e^{**}(0.000537 * \langle P_1 \rangle) - 1) + (e^{**}(0.000871 * \langle P_2 \rangle) - 1) + (e^{**}(0.000487 * \langle P_3 \rangle) - 1)\}\}$;
3. As unidades de medida são diferentes, por ex.: $\mu\text{g}/\text{m}^3$ para os poluentes $\langle P_1 \rangle$ e $\langle P_2 \rangle$ e ppm para o poluente $\langle P_3 \rangle$;
4. A periodicidade de geração dos indicadores está padronizada em $\langle H \rangle$ horas.

6.2 DETERMINAR AS QUESTÕES DE COMPETÊNCIA PARA AVALIAÇÃO DA MODELAGEM

Depois de avaliado o novo cenário proposto, segue-se com a elaboração das questões de competência, as quais servem como parâmetros para desenvolvimento dos padrões de modelagem, mas também para a avaliação da ontologia após sua construção. Assim como proposto nas questões da modelagem anterior, serão utilizados as mesmas variáveis do cenário anterior. As questões descritas anteriormente foram mantidas e estão relacionadas a seguir:

- (Q1) Quais são as medições do poluente $\langle P \rangle$ coletados pela estação $\langle S \rangle$ de responsabilidade do agente $\langle A \rangle$ que ocorreram no período $\langle T \rangle$?
- (Q2) Quais indicadores fazem referência ao poluente $\langle P \rangle$?
- (Q3) Como o indicador $\langle R \rangle$ é calculado a partir das medições?
- (Q4) Como o indicador $\langle R \rangle$ é categorizado?
- (Q5) Qual a categorização para o indicador $\langle R \rangle$ na periodicidade $\langle H \rangle$?
- (Q6) Qual o valor para o indicador $\langle R \rangle$ no período $\langle T \rangle$ com a forma de derivação $\langle D \rangle$?
- (Q7) Qual é a categoria para o valor $\langle V \rangle$ do indicador $\langle R \rangle$?
- (Q8) Qual é a categoria para o valor $\langle V \rangle$ do índice $\langle I \rangle$?
- (Q9) Quais são as categorias definidas para o índice $\langle I \rangle$?
- (Q10) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice $\langle CI \rangle$ e uma categoria do indicador $\langle CR \rangle$?

(Q11) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI>?

(Q12) Qual a unidade de medida <U> utilizada na medição do poluente <P>?

(Q13) Qual é a forma de derivação <D> do índice <I>?

(Q14) Qual a periodicidade <H> e o poluente <P> calculados pelo indicador <R>?

Para efeito didático, a ontologia resultante desta modelagem será tratada pelo nome de AirQualityOnto_B, levará o acrônimo (BQ) e suas classes e propriedades serão prefixadas por “bq:” nas descrições quando necessário.

6.3 ENUMERAR OS COMPONENTES DA MODELAGEM

Nesta etapa, listam-se os termos em um quadro como explicado na Subseção 4.3, os quais devem estar presentes na ontologia. Na relação apresentada no Quadro 17 a seguir, podem ser vistos os termos que se tornarão classes-subclasses de uma hierarquia presentes da ontologia a ser construída com os padrões da modelagem “B”. Lembrando que não estão presentes categorias de indicadores por não serem necessários neste contexto.

Quadro 17 – Enumeração dos termos a serem modelados como classes.

Termo	Nome da Classe	Descrição	Superclasse
Poluente	Pollutant	Poluente em geral	-
Poluente do Ar	Air_pollutant	Poluente do ar	Pollutant
Antropogênico	Antropogenic	Tipo de poluente gerado por atividade humana	Air_Pollutant
Natural	Natural	Tipo de poluente presente na natureza sem a interferência humana	Air_Pollutant
Ozônio	O3	Poluente do ar antropogênico	Antropogenic
Dióxido de Nitrogênio	NO2	Poluente do ar antropogênico	Antropogenic
Dióxido de Enxofre	SO2	Poluente do ar antropogênico	Antropogenic
Partículas Suspensas menores que 2.5µm	PM25	Poluente do ar natural	Natural
Partículas Suspensas menores que 10µm	PM10	Poluente do ar natural	Natural
Estação de coleta	Station	Utilizada para coletar dados de poluentes.	-
Estação automática	Automatic	Estação opera em modo automático.	Station
Estação manual	Manual	Estação que opera em modo manual.	Station
Medição	Measurement	Medição coletada por uma estação de coleta	-
Categorização	Categorizations	Categorizações para uma escala de valores	-
Categoria de índice	IndexCategorization	Categorização para índices	Categorizations
Periodicidade	Periodicity	Intervalo de tempo	-
Indicador	UserIndicator	Indicador de referência ao usuário, a ser obtido de um conjunto de medições	-
Indicador com categorização	Categorized_UI	Indicador de referência ao usuário que possui categorias para enquadramento	UserIndicator

Indicador sem categorização	NonCategorized_UI	Indicador de referência para o usuário sem possuir categorias para enquadramento	UserIndicator
Índice	Index	Índice de referência a ser obtido de um indicador	-
Índice derivável	Index_with_DF	Índice de referência a ser obtido por meio de uma forma de derivação sobre indicadores	Index
Fórmula	Formula	Expressão matemática para gerar o valor resultante de um índice	-
Forma de derivação	DerivationForms	Forma de derivação ou função de agregação sobre um conjunto de dados	-
Média	Average	Forma de derivação	DerivationForms
Máxima	Maximum	Forma de derivação	DerivationForms
Mínima	Minimum	Forma de derivação	DerivationForms

Fonte: Autoria própria.

Em seguida, identificam-se as relações entre as classes descritas acima, em forma de propriedades. No Quadro 18, podem ser vistas estas propriedades e as suas características, fazendo uso do quadro proposto na Seção 4.3.

Quadro 18 – Enumeração dos termos a serem modelados como propriedades

Termo	Nome da propriedade	Descrição	Domínio x Imagem*	Tipo
Momento da medição	hasHappenedOn	Momento no tempo que ocorre a coleta (data/hora)	Measurement, xsd:dateTime	Funcional
Limite inferior do intervalo	hasLowerLimit	Valor mínimo de um intervalo	[Categorizations], xsd:decimal	Funcional
Limite superior do intervalo	hasUpperLimit	Valor máximo de um intervalo	[Categorizations], xsd:decimal	Funcional
Medição	hasMeasurementOf		Measurement, Air_pollutant	Funcional
Periodicidade	hasPeriodicity		[UserIndicator], [Periodicity]	Funcional
Categorização	isCategorizedBy		[Index], [Categorizations]	Não Funcional
Derivação ou agregação	isDerivedBy		[UserIndicator; Index], DerivationForms	Funcional
Indicador	isIndicatorFor		[UserIndicator], Air_pollutant	Funcional
Aparelho físico	takesPlaceAt		[Measurement], Station	Funcional

(*) Objetos entre [] são classes implícitas na definição

Fonte: Autoria própria.

Da mesma forma que na modelagem anterior, neste momento está identificado o cenário para modelagem, estão elaboradas as questões de competência e listados os termos que devem estar presentes na modelagem da ontologia.

6.4 CONSIDERAR O REUSO DE ELEMENTOS DE ONTOLOGIAS *TOP-LEVEL* NA MODELAGEM

Da mesma forma que na modelagem anterior, se busca aplicar o reuso de conceitos e propriedades bem fundamentados de ontologias de alto nível. Dependendo das classes a serem declaradas, são reutilizados componentes de outras ontologias que são avaliadas depois de se efetuar uma busca nos portais previamente informados na Seção 5.4.

Seguiu-se então, com os mesmos passos apresentados em 5.4, isto é: identificação de classes e propriedades relacionadas aos poluentes; aos indicadores e seus aspectos referentes à medição, tempo e proveniência dos dados; e por fim identificação de classes relativas aos sensores.

6.4.1 Identificação das Classes e Relações Associadas aos Poluentes

A identificação relativa aos poluentes se deu exatamente como foi executada na etapa equivalente para a modelagem “A”, com o uso de elementos da ontologia *Pollution.owl* (PO), porém a hierarquia dos poluentes ficou restrita somente aos três poluentes envolvidos neste cenário.

6.4.2 Identificação das Classes e Relações Associadas aos Indicadores

Sobre a identificação das classes e relações associadas aos indicadores, foram utilizados os mesmos conceitos da modelagem “A”, isto é, a classe **om:measure**, as propriedades **om:numerical_value** e **om:unit_of_measure**, todos definidos na ontologia *om-1.6.owl* (OM) que podem ser vistas na Figura 13, na seção 5.4.2.

Em relação ao aspecto temporal, da mesma forma que na modelagem anterior, optou-se pelo reuso da classe **DurationDescription** (Figuras 14 e 15) da ontologia *time.owl* (OT).

6.4.3 Identificação das Classes e Relações Associadas aos Sensores

Nesta etapa, assim como na modelagem “A”, também houve o reuso da classe **Sensor** (Figura 16 na Seção 5.4.3), da ontologia *ssn.owl* (SE) para representação do dispositivo de medição.

6.4.4 Identificação das Classes e Relações Associadas à Proveniência dos Dados

Assim como proposto na modelagem “A”, nas descrições da Seção 5.4.4, a classe a ser reutilizada para a proveniência de dados é **Agent** e sua propriedade **wasAttributedTo**, presentes na ontologia *prov-o.owl* (PR) da W3C que podem ser vistas na Figura 17.

Após o levantamento para aplicação da reusabilidade, percebeu-se que algumas modificações foram feitas nas classes e propriedades inicialmente identificadas, conforme mostram respectivamente, os Quadros 19 e 20.

Quadro 19 – Enumeração dos termos a serem modelados como classes.

Termo	Nome da Classe	Descrição	Superclasse
Poluente	PO:Pollution	Poluente em geral	-
Poluente	Pollutant	Poluente em geral	-
Poluente do Ar	PO:Air_pollution	Poluente do ar	PO:Pollution
Poluente do Ar	Air_pollutant	Poluente do ar	Pollutant
Antropogênico	Antropogenic	Tipo de poluente gerado por atividade humana	Air_Pollutant
Natural	Natural	Tipo de poluente presente na natureza s/ interferência humana	Air_Pollutant
Ozônio	O3	Poluente do ar antropogênico	Antropogenic
Dióxido de Nitrogênio	NO2	Poluente do ar antropogênico	Antropogenic
Partículas Suspensas menores que 2.5µm	PM2.5	Poluente do ar natural	Natural
Sensor	SE:Sensor	Sensor ou dispositivo para coletar dados	DUL:PhysicalObject
Estação de coleta	Station	Utilizada para coletar dados	SE:Sensor
Estação automática	Automatic	Estação que opera em modo automático.	Station
Estação manual	Manual	Estação de operação manual	Station
Medição	Measurement	Medição coletada por uma estação de coleta	OM:Measure
Unidade de medida	OM:Unit_of_measure	Grandeza que caracteriza um valor medido	-
Categorização	Categorizations	Categorizações para uma escala de valores	-
Categoria de índice	IndexCategorization	Categorização para índices	Categorizations
Duração	OT:DurationDescription	Duração de um evento	
Periodicidade	Periodicity	Intervalo de tempo	OT:DurationDescription
Hora	OT:Hour	Duração em quantidade de horas	OT:DurationDescription
Dia	OT:Days	Duração em quantidade de dias	OT:DurationDescription
Indicador	UserIndicator	Indicador de referência para o usuário, a ser obtido de um conjunto de medições	-
Indicador com categorização	Categorized_UI	Indicador de referência para o usuário que possui categorias para enquadramento	UserIndicator
Indicador sem categorização	NonCategorized_UI	Indicador de referência para o usuário sem possuir categorias para enquadramento	UserIndicator

Índice	Index	Índice de referência a ser obtido de um indicador	-
Índice derivável	Index_with_DF	Índice de referência a ser obtido por meio de uma forma de derivação sobre indicadores	Index
Fórmula	Formula	Expressão matemática para gerar o valor resultante de um índice	-
Forma de derivação	DerivationForms	Forma de derivação ou de agregação sobre um conjunto de dados	-
Média	Average	Forma de derivação	DerivationForms
Máxima	Maximum	Forma de derivação	DerivationForms
Mínima	Minimum	Forma de derivação	DerivationForms
Agente	PR:Agent	Agente responsável pela coleta e publicação da medição	-

Fonte: Autoria própria.

Em seguida, pode ser vista uma relação das propriedades no Quadro 20, no qual as propriedades oriundas de reuso estão na cor cinza.

Quadro 20 – Enumeração dos termos a serem modelados como propriedades.

Termo	Nome da propriedade	Descrição	Domínio x Imagem*	Tipo
Momento da medição	hasHappenedOn	Momento no tempo que ocorre a coleta data/hora	Measurement, xsd:dateTime	Funcional
Limite inferior do intervalo	hasLowerLimit	Valor mínimo de um intervalo	[Categorizations], xsd:decimal	Funcional
Limite superior do intervalo	hasUpperLimit	Valor máximo de um intervalo	[Categorizations], xsd:decimal	Funcional
Medição	hasMeasurementOf	Tem a medição de um poluente	Measurement, Air_pollutant	Funcional
Valor da medição	OM:numerical_value	Valor real de uma medição	Measure, xsd:decimal	Funcional
Unidade de medida	OM:unit_of_measure_or_measurement_scale	Unidade de medida ou escala de medida	measure, Unit_of_measure_or_measurement_scale	Funcional
Periodicidade	hasPeriodicity	Periodicidade de	[UserIndicator], [Periodicity]	Funcional
Duração ou periodicidade	OT:days	Dias	ot:DurationDescription, xsd:decimal	Funcional
Duração ou periodicidade	OT:hours	Horas	ot:DurationDescription, xsd:decimal	Funcional
Categorização	isCategorizedBy	É categorizado por	[Index], [Categorizations]	Não Funcional
Derivação ou agregação	isDerivedBy	É derivado de	[UserIndicator; Index_with_DF], DerivationForms	Funcional
Indicador	isIndicatorFor	É indicador para o poluente	[UserIndicator], Air_pollutant	Funcional
Aparelho da medição	takesPlaceAt	Ocorre em	[Measurement], Station	Funcional

(*) Objetos entre [] são classes implícitas na definição

Fonte: Autoria própria.

Reiterando os prefixos das ontologias de reuso presentes nestes objetos pertencem as ontologias: PO-Pollution.owl, SE-Sensor.owl, OM-OM1.6.owl, OT:Time.owl, PR:Prov-o.owl e DUL:DUL.owl.

6.5 ESPECIFICAÇÃO DOS PADRÕES DE MODELAGEM

Repete-se a ordem adotada na Seção 5.5 para realizar a especificação dos padrões da modelagem “B”, iniciando com a classe das medições. Da mesma forma que na modelagem “A” se definiu o padrão **Measurement**, nesta abordagem serão mantidas as mesmas propriedades, visto que, a classe para medições não sofreu alterações em seus requisitos. O padrão **Measurement**, para a modelagem “B” no formato proposto na seção 3, ficou assim definido:

Nome do padrão: **Measurement**

Questões de competência:

(Q1) Quais são as medições do poluente <P> coletados pela estação <S> de responsabilidade do agente <A> que ocorreram no período <T>?

(Q12) Qual a unidade de medida <U> utilizada na medição do poluente <P>?

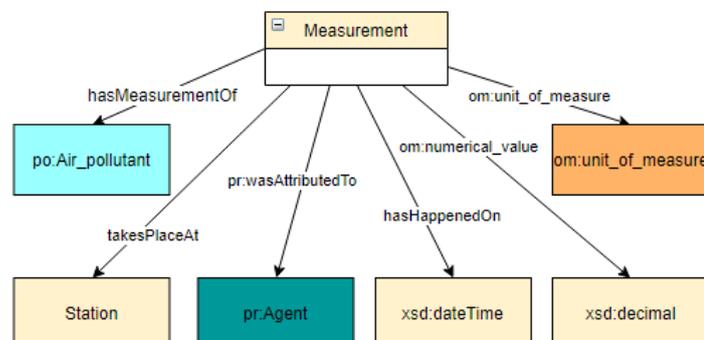
Elementos:

```

Class: Measurement

SubClassOf:
    aq:hasMeasurementOf exactly 1 po:Air_pollutant
    aq:takesPlaceAt exactly 1 aq:Station
    pr:wasAttributedTo exactlty 1 pr:Agent
    aq:hasHappenedOn exactly 1 xsd:dateTime
    om:numerical_value some xsd:decimal
    om:unit_of_measure some om:Unit_of_measure
  
```

Diagrama:

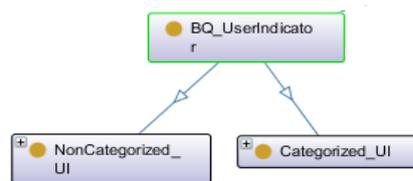


Prosseguiu-se com a definição do padrão para a classe de indicadores, que será tratado nesta modelagem com o nome de **BQ_UserIndicator**. Neste padrão, propõe-se estender a

definição proposta na modelagem anterior, já que pelo cenário proposto nesta abordagem, a necessidade de adequação de indicadores em categorias não é exigida para gerar o seu índice. Escolheu-se então, fazer uma especialização para cada caso, ou seja, um indicador para quando há categorização e outro para quando não há categorização.

Este padrão vai conter as mesmas propriedades descritas para **UserIndicator** (vide Seção 5.5) com exceção da propriedade **isCategorizedBy**. As propriedades remanescentes serão herdadas por duas especializações nomeadas aqui de: **Categorized_UI** e **NonCategorized_UI**, como ilustra o diagrama da Figura 20. Dentre as classes herdeiras, a classe **Categorized_UI** vai conter a propriedade **isCategorizedBy**, para que assim se possa associar um indicador com uma categorização, expandindo o alcance da modelagem.

Figura 20 – Subclasses da classe BQ_UserIndicator.



Fonte: O autor (2018)

A subclasse **NonCategorized_UI** pode ser definida com a mesma propriedade para deixar explícita a ausência de necessidade de uma categoria de indicador, através do uso da cardinalidade **MAX:[0]** ou **EXACTLY:[0]**. Entretanto, os *reasoners* das versões utilizadas do Protégé apresentam problemas para lidar com estas regras de restrição, deixando ao padrão desta classe, duas alternativas de modelagem (vide padrão **NonCategorized_UI** mais adiante).

Devido à experimentação da modelagem na construção da ontologia, se recomenda utilizar a propriedade **isCategorizedBy** somente na subclasse **Categorized_UI**. O padrão da classe de indicadores ficou assim definido.

Nome do padrão: **BQ_UserIndicator**

Questões de competência:

- (Q2) Quais indicadores fazem referência ao poluente <P>?
- (Q3) Como o indicador <R> é calculado a partir das medições?
- (Q4) Como o indicador <R> é categorizado?
- (Q5) Qual a categorização para o indicador <R> na periodicidade <H>?
- (Q14) Qual a periodicidade <H> e o poluente <P> calculados pelo indicador <R>?

Elementos:

```

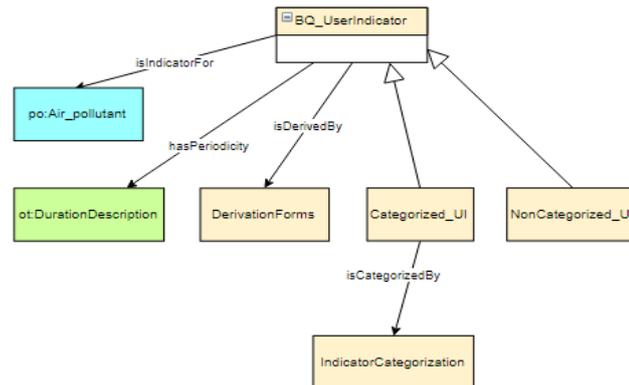
Class: bq:BQ_UserIndicator

SubClassOf:
  bq:hasPeriodicity exactly 1 po:DurationDescription,
  bq:isDerivedBy exactly 1 bq:DerivationForms,
  bq:isIndicatorFor exactly 1 po:Air_pollutant

DisjointUnionOf:
  bq:Categorized_UI, bq:NonCategorized_UI

```

Diagrama:



No parâmetro ‘Elementos’ deste padrão de modelagem, encontram-se as propriedades para as relações do indicador com um poluente, do indicador com uma forma de derivação e, do indicador com a periodicidade na qual é calculado seu valor, pela forma de derivação. A modelagem prossegue com os padrões das duas especializações de **BQ_UserIndicator**.

Nome do padrão: **Categorized_UI**

Questões de competência:

(Q4) Como o indicador <R> é categorizado?

(Q5) Qual a categorização para o indicador <R> no período <T>?

Elementos:

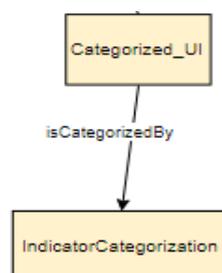
```

Class: Categorized_UI

SubClassOf:
  bq:BQ_UserIndicator
  bq:isCategorizedBy some bq:IndicatorCategorization

```

Diagrama:



Nome do padrão: **NonCategorized_UI**

Questão de competência: N/A

Elementos:

```
Class: NonCategorized_UI
SubClassOf:
  bq:BQ_UserIndicator
  bq:isCategorizedBy max 0 bq:IndicatorCategorization
```

As subclasses **Categorized_UI** e **NonCategorized_UI**, sendo especializações do padrão **BQ_UserIndicator**, têm o papel de informar se um indicador tem que ser enquadrado em categorias antes de ser associado ao índice que o utilizará.

Sobre a classe de categorizações, esta classe irá utilizar mesmo padrão proposto na modelagem “A” identificado pelo termo **Categorizations** que continua sendo composto das subclasses **IndexCategorization** e **IndicatorCategorization**.

Propõe-se que a especialização **IndicatorCategorization** seja mantida neste padrão, pois além de existir a especialização **Categorized_UI** que faz uso dela, possibilita-se a criação de indicadores que tenham que ser categorizados, expandindo a aplicação da ontologia que fizer uso deste padrão de modelagem, apesar de não ser um requerimento do cenário proposto. Assim, o padrão para a classe **Categorizations** que é o mesmo proposto na modelagem anterior, ficou assim definido.

Nome do padrão: **Categorizations**

Questões de competência:

(Q7) Qual é a categoria para o valor <V> do indicador <R>?

(Q8) Qual é a categoria para o valor <V> do índice <I>?

Elementos:

```
Class: Categorizations
SubClassOf:
  hasLowerLimit exactly 1 xsd:decimal,
  hasUpperLimit exactly 1 xsd:decimal

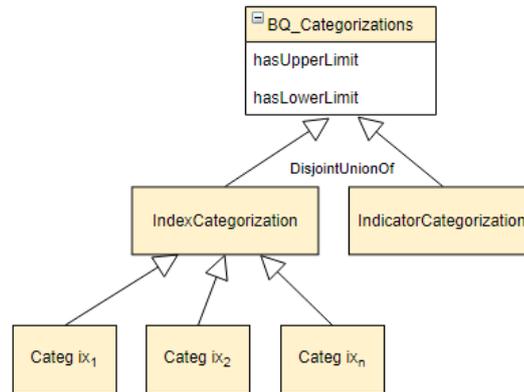
DisjointUnionOf:
  IndexCategorization, IndicatorCategorization

Class: <categoriação>
SubClassOf:
  IndexCategorization ..or.. IndicatorCategorization

DisjointUnionOf:
  <categoria>1, <categoria>2 .. <categoria>n

Class: <caategoria>i
SubClassOf:
  <categoriação>
```

Diagrama:



Lembrando que a partir das medições é possível gerar os valores dos indicadores de poluentes. Isto pode ocorrer pelo uso de consultas em SPARQL, como já explicado na Seção 5.5. O padrão proposto na modelagem “A”, na forma de uma consulta SPARQL será igualmente adotado nesta modelagem, que ficou assim definido.

Nome do padrão: **SPARQL_for_UserIndicator**

Questão de competência:

(Q6) Qual o valor para o indicador <R> no período <T> com a forma de derivação <D>?

Elementos:

```

SELECT ?ui (?<df> (?valor) AS ?'<df>') ?df ?pol
WHERE { ?ui aq:isIndicatorFor ?pol .
        ?ui aq:isDerivedBy ?df .
        ?m aq:hasMeasurementOf ?pol .
        ?m aq:hasHappenedOn ?datahora .
        ?m om:numerical_value ?valor .

FILTER ( ?ui = <userIndicator> &&
        ?datahora >= 'YYYY-MM-DDThh:mm:ss'^xsd:dateTime &&
        ?datahora <= 'YYYY-MM-DDThh:mm:ss'^xsd:dateTime )
}
GROUP BY ?ui ?df ?pol |
  
```

Assim como na modelagem “A”, existe a necessidade de se definir a classe **DerivationForms**, cujo objetivo é indicar a forma de derivação da concentração do poluente a partir das medições coletadas. Também neste caso, pode-se adotar igualmente o padrão anteriormente proposto, que pode ser revisto nas Figuras 18 e 19 da Seção 5.5. Seguindo, deve-se modelar o padrão para a classe de índices, o qual seguirá igualmente a modelagem anterior. O padrão a ser adotado para esta classe pelo formato proposto ficou assim definido.

Nome do padrão: **Index**

Questões de competência: (Q9) Quais são as categorias definidas para o índice <I>?

(Q13) Qual é a forma de derivação <D> do índice <I>?

Elementos:

```

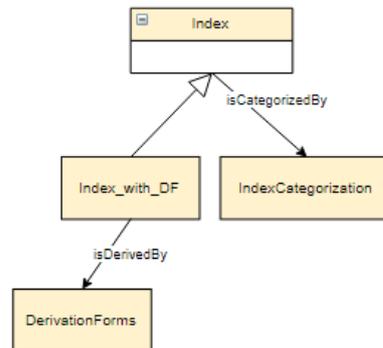
Class: Index
SubClassOf:
  aq:isCategorizedBy some aq:IndexCategorization

Class: aq:AQ_Index_with_DF
EquivalentTo:
  aq:isDerivedBy exactly 1 aq:DerivationForms

SubClassOf:
  aq:Index

```

Diagrama:



Em relação ao padrão para a classe de ligação entre índice-indicador, utilizada na modelagem “A”, para que pudesse ser empregada nesta modelagem, precisaria que a dupla dependência da fórmula com as categorias de indicador e de índice estivesse referenciada. Como não é preciso associar uma categoria de indicador nesta abordagem, se deve buscar uma alternativa de solução.

A alternativa encontrada foi de se criar um novo padrão para tratar somente a correlação entre a categoria de índice e sua fórmula já que se deve tratar a ausência de uma categorização de indicadores neste caso. Este padrão foi chamado de **IndexToFormula** e as suas propriedades estão relacionadas a seguir:

- **makesRefTo**: propriedade que relaciona o índice ao seu conjunto de categorias, representado por **IndexCategorization**;
- **hasFormula**: propriedade que relaciona as categorias do índice com uma fórmula que será representada na forma indicativa pelo tipo de dado **string**.

Prosseguindo com a definição formal, o formato deste padrão ficaria conforme segue:

Nome do padrão: **IndexToFormula**

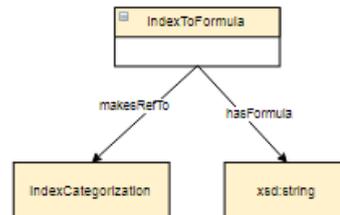
Questão de competência:

(Q11) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI>?

Elementos:

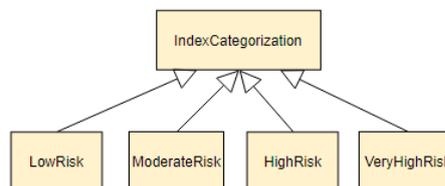
```
Class: IndexToFormula
SubClassOf:
  aq:makesRefTo exactly 1 IndexCategorization
  aq:hasFormula exactly 1 xsd:string
```

Diagrama:



Em relação às categorias com as faixas de valores para o índice proposto no cenário. Estas categorias devem ser declaradas, com o uso dos termos: **LowRisk**, **ModerateRisk**, **HighRisk**, e **VeryHighRisk**. Estas categorias devem então ser subclasses disjuntas da classe **IndexCategorization** como ilustrado na Figura 21.

Figura 21 – Subclasses disjuntas que completam a classe IndexCategorization.



Fonte: O autor (2018)

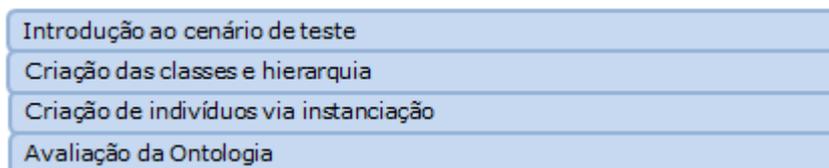
Com a definição desta classe que contém as categorias aplicáveis ao índice do cenário proposto nesta abordagem, se encerra a etapa de modelagem dos padrões. A seguir, dois estudos de caso irão fazer uso dos padrões e classes que foram modelados para a descrição de índices e indicadores em duas ontologias, cada qual representando as modelagens desenvolvidas na instanciação da metodologia proposta neste trabalho.

7 ESTUDOS DE CASO

A avaliação dos padrões de modelagem propostos por meio de estudos de caso é a última etapa da metodologia proposta na Seção 4. Através desta etapa é possível: confirmar a viabilidade das modelagens A e B, isto é, se é possível criar ontologias com os padrões propostos previamente e, também, avaliar o desempenho em relação aos critérios elencados previamente. Recorda-se que as ontologias desenvolvidas se destinam a descrever medições, indicadores e índices. Deve ser possível a uma aplicação que utilize uma destas ontologias, construir visões sobre as medições, isto é, derivar indicadores sobre estas e, por fim, calcular índices e categorizar os resultados obtidos.

O conteúdo desta seção ficou dividido em duas subseções, onde cada uma trata da utilização de uma das modelagens para trabalhar com os padrões propostos de construção das ontologias sobre poluição do ar. Foram propostas duas modelagens diferentes para a representação da relação índice-indicador em ontologias a serem avaliadas com base nos critérios escolhidos, a saber: concisão, expansibilidade, completude e qualidade das saídas produzidas. As subseções seguem uma sequência de passos ilustrados na Figura 22.

Figura 22 – Etapas para desenvolvimento do Estudo de Caso



7.1 DESENVOLVIMENTO DA ONTOLOGIA AIRQUALITYONTO (AQ) PELA MODELAGEM “A”

A modelagem “A”, desenvolvida pelo uso da metodologia proposta, foi aplicada para a construção da ontologia AirQualityOnto, identificada pelo prefixo AQ. Os padrões propostos nesta modelagem foram utilizados para a criação dos elementos desta ontologia, bem como, elementos de ontologias de alto nível foram reutilizados. Os passos percorridos para a sua criação estão descritos a seguir.

7.1.1 Introdução ao Cenário de Teste

Neste cenário, serão considerados os requisitos relativos ao índice IMECA, já comentado na seção sobre o referencial teórico utilizado neste trabalho. Este índice utiliza-se de diferentes fórmulas por poluente, as quais podem variar de acordo com o enquadramento do nível de concentração obtido para seus indicadores através da derivação feita (vide Seção 5.1). As categorias da qualidade do ar adotadas pelo índice IMECA são divididas em cinco níveis, i.e., ‘buena’, ‘regular’, ‘mala’, ‘muy mala’ e ‘extremadamente mala’.

Os poluentes que são monitorados em tempo real são O₃, NO₂, SO₂, CO, PM_{2.5} e PM₁₀. Dependendo da periodicidade definida para cada poluente, uma média é calculada sobre as medições obtidas nos seus intervalos de tempo.

7.1.2 Criação das Classes e Hierarquia

Os termos que fazem parte do domínio relativo ao monitoramento da qualidade do ar quando representam conceitos sobre o assunto escolhido, acabam por se tornar classes dentro da hierarquia da ontologia. Lembrar que a ontologia AQ foi construída com o uso dos padrões da modelagem “A” e também com o reuso de elementos de outras ontologias.

Depois da construção da ontologia AQ ainda precisam ser definidas subclasses ou especializações de classes já criadas na ontologia AQ e relacionadas na Seção 5. No Quadro 21 abaixo, é apresentada a relação dos elementos que foram definidos na ontologia para que o estudo de caso pudesse ser executado.

Quadro 21 – Classes da ontologia AQ

Classe	Descrição	Subclasse de
IMECA_Good	Categoria de qualidade do ar do índice IMECA	IndexCategorization
IMECA_Regular	Categoria de qualidade do ar do índice IMECA	IndexCategorization
IMECA_Poor	Categoria de qualidade do ar do índice IMECA	IndexCategorization
IMECA_Bad	Categoria de qualidade do ar do índice IMECA	IndexCategorization
IMECA_Critical	Categoria de qualidade do ar do índice IMECA	IndexCategorization
PM10_IMECA_Good	Categoria de concentração para um indicador do poluente PM10 do índice IMECA	IndicatorCategorization
PM10_IMECA_Regular	Categoria de concentração para um indicador do poluente PM10 do índice IMECA	IndicatorCategorization
PM10_IMECA_Poor	Categoria de concentração para um indicador do poluente PM10 do índice IMECA	IndicatorCategorization
PM10_IMECA_Bad	Categoria de concentração para um indicador do poluente PM10 do índice IMECA	IndicatorCategorization
PM10_IMECA_Critical	Categoria de concentração para um indicador do poluente PM10 do índice IMECA	IndicatorCategorization

Fonte: Autoria própria.

7.1.3 Criação de Indivíduos Via Instanciação

Os indivíduos que são inseridos na ontologia, de forma automática ou manual, devem ser classificados segundo os termos da hierarquia de classes. Deste modo, irão se tornar instâncias membros de uma das classes do modelo, por meio da tipificação, isto é, atribuição explícita de um tipo para o indivíduo. Os tipos dos indivíduos, se definidos explicitamente, podem ser descritos por meio de classes da hierarquia, propriedades de objetos ou por meio de propriedades de dados.

Mais apropriado do que tipificar é declarar as propriedades (de objetos ou de dados) com suas características, que pertençam aos critérios de definição da classe (tipo do indivíduo). Neste caso, o *reasoner* realiza a inferência, identificando sua classificação. A seguir, no Quadro 22, estão listadas as instâncias que foram inseridas em AQ em conjunto com as propriedades e seus valores.

Quadro 22 – Instâncias da ontologia AQ.

Classe do indivíduo	Nome	Propriedade	Valor da propriedade
PO:Air_pollutant	PM10	AQ:hasPeriodicity	24
	O3	AQ:hasPeriodicity	1
	NO2	AQ:hasPeriodicity	1
	SO2	AQ:hasPeriodicity	24
	CO	AQ:hasPeriodicity	8
	PM2.5	AQ:hasPeriodicity	24
OT:DurationDescription	1	OT:hours	1
	4	OT:hours	4
	8	OT:hours	8
	24	OT:hours	24
AQ:Station	ASSIS	operatesInMode	Manual
PR:Agent	Governo_CM		
AQ:DerivationForms	Average		
	Maximum		
	Minimum		
OM:unit_of_measure	microgram_per_cubic_meter	OM:dimension	Density dimension
		OM:numerator	Microgram
		OM:denominator	Cubic meter
		OM:symbol	“µg/m3”
Measurement	AQM_PM10_ASSIS_3981_2017	AQ:hasMeasurementOf	PM10
		AQ:takesPlaceAt	ASSIS
		PR:wasAttributedTo	Governo_CM
		OM:numerical_value	37
		AQ:hasHappenedOn	2017-06-17T09:00:00
		OM:unit of measure	Microgram_per_cubic_meter
Measurement	AQM_PM10_ASSIS_3982_2017	AQ:hasMeasurementOf	PM10
		AQ:takesPlaceAt	ASSIS
		PR:wasAttributedTo	Governo_CM
		OM:numerical_value	65
		AQ:hasHappenedOn	2017-06-17T10:00:00
		OM:unit of measure	Microgram_per_cubic_meter

BQ:IndexCategorization	IMECA_good	AQ:hasLowerLimit	0
		AQ:hasUpperLimit	50
	IMECA_regular	AQ:hasLowerLimit	51
		AQ:hasUpperLimit	100
	IMECA_poor	AQ:hasLowerLimit	101
		AQ:hasUpperLimit	150
	IMECA_bad	AQ:hasLowerLimit	151
		AQ:hasUpperLimit	200
	IMECA_critical	AQ:hasLowerLimit	201
		AQ:hasUpperLimit	500
BQ:IndicatorCategorization	PM10_IMECA_good	AQ:hasLowerLimit	0
		AQ:hasUpperLimit	60
	PM10_IMECA_regular	AQ:hasLowerLimit	61
		AQ:hasUpperLimit	120
	PM10_IMECA_poor	AQ:hasLowerLimit	121
		AQ:hasUpperLimit	220
UserIndicator	UI_4_IMECA	AQ:hasPeriodicity	24
		AQ:isCategorizedBy	PM10_IMECA_good
		AQ:isIndicatorFor	PM10
		AQ:isDerivedBy	Average
AQ:IndicatorToIndexFormula	UI_4_IMECA_to_IMECA_Mexico	AQ:makesRefTo	IMECA_regular
		AQ:makesRefTo	PM10_IMECA_good
		AQ:hasFormula	"C_PM10 * (5/6)"
Index	IMECA_Mexico	AQ:isDerivedBy	Average
		AQ:isCategorizedBy	[IMECA_good]

Fonte: Autoria própria.

É importante comentar que os valores utilizados nas propriedades e os nomes de classes instanciadas na ontologia são fictícios, entretanto, estes conteúdos se baseiam no material pesquisado do referencial teórico. A sequência para o cálculo do índice utilizando esta ontologia executa os seguintes passos:

- 1) as medições lidas dos sensores são instanciadas como indivíduos da classe Measurement;
- 2) o usuário então requer a criação de um indicador pela instanciação da classe UserIndicator;
- 3) o usuário define a periodicidade, o poluente e a forma de derivação deste indicador;
- 4) a aplicação executa a query SPARQL do padrão proposto utilizando um intervalo equivalente à periodicidade definida selecionando as medições do poluente ligado ao indicador e aplicando a forma de agregação que também foi atribuído ao indicador:
 - a. o processamento deste cálculo vai selecionar as duas medições presentes e aplicar a média sobre seus valores obtendo $65 + 37 = 102/2 = 51$;
- 5) a aplicação recebe o valor da concentração para o indicador e então faz o enquadramento na categoria que o abrange, obtendo assim a categorização do indicador:

- a. este valor será enquadrado nas categorias de indicador como PM10_IMECA_good;
- 6) a aplicação utiliza o mesmo valor para enquadrá-lo nas categorias do índice, obtendo assim sua categorização:
- a. este valor será enquadrado nas categorias de índice como IMECA_regular;
- 7) com as duas categorias identificadas, a aplicação interpreta a fórmula correspondente para cálculo do valor do índice:
- a. a fórmula associada a esta combinação é dada pela expressão: $C_{PM10} * (5/6)$;
- 8) a aplicação executa a fórmula resultando no valor final, o qual deve ser novamente enquadrado nas categorias do índice para descobrir qual delas representa o nível de qualidade do ar para o poluente escolhido no passo 3;
- a. a expressão é executada como $I = 51 * (5/6) = 42,50$ e este valor é enquadrado na categoria do índice de nome IMECA_good.

7.1.4 Avaliação da Ontologia (AQ)

A avaliação da ontologia AQ criada com a aplicação da modelagem “A” é apresentada no Quadro 23 a seguir. Procurou-se avaliar se as questões de competência são respondidas corretamente. Adicionalmente, foi incluída em qual linguagem semântica foi construída a *query* que representa a questão de competência. Como existe ao menos um padrão que atendeu corretamente as questões de competência propostas, pode-se dizer que a modelagem produziu um resultado satisfatório. As consultas aplicadas nas linguagens DL ou SPARQL, para esta avaliação, estão disponíveis no Apêndice A.

Quadro 23 – Avaliação dos Padrões da Modelagem “A” pelas Questões de Competência

Questões de competência	Padrão de modelagem “A”	Forma de verificação
(Q1), (Q12)	Measurement	SPARQL query
(Q2)*, (Q3), (Q4), (Q5)**, (Q14)**	UserIndicator	(*)DL query SPARQL queries (**) ot:DurationDescription
(Q6)	SPARQL_for_UserIndicator	SPARQL query
(Q7), (Q8)	Categorizations	SPARQL query
(Q9), (Q13)*	Index	SPARQL query (*) Index_with_DF
(Q10), (Q11)	IndicatorToIndexFormula	SPARQL query

Fonte: Autoria própria.

Pela aplicação dos padrões da metodologia “A” e aplicação dos passos propostos na introdução desta seção (Figura 22), foi possível criar a ontologia AirQualityOnto (AQ). As classes e propriedades que fazem parte das declarações deste modelo foram utilizadas e indivíduos foram instanciados na ontologia com sucesso.

Contudo, observou-se que não foi possível instanciar um indicador que não fosse enquadrado em categorias, evidenciando uma característica negativa por possuir definições incompletas, como apontado em (GÓMEZ-PÉREZ, 1996). A falta de um tratamento para indicadores sem categorização como aqueles utilizados pelo índice do Canadá (AQHI) demonstrou que é possível a alteração da modelagem para que trate esta necessidade. Fato este que evidencia a possibilidade do modelo em ser expandido, porém, ilustra uma deficiência no critério da completude.

Sobre o critério da concisão, segundo Gómez-Pérez (1996), deve-se trabalhar para que não existam definições não aproveitadas ou redundantes. Neste critério, pode-se dizer que a modelagem atendeu parcialmente a condição, pois muitas classes e subclasses, além também de muitas propriedades, não foram utilizadas como, por exemplo: *pr:Organization*, *se:Output*, *om:phenomenon*, entre outras. Por outro lado, não existem definições redundantes no modelo. Em relação ao critério sobre a qualidade das saídas geradas pela ontologia quando inquirida ou por *queries* em DL ou em SPARQL, ela trouxe as respostas esperadas.

Dadas estas considerações sobre os critérios de avaliação, fez-se uso do quadro 11, proposto na metodologia (Seção 4.6), para demonstrar os critérios, as avaliações feitas e a pontuação dada de uma escala entre 1 a 3. Reiterando que as pontuações dos critérios significam respectivamente: não atendido, parcialmente atendido e atendido. O Quadro 24 a seguir, compreende as avaliações aplicadas para a modelagem “A” e sua ontologia correspondente, considerando os critérios de avaliação escolhidos, o desempenho da ontologia em responder as questões de competência e, o desempenho do *reasoner* em classificar as instâncias corretamente.

Quadro 24 – Avaliação da ontologia AQ pela modelagem “A”

Ontologia AQ: AirQualityOnto – Modelagem “A”			
Questões s/ critérios	Qtde	Observação	Pontuação
Ontologia foi construída?	-	Sim.	N/A
Tem completude?	-	Faltam alguns axiomas, mas funcionou para o estudo de caso.	2
Permite expansão?	-	Sim, nas classes Categorizations e Index.	3
Modelo é conciso?	-	Falta remover classes e propriedades sem uso. Mas não existem definições prolixas.	2

Respondeu todas as questões de competência?	-	Sim.	N/A
Classificação pelo <i>reasoner</i> apresentou erro? Quantos?	0	Não. Ele classificou corretamente os indivíduos.	3
Qualidade das saídas está satisfatória?	-	Sim. As fórmulas e formas derivadas corretamente. O valor calculado do índice indicou a categoria correta da qualidade do ar.	3
-	-	Média Geral	2,6

Fonte: Autoria própria.

Nota. Foi disponibilizada no Apêndice B, uma estatística copiada da ferramenta Protégé, sobre os elementos presentes nas ontologias criadas nos dois estudos de caso.

7.2 DESENVOLVIMENTO DA ONTOLOGIA AIRQUALITYONTO (BQ) PELA MODELAGEM “B”

7.2.1 Introdução ao Cenário de Teste

Nesta prova de conceito será utilizada a construção do índice do Canadá, conhecido por AQHI. Este índice está baseado somente em indicadores de três poluentes e não utiliza faixas de categorização para eles. Tem apenas uma fórmula de cálculo a ser executada, cujo valor final deve ser enquadrado em uma das quatro categorias de risco de poluição do ar, que são: *‘low risk’*, *‘moderate risk’*, *‘high risk’* e *‘very high risk’*. Os poluentes que são monitorados são NO₂, O₃ e PM_{2,5}, sendo que os dois primeiros são medidos em ppb e o último, em µg/m³. A periodicidade para emissão dos indicadores é de 3h.

7.2.2 Criação das Classes e Hierarquia

Seguindo a mesma abordagem dada para o estudo de caso anterior, somente as especializações das classes já criadas na ontologia BQ necessitam ser definidos. A ontologia foi criada com o nome **AirQualityOnto_B** (BQ) e seus termos são semelhantes aos da relação apresentada na ontologia (AQ). A seguir, é apresentada uma relação com as classes de sua hierarquia no Quadro 25.

Quadro 25 – Classes da ontologia BQ

Classe	Descrição	Subclasse de
NonCategorized_UI	Indicador sem categorização, derivado de um conjunto de medições por uma forma de derivação matemática	BQ_UserIndicator
Categorized_UI	Indicador categorizável, derivado de	BQ_UserIndicator

	um conjunto de medições por uma forma de derivação matemática	
AQHI_LowRisk	Categoria de risco baixo do índice AQHI	IndexCategorization
AQHI_ModerateRisk	Categoria de risco moderado do índice AQHI	IndexCategorization
AQHI_HighRisk	Categoria de risco alto do índice AQHI	IndexCategorization
AQHI_VeryHighRisk	Categoria de risco muito alto do índice AQHI	IndexCategorization
IndexToFormula	Ligação entre o índice e sua fórmula	-

Fonte: Autoria própria.

7.2.3 Criação de Indivíduos Via Instanciação

Como explicado na modelagem anterior, os indivíduos inseridos na ontologia, de forma automática ou manual, se tornam instâncias membros de uma das classes do modelo. Ou por meio de sua tipificação, ou pelas propriedades adicionadas em suas declarações. No Quadro 26, podem ser visualizadas as principais instâncias inseridas na ontologia BQ fazendo uso das propriedades que definem as classes de sua hierarquia. Estas instâncias compreendem o conteúdo necessário para processar o índice AQHI por meio deste estudo de caso.

Quadro 26 – Instâncias da ontologia BQ

Classe do indivíduo	Nome	Propriedade	Valor da propriedade
PO:Air_pollutant	PM2.5	BQ:hasPeriodicity	24
	O3	BQ:hasPeriodicity	1
	NO2	BQ:hasPeriodicity	1
	CO	BQ:hasPeriodicity	8
OT:DurationDescription	1	OT:hours	1
	3	OT:hours	3
	8	OT:hours	8
	24	OT:hours	24
BQ:Station	TOR_DWNT	BQ:isStationOf	Toronto
		BQ:operatesInMode	Automatic
PR:Agent	EnvironmentCanada		
BQ:DerivationForms	Average		
OM:unit_of_measure	microgram_per_cubic_meter	OM:dimension	Density dimension
		OM:numerator	Microgram
		OM:denominatot	Cubic meter
		OM:symbol	“µg/m3”
OM:unit_of_measure	Parts_per_billion	OM:dimension	Density dimension
		OM:numerator	Microgram
		OM:denominatot	Liter
		OM:symbol	“ppb”
Measurement	BQM_PM2.5_TOR_DWNT_9_2018	BQ:hasMeasurementOf	PM2.5
		BQ:takesPlaceAt	TOR_DWNT
		PR:wasAttributedTo	EnvironmentCanada
		OM:numerical_value	31
		BQ:hasHappenedOn	2018-01-01T09:00:00
OM:unit of measure	Microgram_per_cubic_meter		

Measurement	BQM_NO2_TOR_D WNT_8_2018	BQ:hasMeasurementOf	NO2
		BQ:takesPlaceAt	TOR_DWNT
		PR:wasAttributedTo	EnvironmentCanada
		OM:numerical_value	18
		BQ:hasHappenedOn	2018-01-01T08:00:00
		OM:unit of measure	Parts_per_billion
Measurement	BQM_NO2_TOR_D WNT_9_2018	BQ:hasMeasurementOf	NO2
		BQ:takesPlaceAt	TOR_DWNT
		PR:wasAttributedTo	EnvironmentCanada
		OM:numerical_value	20
		BQ:hasHappenedOn	2018-01-01T09:00:00
		OM:unit of measure	Parts_per_billion
Measurement	BQM_O3_TOR_DW NT_9_2018	BQ:hasMeasurementOf	O3
		BQ:takesPlaceAt	TOR_DWNT
		PR:wasAttributedTo	EnvironmentCanada
		OM:numerical_value	17
		BQ:hasHappenedOn	2018-01-01T09:00:00
		OM:unit of measure	Parts_per_billion
BQ:IndexCategorization	AQHI_LowRisk	BQ:hasLowerLimit	1
		BQ:hasUpperLimit	3.99
	AQHI_ModerateRisk	BQ:hasLowerLimit	4
		BQ:hasUpperLimit	6.99
	AQHI_HighRisk	BQ:hasLowerLimit	7
		BQ:hasUpperLimit	10.99
	AQHI_VeryHighRisk	BQ:hasLowerLimit	11
		BQ:hasUpperLimit	100
BQ:NonCategorized_UI	UI_AQHI_PM25_TO R 31	BQ:hasPeriodicity	3
		BQ:isIndicatorFor	PM2.5
		BQ:isDerivedBy	Average
BQ:NonCategorized_UI	UI_AQHI_NO2_TO R 19	BQ:hasPeriodicity	3
		BQ:isIndicatorFor	NO2
		BQ:isDerivedBy	Average
BQ:NonCategorized_UI	UI_AQHI_O3_TOR 17	BQ:hasPeriodicity	3
		BQ:isIndicatorFor	O3
		BQ:isDerivedBy	Average
BQ:IndexToFormula	I2F_AQHI_TOR	BQ:makesRefTo	AQHI_VeryHighRisk
		BQ:hasFormula	$\{(1000/10.4)^{\{(e^{*(0.00537*O3)}-1) + (e^{*(0.000871*NO2)}-1) + (e^{*(0.000487*PM2.5)}-1)\}}$
Index	AQHI_Canada	BQ:isDerivedBy	Average
		BQ:isCategorizedBy	AQHI_LowRisk
BQ:AQHI_LowRisk	AQHI_Canada	BQ:indexValuels	[3.95]

Fonte: Autoria própria.

Nota. Os valores nas propriedades e nomes de classes instanciados na ontologia não são oficiais entretanto, são conteúdos que se baseiam no material lido do referencial teórico. A sequência para o cálculo deste índice utilizando a ontologia BQ executa os seguintes passos:

- 1) as medições são lidas dos sensores e instanciadas na classe Measurement;
- 2) o usuário então requer a criação de três indicadores da classe BQ_UserIndicator, um por poluente;

- 3) o usuário define a periodicidade, o poluente e a forma de derivação destes indicadores;
- 4) a aplicação executa a query SPARQL do padrão proposto utilizando um intervalo equivalente à periodicidade definida, selecionando as medições do poluente ligado ao indicador e aplicando a forma de agregação que também foi atribuída ao indicador – executa esta query para cada um dos poluentes:
 - a. o processamento deste cálculo vai selecionar as duas medições de NO₂ e encontrar seu valor médio logo, entre 18 e 20 tem-se 19 no indicador de NO₂, depois vai selecionar cada medição restante sendo, uma de 31 para o poluente PM₂₅ e 17 para o poluente O₃;
- 5) a aplicação utiliza o valor de concentração de cada indicador para enquadrá-lo nas categorias do índice, obtendo assim sua categorização para interpretar a fórmula correspondente para cálculo do valor do índice:
 - a. estes valores devem ser enquadrados nas categorias de índice, cujo resultado é “VeryHighRisk”;
 - b. não há necessidade de se enquadrar em categorias de indicador pois não existem neste caso;
 - c. a fórmula deste índice é única, portanto, seu conteúdo será o mesmo para todas as categorias de índices;
 - d. a fórmula associada com a categoria de índice é dada pela expressão:

$$\{(1000/10.4)*\{e^{(0.000537*O_3)-1} + e^{(0.000871*NO_2)-1} + e^{(0.000487*PM_{2.5})-1}\}\};$$
- 6) a aplicação executa a fórmula resultando no valor final, o qual deve ser novamente enquadrado nas categorias do índice para descobrir qual delas representa o nível de risco de poluição do ar dentre os possíveis:
 - a. a expressão matemática é executada, resultando no valor igual a 3,95 que, ao ser enquadrado nas categorias do índice, encontrará o nível de risco: AQHI_LowRisk.

7.2.4 Avaliação da Ontologia (BQ)

A avaliação se verifica pela ideia aplicada na modelagem “A”, onde as questões de competência são confrontadas com os padrões de modelagem para se validar a qualidade da modelagem “B”. O Quadro 27 a seguir, contém o resultado obtido para esta avaliação baseada

nas questões de competência. As consultas aplicadas em linguagem DL ou SPARQL desta avaliação, estão disponíveis no Apêndice A.

Quadro 27 – Avaliação dos Padrões da Modelagem “B” pelas Questões de Competência

Questões de competência	Padrão de modelagem “B”	Forma de verificação
(Q1), (Q12)	Measurement	SPARQL query
(Q2)*, (Q3), (Q4), (Q5)**, (Q14)**	BQ_UserIndicator	(*) DL query SPARQL query (**) ot:DurationDescription
(Q6)	SPARQL_for_UserIndicator	SPARQL query
(Q7), (Q8)	Categorizations	SPARQL query
(Q9), (Q13)*	Index	SPARQL query (*) Index_with_DF
(Q10)	IndicatorToIndexFormula	SPARQL query
(Q11)	IndexToFormula	SPARQL query

Fonte: Autoria própria.

Seguindo os passos para construção de uma ontologia com os padrões da modelagem “B”, a ontologia AirQualityOnto_B (BQ) foi construída. A opção de modelagem “B”, ao tratar de um cenário diferente, teve que tratar algumas condições novas que diferenciam o desempenho da ontologia BQ em relação à ontologia AQ.

A possibilidade de se tratar um índice como o AQHI que não categoriza seus indicadores amplia seu grau de atingimento, deixando o critério da expansibilidade em um nível mais satisfatório. Pode-se afirmar que há mais completude na ontologia BQ do que na ontologia AQ. Embora não tenha sido apresentada a criação do índice IMECA na ontologia BQ, como ambos os modelos são similares em seus construtos se percebe que tal índice mesmo tendo um indicador categorizado pode ser também, criado em BQ, com o uso da classe **Categorized_UI** especialização de **BQ_UserIndicator**.

Sobre o critério da concisão, da mesma forma que na ontologia AQ, pode-se afirmar que a modelagem atende esta condição parcialmente, pois possui classes e subclasses não utilizadas, principalmente oriundas do reuso. Em contrapartida não possui definições redundantes para as suas classes.

E por fim, quanto à qualidade das saídas geradas pela ontologia BQ, obteve-se o mesmo parecer que ocorreu com a ontologia AQ, deixando este critério em nível satisfatório. O Quadro 28 a seguir, contém a avaliação da modelagem “B” e sua correspondente ontologia.

Quadro 28 – Avaliação da ontologia BQ pela modelagem “B”

Ontologia BQ: AirQualityOnto_B - Modelagem “B”			
Questões s/ critérios	Qtde	Observação	Pontuação
Ontologia foi construída?	-	Sim.	N/A
Tem completude?	-	Os axiomas cobrem mais possibilidades (ex. IndexToFormula e Categ/NonCateg_UI).	3
Permite expansão?	-	Sim. As Classes Categorizations e Index foram estendidas.	3
Modelo é conciso?	-	Falta remover classes e propriedades sem uso. Mas não existem definições prolixas.	2
Respondeu todas as questões de competência?	-	Sim.	N/A
Classificação pelo <i>reasoner</i> apresentou erro? Quantos?	0	Não. Ele classificou corretamente os indivíduos.	3
Qualidade das saídas está satisfatória?	-	Sim. As fórmulas e formas derivadas corretamente. O valor calculado do índice indicou a categoria correta da qualidade do ar.	3
-	-	Média Geral	2,8

Fonte: Autoria própria.

8 RESULTADOS E DISCUSSÃO

Pela aplicação das etapas da metodologia proposta, foram produzidos padrões de modelagem para a criação de ontologias de monitoração da qualidade do ar. Estes padrões visam solucionar os problemas encontrados para representar a relação entre índices e indicadores necessários para realizar esta monitoração. Estes padrões foram divididos em duas abordagens (A e B) que foram testadas por meio de estudos de casos para construir ontologias que foram avaliadas pelos critérios escolhidos. Embora tenham semelhanças, os cenários propostos continham algumas diferenças de modo a permitir que mudanças na modelagem fossem exploradas e aplicadas. Estas diferenças, presentes em alguns dos padrões desenvolvidos, provocaram resultados um pouco distintos em alguns dos critérios de avaliação propostos que estão detalhados na coluna ‘observações’ do quadro comparativo. Os resultados suscitam discussões a respeito do uso de recursos ou da aplicação de uma abordagem ou de outra para solucionar os problemas encontrados durante o desenvolvimento das modelagens. No Quadro 29, é demonstrado um comparativo entre os modelos.

Quadro 29 – Comparativo das avaliações das modelagens

Comparativo das avaliações entre as ontologias AQ e BQ			
Descrição	AQ	BQ	Observação
Ontologia foi construída	Sim	Sim	Seguindo os passos da metodologia foi possível criar ambas as ontologias.
Tem completude?	Não	Sim	Devido à possibilidade que a ontologia BQ possui para descrever índices com ou sem categorias para seus indicadores, considerou-se que BQ é mais completa que a ontologia AQ.
Permite expansão?	Sim	Sim	Ambas possibilitam que sejam estendidas, ou pela alteração nas definições existentes ou pela adição de novos axiomas.
Modelo é conciso?	Não	Não	Apesar da ausência de definições redundantes nas ontologias, observou-se que faltou a remoção de muitos elementos não utilizados nos estudos de casos.
Classificação pelo reasoner apresentou erro?	Não	Não	A classificação das instâncias inseridas nas ontologias foi correta.

Fonte: Autoria própria.

A ontologia BQ sofreu uma influência mais positiva no critério da completude porque houve uma modelagem adicional para que a classe de categorizações dos indicadores pudesse trabalhar com e sem a existência de categorias como suas subclasses. A modelagem “B” ao

propor um padrão para o indicador de usuário com as especializações **Categorized_UI** e **NonCategorized_UI**, vai permitir que mais tipos de índices sejam atendidos por esta modelagem.

Em outras palavras, tornou-se possível expressar um indicador que não depende de uma categorização para derivação da fórmula de seu índice. A adoção desta abordagem gerou um efeito sobre o padrão da modelagem “A” de nome **IndicatorToIndexFormula** que teve que ser remodelado em “B”.

O padrão proposto em “B” denominado **IndexToFormula** sofreu a remoção da propriedade ligada à categoria do indicador. Em relação a esta decisão, é recomendável que ambos os padrões sejam utilizados na construção de uma ontologia. De modo que tanto índices de qualidade do ar que especificam uma fórmula dependente das categorizações dos dois conceitos (índice e indicador), sejam atendidos pelo padrão **IndicatorToIndexFormula** e que índices que dependem somente de uma categorização seja atendido pelo padrão **IndexToFormula**.

Além disto, surgiu uma dúvida decorrente dos requerimentos do índice AQHI. Este índice é dado por somente uma fórmula que deve ser utilizada para todas as categorias de seu índice. Isto significaria dizer que a classe **IndexToFormula** poderia ser definida fazendo uso de duas especializações, uma para fazer referência a uma categoria e outra para representar a ausência desta necessidade, de modo que, independentemente da categoria do índice, a fórmula a ser utilizada é a mesma. Esta proposta de solução deixaria o padrão flexível para atender os dois casos.

É importante comentar que a solução dada para a modelagem ‘A’ atende de forma eficiente, o índice de qualidade do ar da Cidade do México, contudo, os demais índices de qualidade provavelmente não fossem totalmente atendidos com o uso restrito de seus padrões. Isto ficaria evidente ao se aplicar uma prova de conceito para a descrição de outro índice de poluição do ar que não faz uso de categorias para indicadores.

9 CONCLUSÕES

Para permitir que usuários ou agentes computacionais sejam capazes de descrever índices e indicadores de poluição do ar fazendo uso da tecnologia semântica, mais especificamente por meio de ontologias, buscou-se conhecimento sobre como elas são construídas. Ficou claro que as ontologias têm um papel fundamental nesta tecnologia porque permitem que conhecimento seja representado de modo que seja compreendido e compartilhado, viabilizando a expansão da Web Semântica. Para desvendar como as ontologias são modeladas, este trabalho apresentou, sucintamente, as linguagens para descrição lógica utilizada até então e deu ênfase na pesquisa dos recursos da linguagem OWL-DL aplicáveis na representação de conhecimento.

Além disto, buscou-se na revisão bibliográfica, como ontologias sobre meio ambiente representam os conceitos e relações envolvidos no domínio dos índices sobre qualidade do ar. Desta pesquisa, foi possível perceber que não existem ontologias que abordam as características dos índices e indicadores sobre poluição do ar que remetem ao objetivo principal proposto neste trabalho.

Desta forma, este trabalho buscou recursos nas lógicas de descrição, padrões de modelagem semântica e encontrou soluções para a maior parte dos problemas identificados. Adicionalmente, para que as modelagens desenvolvidas com esta pesquisa fossem avaliadas, analisou-se quais são as formas utilizadas para avaliação de ontologias, e extraíram-se delas os critérios para esta avaliação.

Então, como a pesquisa demonstrou, foram propostos padrões de modelagem para tratar a relação parte-todo entre os conceitos principais (índices-indicadores), para representar relações entre conceitos por meio de propriedades, para representar classes compostas por subclasses que devem ser disjuntas como também, como categorias com faixas de valores podem ser definidas. Por fim, também foi indicada uma solução para representar formas e fórmulas de cálculo em ontologias. Em relação às fórmulas, apesar da solução adotada ter suprido a necessidade do trabalho, ainda se requer uma pesquisa mais aprofundada sobre a questão para que uma melhor solução seja encontrada.

Deve-se reforçar aqui que alguns dos passos aplicados na metodologia, obtidos da fundamentação teórica, foram muito importantes no desenvolvimento dos padrões de modelagem assim como para trazer mais qualidade semântica às definições. Destes passos, destacam-se a elaboração de questões de competência e utilização do reuso de conceitos e relações bem elaborados das ontologias de alto nível.

Sobre a reusabilidade, cabe aqui um questionamento sobre as consequências de sua adoção por causa da necessidade em se tratar os objetos não utilizados que são importados com aqueles que interessam às definições. Neste sentido, se identificou uma possível ação adicional que poderia ser adotada para fazer os recortes dos conceitos e relações não utilizados. Em todo o caso, se deve levar em conta, a grande vantagem na adoção do reuso de ontologias de alto nível para ampliar as possibilidades de compartilhamento semântico de informações na *internet*.

Adicionalmente, este trabalho propôs a adoção de um formato para apresentação de padrões, derivado dos parâmetros do catálogo indicados por Gangemi e Presutti (2009) e, efetuou uma avaliação dos modelos construídos pelo uso de critérios extraídos dos estudos realizados por autores especialistas em modelagem e avaliação de ontologias.

Além destas contribuições, vale comentar que o propósito principal da dissertação foi o desenvolvimento de padrões para a descrição de índices e indicadores de poluição do ar em ontologias, deixando em aberto a utilização prática de seus modelos. Apesar de não ser objetivo deste trabalho, dependendo do interesse do modelador, pode-se definir uma propriedade associada à um tipo de dado numérico (decimal ou equivalente) para armazenar o valor resultante dos cálculos e associá-la às classes do índice e do indicador. Depois, completar com uma propriedade associada com uma definição que referêncie as categorias destas duas classes, de modo que se possa derivar a categoria correta pela inferência do *reasoner*.

9.1 LIMITAÇÕES

Em relação às limitações encontradas no desenvolvimento deste trabalho, nas duas alternativas de modelagens se adotou a representação da fórmula de cálculo para obtenção do índice de qualidade pelo uso de um tipo de dados primitivo como o **String**. Esta solução resolveu o problema para representar cada um dos termos da fórmula, mas transferiu a validação sintática, interpretação da fórmula e sua execução para a aplicação. Em outra abordagem experimentada pelo autor na qual a formulação matemática foi realizada por meio de axiomas, foi possível a execução de operações exclusivamente binárias, i.e., com somente dois argumentos pelo uso da linguagem SWRL¹⁴. Contudo, esta solução foi preterida por implicar em uma intensa comunicação entre a aplicação e a ontologia para se chegar ao resultado final (ver modelo em SWRL no Apêndice C). Esta comunicação entre os dois

¹⁴SWRL: Semantic Web Rule Language.

componentes do sistema precisa ocorrer para cada segmento da fórmula somente para transferir a execução dos cálculos para a ontologia com auxílio de um *reasoner*.

Além dessa dificuldade, foi possível notar também, o problema que a ontologia teria em executar uma expressão como a existente na fórmula do índice do Canadá. Nesta fórmula está presente a função exponencial e^{**} .

As experimentações até o momento com SWRL foram eficientes com as operações básicas, deixando aqui uma questão em relação ao atendimento desta linguagem semântica para a execução de operações complexas.

Outro caso é o da representação de formas de derivação que de acordo com a solução adotada ficaram limitadas aos operadores aceitos pela linguagem SPARQL, i.e., MAX, MIN, AVG, COUNT e SUM.

Outra questão interessante referente à modelagem de padrões, é a necessidade de se especificar limites abertos em intervalos de valores. Alguns índices de poluição do ar contém a categoria mais crítica com o limite superior em aberto. Isto implica na proposição de um recurso de modelagem, de modo a tratar este requerimento. A solução adotada em ambos os estudos de caso foi de atribuir um valor alto o suficiente, naturalmente baseado nas especificações dos índices de qualidade do ar, para o limite superior da camada mais crítica. Entretanto, uma outra solução poderia ter sido adotada.

Outra situação é vista no momento que se quer definir uma relação binária na qual uma das cardinalidades pode assumir um valor entre [0..1]. A ferramenta utilizada não prevê tal valor para atribuição nas cardinalidades das propriedades. Isto acabou implicando na utilização de um contorno pela especialização da classe de categorias dos indicadores em duas subclasses.

9.2 TRABALHOS FUTUROS

Os cenários utilizados como requerimentos da modelagem dos padrões tinham em comum a relação unária entre indicadores e poluentes. Então, uma questão a ser explorada, seria a modelagem de um padrão para indicadores cuja relação com os poluentes seja n -ária.

Outra questão que precisa de refinamento diz respeito à representação de fórmulas matemáticas em ontologias. As pesquisas nos portais¹⁵ sobre padronização de modelagem não apresentaram nenhuma indicação de que uma solução esteja a caminho.

¹⁵Portais pesquisados: ontologydesignpatterns.org, odps.sourceforge.net e www.w3.org.

Também sobre a representação de fórmulas, permanece a dúvida sobre a abrangência da linguagem SWRL para realizar operações mais complexas como já foi comentado na seção anterior.

Sobre os intervalos de valores que precisem indicar o valor ∞ (infinito), como representar esta grandeza é outra questão a ser abordada. Uma possibilidade é definir subclasses de intervalos: intervalo de $-\infty$ até um limite superior, entre um limite inferior e superior (ambos diferentes de infinito) e um de limite inferior até $+\infty$.

Em relação ao armazenamento de grandes volumes de dados, não foi explorado o aspecto sobre esta capacidade. Supondo que houvesse uma funcionalidade da aplicação para manter comunicação com sensores de coleta e, cada medição fosse imediatamente carregada como uma nova instância no modelo, pergunta-se: qual seria o limite desse armazenamento? Qual dos componentes seria o delimitador para este armazenamento?

Outra questão relativa ao armazenamento se refere ao tratamento dos dados históricos, qual a delimitação para tratar os dados como antigos? E como endereçá-los quando for preciso? A adição de outra ontologia ao sistema seria uma solução adequada?

Estas reflexões decorrentes desta pesquisa deixadas nesta seção e outras que devem surgir com a sua consulta e avaliação, poderão servir como motivação para novos trabalhos no campo da modelagem semântica, em especial, na modelagem para descrição de índices e indicadores da poluição do ar. Portanto, espera-se que o conteúdo aqui apresentado possa contribuir para o enriquecimento desta área de pesquisa abrindo novas possibilidades para o aparecimento de conteúdos ainda mais interessantes.

REFERÊNCIAS

- ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, v. 32, p. 7-20, 2003.
- APACHE JENA. Disponível em: <https://jena.apache.org/>. Acesso em: 04 out. 2018.
- AQHI. Environment Canada. Disponível em: <https://www.canada.ca/en/environment-climate-change/services/air-quality-health-index.html>. Acesso em: 04 out. 2018.
- BREWSTER, C. et al. Data driven ontology evaluation. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 2004, Lisboa. **Anais...** Lisboa, 2004.
- DAHLEH, D.; FOX, M. S. **An Environmental Ontology for Global City Indicators (ISO 37120)**. 2016.
- ENVIRONMENTAL PROTECTION AGENCY. Air Now. Air Quality Index Basics. Disponível em: <http://airnow.gov/index.cfm?action=aqibasics.aqi>. Acesso em: 04 out. 2018.
- FENSEL, D. et al. OIL: An ontology infrastructure for the semantic web IEEE intelligent systems, **IEEE Intelligent Systems**, v. 16, n. 2, p. 38-45, 2001.
- FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A.; SUÁREZ-FIGUEROA, M. C. Methodological guidelines for reusing general ontologies. **Data & Knowledge Engineering**, v. 86, p. 242-275, 2013.
- FERNÁNDEZ-LÓPEZ, M.; JURISTO, N. Methontology: from ontological art towards ontological engineering. **AAAI Technical Report**, 1997.
- FOX, M. S. **An education ontology for global city indicators (ISO 37120)**. 2014.
- FOX, M. S. The role of ontologies in publishing and analyzing city indicators **Computers, Environment and Urban Systems**, v. 54, p. 266-279, 2015.
- GANGEMI, A. Ontology design patterns for semantic web content. In: INTERNATIONAL SEMANTIC WEB CONFERENCE, 4., 2005, Galway, Ireland. **Anais...** Galway, 2005. p. 262-276.
- GANGEMI, A. **What is a pattern?** 2010. Disponível em: <http://ontologydesignpatterns.org/wiki/Odp:WhatIsAPattern>. Acesso em: 04 out. 2018.
- GANGEMI, A.; PRESUTTI, V. Ontology Design Patterns. In: STAAB, S.; STUDER, R. (Eds.). **Handbook on Ontologies**. Berlin: Springer, 2009. p. 221-243.
- GHAHREMANLOO, L.; THOM, J. A.; MAGEE, L. An Ontology Derived from Heterogeneous Sustainability Indicator Set Documents. In: PROCEEDINGS OF THE

SEVENTEENTH AUSTRALASIAN DOCUMENT COMPUTING SYMPOSIUM, 2012, New Zealand. **Anais...** New Zealand, 2012. p. 72-79.

GITHUB. Disponível em: <https://github.com/>. Acesso em: 04 out. 2018.

GÓMEZ-PÉREZ, A. Towards a framework to verify knowledge sharing technology. **Expert Systems with applications**, v. 11, p. 519-529, 1996.

GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing? **International Journal of Human-Computer Studies**, v. 43, n. 5-6, p. 907-928, 1995.

GRÜNINGER, M.; FOX, M. S. The role of competency questions in enterprise engineering. In: ROLSTADÁS, A. (Ed.). **Benchmarking – Theory and practice**. Berlin: Springer, 1995. p. 22-31.

GUARINO, N.; GIARETTA, P. Ontologies and Knowledge Bases: Towards a Terminological Clarification. In: MARS, N. J. I. **Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing**. Amsterdam: IOS Press, 1995. p. 25-32.

GUIZZARDI, G. The Role of Foundational Ontologies for Conceptual Modeling and Domain Ontology Representation. **IEEE**, 2006.

HOORNWEG, D. et al. **City indicators: now to Nanjing**. 2007.

HORRIDGE, M. et al. The Manchester OWL syntax. **OWLed**, v. 216, 2006.

HORROCKS, I.; PATEL-SCHNEIDER, P. F.; VAN HARMELEN, F. Reviewing the design of DAML+OIL: an ontology language for the semantic web. **AAAI/IAAI**, p. 792-797, 2002.

IMECA: Índice Metropolitano de la Calidad del Aire. Disponível em: <<https://www.gob.mx/comisionambiental/es/articulos/imeca-indice-metropolitano-de-la-calidad-del-aire?idiom=es>>. Acesso em: 04 out. 2018.

INSTITUTO AMBIENTAL DO PARANÁ. **Indicadores da Qualidade do Ar**. Disponível em: <<http://www.iap.pr.gov.br/modules/conteudo/conteudo.php?conteudo=59>>. Acesso em: 04 out. 2018.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 704:2009**. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:704:ed-3:v1:en>>. Acesso em: 03 out. 2018.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 37120:2018**. Sustainable cities and communities - Indicators for city services and quality of life. 2nd ed. Genebra: ISO, 2018.

KEET, C. M. The Use of Foundational Ontologies in Ontology Development: An Empirical Assessment. In: **ESWC 2011: THE SEMANTIC WEB: RESEARCH AND APPLICATIONS**, 8., 2011. **Proceedings...** Berlin: Springer, 2011. p. 321-335.

KENT, R. E. The information flow foundation for conceptual knowledge organization. **Advances in Knowledge Organization**, v. 7, p. 111-117, 2000.

KRÖTZSCH, M. OWL 2 Profiles: An Introduction to Lightweight Ontology Languages. In: INTERNATIONAL SUMMER SCHOOL 2012: REASONING WEB. SEMANTIC TECHNOLOGIES FOR ADVANCED QUERY ANSWERING, 8., 2012, Vienna. **Proceedings...** Vienna: Springer, 2012. p. 112-183.

KROTZSCH, M.; SIMANCIK, F.; HORROCKS, I. A Description Logic Primer. **CoRR**, 2012.

MAEDCHE, A.; STAAB, S. Measuring similarity between ontologies. In: EKAW 2002: KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT: ONTOLOGIES AND THE SEMANTIC WEB, 13., 2002, Sigüenza. **Proceedings...** Sigüenza: Springer, 2002. p. 251-263.

MÉXICO. Gobierno del Distrito Federal. Secretaría del Medio Ambiente. **Norma ambiental para el Distrito Federal, NADF-009-AIRE-2006**. [S.l.]: Gaceta Oficial del Distrito Federal, 2006.

MUSEN, M. The protégé project: a look back and a look forward. **AI Matters**, v. 1, n. 4, p. 4-12, 2015.

NOY, N. F.; MCGUINNESS, D. L. **Ontology Development 101: A Guide to Creating Your First Ontology**. 2001. Disponível em: <https://protege.stanford.edu/publications/ontology_development/ontology101.pdf>. Acesso em: 04 out. 2018.

ONTOLOGY DESIGN PATTERNS. Disponível em: http://ontologydesignpatterns.org/wiki/Main_Page. Acesso em: 04 out. 2018.

OPREA, M. M. AIR_POLLUTION_Onto: an Ontology for Air Pollution Analysis and Control. In: ILIADIS et al. Artificial Intelligence Applications and Innovations III. IFIPAICT. Boston: Springer, 2009. v. 296.

PALMER S. R.; FELSING, M. **A practical guide to feature-driven development**. Upper Saddle River: Pearson Education, 2001.

PANETTO, H. **Ontology Engineering**. 2016. University of Lorraine, TELECOM Nancy, France. 247 slides.

PORZEL, R.; MALAKA, R. A task-based approach for ontology evaluation. In: ECAI Workshop on Ontology Learning and Population, 2004, Valencia. **Proceedings...** Valencia, 2004. p. 1-6.

REN, Y. et al. Towards Competency Question-driven Ontology Authoring. In: ESWC 2014: THE SEMANTIC WEB: TRENDS AND CHALLENGES, 11., 2014, Crete. **Proceedings...** Crete: Springer, 2014. p. 752-767.

SICHE, R. et al. Índices versus indicadores: precisões conceituais na discussão da sustentabilidade de países. **Ambiente & Sociedade**, v. 10, n. 2, p. 137-148, 2007. Disponível em: <<https://dx.doi.org/10.1590/S1414-753X2007000200009>>. Acesso em: 4 out. 2018.

SILVA, A. Sistemas inteligentes. Disponível em: <http://www.inf.ufsc.br/~alexandre.goncalves.silva/courses/14s2/ine5633/slides/aula1022.pdf>. Acesso em: 04 out. 2018.

W3C. SPARQL Query Language for RDF. Disponível : <<https://www.w3.org/TR/rdf-sparql-query/>>. Acesso em: 04 out. 2018.

STIEB, D. M. et al. A New Multipollutant, No-Threshold Air Quality Health Index Based on Short-Term Associations Observed in Daily Time-Series Analyses. **Journal of the Air & Waste Management Association**, v. 58, n. 3, p. 435-450, 2008.

UK AIR. **Daily Air Quality Index**. Disponível em: <https://uk-air.defra.gov.uk/air-pollution/daqi>. Acesso em: 04 out. 2018.

VILLAZÓN-TERRAZAS, B. et al. Methodological Guidelines for Publishing Government Linked Data. In: WOOD, D. (ed.). **Linking Government Data**. New York: Springer, 2011. p. 27-49.

WURVOC. Disponível em: <<http://www.wurvoc.org/>>. Acesso em: 04 out. 2018.

YU, J.; THOM, J. A.; TAM, A. Requirements-oriented methodology for evaluating ontologies. **Information Systems**, v. 34, p. 766-791, 2009.

APÊNDICE A – CONSULTAS EM LINGUAGEM SPARQL

Consultas realizadas pela ferramenta Protegé, utilizando o *reasoner* Hermit versão 1.3.8.413 nas ontologias AQ e BQ. Linguagens utilizadas: DL Query e SPARQL.

----- PREFIX para SPARQL -----

```

PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX aq: <http://www.semanticweb.org/ricardo/ontologies/2017/5/AirQualityOnto.owl#>
PREFIX prov: <http://www.w3.org/ns/prov#>
PREFIX om-11: <http://www.wurvoc.org/vocabularies/om-1.6/>

```

AQO3.owl (AQ)

(Q1) Quais as medições do poluente <P> coletados pela estação <S> de responsabilidade do agente <A> que ocorreram no período <T>?

```

SELECT ?station ?datahora ?poluente ?valor ?agente
WHERE { ?m aq:hasHappenedOn ?datahora .
       ?m aq:takesPlaceAt ?station .
       ?m aq:wasAttributedTo ?agente .
       ?m aq:hasMeasurementOf ?poluente .
       ?m om-11:numerical_value ?valor .
FILTER ( ?poluente = aq:O3 &&
         ?station = aq:ASSIS &&
         ?agente = aq:IAP_InstAmbParana &&
         ?datahora >= '2017-06-17T08:00:00'^^xsd:dateTime &&
         ?datahora <= '2017-06-17T18:00:00'^^xsd:dateTime ) }
GROUP BY ?station ?poluente ?datahora ?valor
ORDER BY ?poluente ?valor

```

?station	?datahora	?poluente	?valor	?agente
aq:ASSIS	2017-06-17T09:00:00^^xsd:dateTime	aq:O3	17	
aq:ASSIS	2017-06-17T18:00:00^^xsd:dateTime	aq:O3	46	
	aq:IAP_InstAmbParana			

(Q1_variante) somente filtrando o periodo

```

SELECT ?station ?datahora ?poluente ?valor ?agente
WHERE { ?m aq:hasHappenedOn ?datahora .
       ?m aq:takesPlaceAt ?station .
       ?m aq:wasAttributedTo ?agente .
       ?m aq:hasMeasurementOf ?poluente .
       ?m om-11:numerical_value ?valor .
FILTER ( ?datahora >= '2017-06-17T08:00:00'^^xsd:dateTime &&
         ?datahora <= '2017-06-17T18:00:00'^^xsd:dateTime ) }
GROUP BY ?station ?poluente ?datahora ?valor
ORDER BY ?poluente ?valor

```

?station	?datahora	?poluente	?valor	?agente
aq:ASSIS	2017-06-17T09:00:00^^xsd:dateTime	aq:PM10	31	
	aq:IAP_InstAmbParana			

aq:ASSIS 2017-06-17T10:00:00^^xsd:dateTime aq:PM10 65
aq:IAP_InstAmbParana

(Q2) Quais indicadores fazem referência ao poluente <P>?

DLQUERY: isIndicatorFor value PM10
results: UI_4_IMECA
file : AQO3_Q2.jpg

(Q3) Como o indicador <R> é calculado a partir das medições?

```
SELECT ?m ?pol ?derivform
WHERE { ?m aq:isIndicatorFor ?pol .
        ?m aq:isDerivedBy ?derivform .
FILTER ( ?m = aq:UI_4_IMECA ) }
```

?m	?pol	?derivform
aq:UI_4_IMECA	aq:PM10	aq:average

(Q4) Como o indicador <R> é categorizado?

```
SELECT ?ui ?categ
WHERE { ?ui aq:isCategorizedBy ?categ .
FILTER ( ?ui = aq:UI_4_IMECA ) }
```

?ui	?categ
aq:UI_4_IMECA	aq:PM10_IMECA_good
aq:UI_4_IMECA	aq:PM10_IMECA_regular

(Q5) Qual a categorização para o indicador <R> na periodicidade <H>?

```
SELECT ?ui ?categ ?pol ?h
WHERE { ?ui aq:isCategorizedBy ?categ .
        ?ui aq:isIndicatorFor ?pol .
        ?ui aq:hasPeriodicity ?period .
        ?period aq:hours ?h .
FILTER ( ?ui = aq:UI_4_IMECA &&
        ?h = "24"^^xsd:decimal ) }
```

?ui	?categ	?pol	?h
aq:UI_4_IMECA	aq:PM10_IMECA_good	aq:PM10	24
aq:UI_4_IMECA	aq:PM10_IMECA_regular	aq:PM10	24

(Q6) Qual o valor para o indicador <R> no período <T> com a forma de derivação <D>?
D='Average'

```
SELECT ?ui (AVG (?valor) AS ?media) ?df ?pol
WHERE { ?ui aq:isIndicatorFor ?pol .
        ?ui aq:isDerivedBy ?df .
        ?m aq:hasMeasurementOf ?pol .
        ?m aq:hasHappenedOn ?datahora .
        ?m om-11:numerical_value ?valor .
```

```

FILTER ( ?ui = aq:UI_4_IMECA &&
        ?datahora >= '2017-06-17T08:00:00'^^xsd:dateTime &&
        ?datahora <= '2017-06-17T12:00:00'^^xsd:dateTime ) }
GROUP BY ?ui ?df ?pol

```

```

?ui          ?media ?df          ?pol
aq:UI_4_IMECA      65.0    aq:Maximum    aq:PM10

```

(Q7) Qual é a categoria para o valor <V> do indicador <R>?
V=40

```

SELECT ?categ ?min ?max
WHERE { ?ui aq:isCategorizedBy ?categ .
        ?categ rdf:type aq:IndicatorCategorization .
        ?categ aq:hasLowerLimit ?min .
        ?categ aq:hasUpperLimit ?max .
FILTER ( ?ui = aq:UI_4_IMECA &&
        ?min <= '40'^^xsd:decimal &&
        ?max >= '40'^^xsd:decimal ) }

```

```

?categ          ?min    ?max
aq:PM10_IMECA_good 0      60

```

(Q8) Qual é a categoria para o valor <V> do índice <I>?
V=40

```

SELECT ?categ ?min ?max
WHERE { ?ix aq:isCategorizedBy ?categ .
        ?categ rdf:type aq:IndexCategorization .
        ?categ aq:hasLowerLimit ?min .
        ?categ aq:hasUpperLimit ?max .
FILTER ( ?ix = aq:IMECA_Mexico &&
        ?min <= '40'^^xsd:decimal &&
        ?max >= '40'^^xsd:decimal ) }

```

```

?categ          ?min    ?max
aq:IMECA_good      0      50

```

(Q9) Quais são as categorias definidas para o índice <I>?

```

SELECT ?ix ?categ
WHERE { ?ix aq:isCategorizedBy ?categ .
FILTER ( ?ix = aq:IMECA_Mexico ) }

```

```

?ix          ?categ
aq:IMECA_Mexico    aq:IMECA_good
aq:IMECA_Mexico    aq:IMECA_regular
aq:IMECA_Mexico    aq:IMECA_critical

```

(Q10) Qual a fórmula de cálculo a ser utilizada para uma categoria de índice <CI> e uma categoria de indicador <CR>?

```

SELECT ?categ ?formula
WHERE { ?ui2ix rdf:type      aq:IndicatorToIndexFormula .
        ?ui2ix aq:hasFormula ?formula .
        ?ui2ix aq:makesRefTo ?categ .
FILTER ( ?categ = aq:PM10_IMECA_good ||
        ?categ = aq:IMECA_regular ) }

```

```

?categ      ?formula
aq:PM10_IMECA_good C_PM10 * ( 5 / 6 )^^xsd:string
aq:IMECA_regular   C_PM10 * ( 5 / 6 )^^xsd:string

```

(Q10_variante)

```

SELECT ?ui2ix ?categ ?formula
WHERE { ?ui2ix rdf:type      aq:IndicatorToIndexFormula .
        ?ui2ix aq:hasFormula ?formula .
        ?ui2ix aq:makesRefTo ?categ .

```

```

FILTER ( ?ui2ix = aq:UI_4_IMECA_to_IMECA_Mexico ) }

```

```

?ui2ix      ?categ      ?formula
aq:UI_4_IMECA_to_IMECA_Mexico  aq:PM10_IMECA_good C_PM10 * ( 5 / 6 )^^xsd:string
aq:UI_4_IMECA_to_IMECA_Mexico  aq:IMECA_regular   C_PM10 * ( 5 / 6 )^^xsd:string

```

(Q11) Qual a fórmula de cálculo a ser utilizada para uma categoria de índice <CI>?

```

SELECT ?categ ?formula
WHERE { ?ui2ix rdf:type      aq:IndicatorToIndexFormula .
        ?ui2ix aq:hasFormula ?formula .
        ?ui2ix aq:makesRefTo ?categ .
FILTER ( ?categ = aq:IMECA_regular ) }

```

```

?categ      ?formula
aq:IMECA_regular   C_PM10 * ( 5 / 6 )^^xsd:string

```

(Q12) Qual a unidade de medida <U> utilizada na medição do poluente <P> no período <T>?

```

SELECT ?poluente ?unidade
WHERE { ?m aq:hasHappenedOn ?datahora .
        ?m aq:hasMeasurementOf ?poluente .
        ?m om-11:numerical_value ?valor .
        ?m om-11:unit_of_measure_or_measurement_scale ?unidade .
FILTER ( ?poluente = aq:O3      &&
        ?datahora = '2017-06-17T09:00:00'^^xsd:dateTime ) }
GROUP BY ?poluente ?unidade

```

```

?poluente      ?unidade
aq:O3          aq:microgram_per_cubic_metre

```

(Q13) Qual é a forma de derivação <D> do índice <I>?

```

SELECT ?ix ?df
WHERE { ?ix aq:isDerivedBy ?df .
FILTER ( ?ix = aq:IMECA_Mexico ) }

```

```
?ix          ?df
aq:IMECA_Mexico    aq:Average
```

(Q14) Qual a periodicidade <H> e o poluente <P> calculados pelo indicador <R>?

```
SELECT ?ui ?pol ?period
WHERE { ?ui aq:isIndicatorFor ?pol .
        ?ui aq:hasPeriodicity ?dd .
        ?dd aq:hours ?period .
FILTER ( ?ui = aq:UI_4_IMECA ) }
```

```
?ui          ?pol          ?period
aq:UI_4_IMECA    aq:PM10          24
```

----- FIM AQ -----

----- PREFIX para SPARQL -----

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX bq: <http://www.semanticweb.org/ricardo/ontologies/2017/5/AirQualityOnto_B.owl#>
PREFIX prov: <http://www.w3.org/ns/prov#>
PREFIX om-11: <http://www.wurvoc.org/vocabularies/om-1.6/>
```

----- BQO3.owl (BQ) -----

(Q1) Quais são as medições do poluente <P> coletados pela estação <S> de responsabilidade do agente <A> que ocorreram no período <T>?

```
SELECT ?station ?datahora ?poluente ?valor ?agente
WHERE { ?m bq:hasHappenedOn ?datahora .
        ?m bq:takesPlaceAt ?station .
        ?m prov:wasAttributedTo ?agente .
        ?m bq:hasMeasurementOf ?poluente .
        ?m om-11:numerical_value ?valor .
FILTER ( ?poluente = bq:NO2 &&
        ?station = bq:TOR_DWNT &&
        ?agente = bq:EnvironmentCanada &&
        ?datahora >= '2018-01-01T08:00:00'^^xsd:dateTime &&
        ?datahora <= '2018-01-01T10:00:00'^^xsd:dateTime ) }
GROUP BY ?station ?poluente ?datahora ?valor ?agente
ORDER BY ?poluente ?valor
```

```
?station ?datahora          ?poluente    ?valor  ?agente
bq:TOR_DWNT2018-01-01T08:00:00^^xsd:dateTime    bq:NO2    18
    bq:EnvironmentCanada
bq:TOR_DWNT2018-01-01T09:00:00^^xsd:dateTime    bq:NO2    20
    bq:EnvironmentCanada
```

----- (Q2) Quais indicadores fazem referência ao poluente <P>? -----

```
DLQUERY: isIndicatorFor value O3
results: UI_AQHI_O3_TOR
```

file : BQO3_Q2.jpg

 (Q3) Como o indicador <R> é calculado a partir das medições?

```
SELECT ?m ?pol ?derivform
WHERE { ?m bq:isIndicatorFor ?pol .
        ?m bq:isDerivedBy ?derivform .
FILTER ( ?m = bq:UI_AQHI_O3_TOR ) }
```

```
?m ?pol ?derivform
bq:UI_AQHI_O3_TOR bq:O3 bq:Average
```

 (Q4) Como o indicador <R> é categorizado?

```
SELECT ?ui ?categ
WHERE { ?ui bq:isCategorizedBy ?categ .
FILTER ( ?ui = bq:UI_4_IMECA ) }
```

```
?ui ?categ
bq:UI_4_IMECA bq:PM10_IMECA_good
```

```
SELECT ?ui ?categ
WHERE { ?ui bq:isCategorizedBy ?categ .
FILTER ( ?ui = bq:UI_AQHI_NO2_TOR ) }
```

```
?ui ?categ
(0)
```

 (Q5) Qual a categorização para o indicador <R> na periodicidade <H>?

```
SELECT ?ui ?categ ?pol ?h
WHERE { ?ui bq:isCategorizedBy ?categ .
        ?ui bq:isIndicatorFor ?pol .
        ?ui bq:hasPeriodicity ?period .
        ?period bq:hours ?h .
FILTER ( ?ui = bq:UI_4_IMECA &&
        ?h = "24"^^xsd:decimal ) }
```

```
?ui ?pol ?h ?categ
bq:UI_4_IMECA bq:PM10 24 bq:PM10_IMECA_good
```

ou (Indicador NonCategorized_UI)

```
SELECT ?ui ?categ ?pol ?h
WHERE { ?ui bq:isCategorizedBy ?categ .
        ?ui bq:isIndicatorFor ?pol .
        ?ui bq:hasPeriodicity ?period .
        ?period bq:hours ?h .
FILTER ( ?ui = bq:UI_AQHI_PM25_TOR &&
        ?h = "3"^^xsd:decimal ) }
```

```
?ui ?categ ?pol ?h
(0)
```

(Q6) Qual o valor para o indicador <R> no período <T> com a forma de derivação <D>?

```
SELECT ?ui (AVG (?valor) AS ?media) ?df ?pol
WHERE { ?ui bq:isIndicatorFor ?pol .
        ?ui bq:isDerivedBy ?df .
        ?m bq:hasMeasurementOf ?pol .
        ?m bq:hasHappenedOn ?datahora .
        ?m om-11:numerical_value ?valor .
FILTER ( ?ui = bq:UI_AQHI_NO2_TOR &&
        ?datahora >= '2017-06-17T08:00:00'^^xsd:dateTime &&
        ?datahora <= '2017-06-17T12:00:00'^^xsd:dateTime ) }
GROUP BY ?ui ?df ?pol
```

```
?ui ?media ?df ?pol
bq:UI_AQHI_NO2_TOR 19.0 bq:Average bq:NO2
```

(Q7) Qual é a categoria para o valor <V> do indicador <R>?

```
SELECT ?categ ?min ?max
WHERE { ?ui bq:isCategorizedBy ?categ .
        ?categ rdf:type bq:IndicatorCategorization .
        ?categ bq:hasLowerLimit ?min .
        ?categ bq:hasUpperLimit ?max .
FILTER ( ?ui = bq:UI_4_IMECA &&
        ?min <= '41'^^xsd:decimal &&
        ?max >= '41'^^xsd:decimal ) }
```

```
?categ ?min ?max
bq:PM10_IMECA_good 0 60
```

ou

```
SELECT ?categ ?min ?max
WHERE { ?ui bq:isCategorizedBy ?categ .
        ?categ rdf:type bq:IndicatorCategorization .
        ?categ bq:hasLowerLimit ?min .
        ?categ bq:hasUpperLimit ?max .
FILTER ( ?ui = bq:UI_AQHI_O3_TOR &&
        ?min <= '17'^^xsd:decimal &&
        ?max >= '17'^^xsd:decimal ) }
```

```
?categ ?min ?max
(0)
```

(Q8) Qual é a categoria para o valor <V> do índice <I>?

```
SELECT ?categ ?min ?max
WHERE { ?ix bq:isCategorizedBy ?categ .
        ?categ rdf:type bq:IndexCategorization .
        ?categ bq:hasLowerLimit ?min .
        ?categ bq:hasUpperLimit ?max .
FILTER ( ?ix = bq:AQHI_Canada &&
        ?min <= '3.4'^^xsd:decimal &&
        ?max >= '3.4'^^xsd:decimal ) }
```

?categ	?min	?max
bq:LowRisk	1	3.99

(Q9) Quais são as categorias definidas para o índice <I>?

```
SELECT ?ix ?categ
WHERE { ?ix bq:isCategorizedBy ?categ .
FILTER ( ?ix = bq:AQHI_Canada ) }
```

?ix	?categ
bq:AQHI_Canada	bq:LowRisk
bq:AQHI_Canada	bq:ModerateRisk
bq:AQHI_Canada	bq:HighRisk

(Q10) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI> e uma categoria do indicador <CR>?

```
SELECT ?categ ?formula
WHERE { ?ui2ix rdf:type bq:IndicatorToIndexFormula .
?ui2ix bq:hasFormula ?formula .
?ui2ix bq:makesRefTo ?categ .
FILTER ( ?categ = bq:PM10_IMECA_good ||
?categ = bq:IMECA_regular ) }
```

?categ	?formula
bq:IMECA_regular	C_PM10 * (5 / 6)^^xsd:string
bq:PM10_IMECA_good	C_PM10 * (5 / 6)^^xsd:string

(Q11) Qual a fórmula de cálculo a ser utilizada para uma categoria do índice <CI>?

```
SELECT ?ui2ix ?categ ?formula
WHERE { ?ui2ix rdf:type bq:IndexToFormula .
?ui2ix bq:hasFormula ?formula .
?ui2ix bq:makesRefTo ?categ .
FILTER ( ?categ = bq:LowRisk ) }
```

?ui2ix	?categ
bq:I2F_AQHI_TOR	bq:LowRisk

```
?ix2f
{(1000/10.4)*{(e**(0.000537xO3)-1) + (e**(0.000871xNO2)-1) + (e**(0.000487xPM2.5)-1)}}^^xsd:string
```

ou

```
SELECT ?ix ?categ ?ix2f ?formula
WHERE { ?ix bq:isCategorizedBy ?categ .
?ix2f bq:makesRefTo ?categ .
?ix2f rdf:type bq:IndexToFormula .
?ix2f bq:hasFormula ?formula .
?ix2f bq:makesRefTo ?categ .
FILTER ( ?ix = bq:AQHI_Canada ) }
```

?ix	?categ	?ix2f
-----	--------	-------

bq:AQHI_Canada bq:LowRisk bq:I2F_AQHI_TOR

?formula

{(1000/10.4)*{(e**(0.000537x03)-1) + (e**(0.000871xNO2)-1) + (e**(0.000487xPM2.5)-1)}^^xsd:string

 (Q12) Qual a unidade de medida <U> utilizada na medição do poluente <P>?

```
SELECT ?poluente ?unidade
WHERE { ?m bq:hasMeasurementOf ?poluente .
       ?m om-11:numerical_value ?valor .
       ?m om-11:unit_of_measure_or_measurement_scale ?unidade .
FILTER ( ?poluente = bq:PM2.5 ) }
```

?poluente ?unidade
 bq:PM2.5 bq:microgram_per_cubic_metre

 (Q13) Qual é a forma de derivação <D> do índice <I>?

```
SELECT ?ix ?df
WHERE { ?ix bq:isDerivedBy ?df .
FILTER ( ?ix = bq:IMECA_Mexico ) }
```

?ix ?df
 bq:IMECA_Mexico bq:Average

 (Q14) Qual a periodicidade <H> e o poluente <P> calculados pelo indicador <R>?

```
SELECT ?ui ?pol ?period
WHERE { ?ui bq:isIndicatorFor ?pol .
       ?ui bq:hasPeriodicity ?dd .
       ?dd bq:hours ?period .
FILTER ( ?ui = bq:UI_4_IMECA ) }
```

?ui ?pol ?period
 bq:UI_4_IMECA bq:PM10 24

----- FIM BQ -----

APÊNDICE B – ESTATÍSTICAS DAS ONTOLOGIAS

Estatísticas extraídas da ferramenta Protégé, pela consulta à função ‘ontology metrics’ que disponibiliza uma contagem dos elementos presentes na ontologia divididos em grupos que dependem da natureza dos elementos.

Ontologia AQ: AirQualityOnto – Modelagem “A”		
Métricas	Qtde	Observação
Quantidade de classes	759	Incluindo classes de ontologias importadas
Quantidade de classes importadas e reutilizadas	8	Subclasses não foram computadas
Quantidade de propriedades de objetos (OP)	142	Incluindo propriedades de ontologias importadas
Quantidade de propriedades de dados (DP)	35	Idem
Quantidade de indivíduos	136	Criados para o estudo de caso
Quantidade total de elementos	1072	Classes, OPs, DPs e Indivíduos

Ontologia BQ: AirQualityOnto_B – Modelagem “B”		
Métricas	Qtde	Observação
Quantidade de classes	711	Incluindo classes de ontologias importadas
Quantidade de classes importadas e reutilizadas	8	Subclasses não foram computadas
Quantidade de propriedades de objetos (OP)	139	Incluindo propriedades de ontologias importadas
Quantidade de propriedades de dados (DP)	34	Idem
Quantidade de indivíduos	136	Criados para o estudo de caso
Quantidade total de elementos	1020	Classes, OPs, DPs e Indivíduos

APÊNDICE C – CONTEÚDO DA ONTOLOGIA SWRL

Ontologia desenvolvida para uso da linguagem SWRL para executar operações matemáticas. Conteúdo descrito no formato Manchester Syntax.

Prefix: : <<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#>>

Prefix: owl: <<http://www.w3.org/2002/07/owl#>>

Prefix: rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

Prefix: rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

Prefix: swrl: <<http://www.w3.org/2003/11/swrl#>>

Prefix: swrla: <<http://swrl.stanford.edu/ontologies/3.3/swrla.owl#>>

Prefix: swrlb: <<http://www.w3.org/2003/11/swrlb#>>

Prefix: xml: <<http://www.w3.org/XML/1998/namespace>>

Prefix: xsd: <<http://www.w3.org/2001/XMLSchema#>>

Ontology: <<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23>>

Annotations:

rdfs:comment "Exemplo de uso com SWRL"

Com operações de soma e multiplicação.

Autor: Ricardo Martins Taques"

AnnotationProperty: rdfs:comment

AnnotationProperty: rdfs:label

AnnotationProperty: swrla:isRuleEnabled

Datatype: xsd:boolean

Datatype: xsd:integer

Datatype: xsd:string

ObjectProperty: temSys1

Domain:

InterSystem,
temSys1 exactly 1 Sys1,
temSys2 exactly 1 Sys2

Range:

Sys1,
System

ObjectProperty: temSys2

Domain:

InterSystem,
temSys1 exactly 1 Sys1,
temSys2 exactly 1 Sys2

Range:

Sys2,
System

DataProperty: temCompleto

DataProperty: temOmega

DataProperty: temResultDeMultiplicacao

DataProperty: temResultDeSoma

DataProperty: temValor

Class: InterSystem

EquivalentTo:

temSys1 exactly 1 Sys1,

temSys2 exactly 1 Sys2

Class: Multiplicar

SubClassOf:

Operacao

Class: Operacao

Class: Somar

SubClassOf:

Operacao

Class: Sys1

SubClassOf:

System

Class: Sys2

SubClassOf:

System

Class: System

Individual: intersystem

Types:

InterSystem,

temSys1 exactly 1 Sys1,

temSys2 exactly 1 Sys2

Facts:

temSys1 s1

Individual: operando1

Types:

Multiplicar

Facts:

temValor 51

Individual: operando2

Types:
Somar

Facts:
temValor 1

Individual: s1

Types:
Sys1,
System

Facts:
temCompleto 42

Individual: s2

Types:
Sys2,
System

Rule:

```
InterSystem(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#i>),
Sys1(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#x1>),
temSys1(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#i>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#x1>),
temCompleto(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#x1>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#complete>),
swrlb:multiply(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#complete>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#omega>, 1) ->
temOmega(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#i>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#omega>)
```

Rule:

```
temValor(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#sm1>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#valors>),
Somar(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#sm1>),
swrlb:add(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#result>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#valorm>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#valors>),
temValor(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#m1>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#valorm>),
Operacao(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#op>),
Multiplicar(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#m1>) ->
temResultDeSoma(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#op>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#result>)
```

Rule:

```
swrlb:multiply(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#result>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#valor>, 2),
temValor(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#m1>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#valor>),
Operacao(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#op>),
Multiplicar(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#m1>) ->
temResultDeMultiplicacao(?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#op>,
?<http://www.semanticweb.org/ricardo/ontologies/2017/6/untitled-ontology-23#result>)
```