

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Vinícius Gama Valory Frauches

**Uma Abordagem Baseada em Ontologias para
Obtenção de Indicadores a partir de Dados
Abertos**

VITÓRIA
2014

Vinícius Gama Valory Frauches

Uma Abordagem Baseada em Ontologias para Obtenção de Indicadores a Partir de Dados Abertos

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Orientadora: Monalessa Perini Barcellos

Coorientador: Ricardo de Almeida Falbo

VITÓRIA
2014

Vinícius Gama Valory Frauches

Uma Abordagem Baseada em Ontologias para Obtenção de Indicadores a Partir de Dados Abertos

COMISSÃO EXAMINADORA

Prof. Monalessa Perini Barcellos, D. Sc.
Universidade Federal do Espírito Santo (UFES)

Prof. Ricardo de Almeida Falbo, D. Sc.
Universidade Federal do Espírito Santo (UFES)

Prof. Vítor Estêvão Silva Souza, Ph. D.
Universidade Federal do Espírito Santo (UFES)

Profa. Andreia Malucelli, D. Sc.
Pontifícia Universidade Católica do Paraná (PUCPR)

Vitória, 28 de Agosto de 2014

Ao meu filho Diogo e minha esposa Juliana, minha motivação.

AGRADECIMENTOS

Agradeço a Deus, minha filosofia, pois é através de Sua doutrina que me mantenho resiliente e pronto .

Agradeço aos meus pais, Fátima e Ronaldo, pois foi sua educação que me fez hoje quem sou.

A todos os meus professores que proveram o conhecimento que tenho hoje e que foram imprescindíveis para chegar onde cheguei, inclusive para a conclusão deste trabalho.

À minha esposa Juliana, pelo companheirismo, apoio e compreensão durante minha ausência e trabalho.

À professora Monalessa, que mais do que professora e orientadora, foi amiga e me incentivou quando precisei.

Ao professor Falbo por suas valiosas contribuições e orientações para com este trabalho.

À minha família, pela torcida e por proverem o tempo necessário para o desenvolvimento desta dissertação.

Aos colegas Vitor, Cassio, Freddy, Laylla, Filipe e Carlos, por suas opiniões, contribuições, palavras, e principalmente pelo café.

Aos professores Vítor Souza e Andreia Malucelli por aceitarem participar da banca e por suas valiosas contribuições.

Ao corpo administrativo do PPGI pelo bom atendimento e apoio.

Muito obrigado a todos que contribuíram de alguma forma para que este momento se tornasse realidade.

RESUMO

Em uma organização diversas decisões são tomadas diariamente, desde o nível estratégico até o nível operacional. Independente do nível em que as decisões são tomadas, elas devem ser baseadas em dados que, quando analisados, fornecem informações que guiam a tomada de decisão. No entanto, apesar de as organizações coletarem e armazenarem dados para auxiliar a tomada de decisão, eles nem sempre são utilizados adequadamente. Esse problema é ampliado quando se trata de dados abertos, onde um grande volume de dados é disponibilizado, mas a identificação e acesso aos dados que, quando analisados, podem prover informações úteis e apoiar tomadas de decisão, não é trivial. Independente do domínio, muitas vezes, a obtenção de informações se dá a partir dos resultados de um processo de medição, no qual definem-se as medidas para as quais se deseja coletar dados e realizam-se a coleta e a análise dos dados. Considerando esse cenário, este trabalho propõe uma abordagem para obtenção de indicadores úteis à tomada de decisão a partir de dados abertos. Apesar de a abordagem ter sido desenvolvida considerando o uso de dados abertos, ela pode ser utilizada para outros tipos de fontes de dados. A abordagem foi definida com base em duas ontologias, também propostas neste trabalho: uma *Core Ontology on Measurement*, que apresenta a conceituação central de medição independente do domínio de aplicação, incluindo aspectos relacionados à definição de medidas de maneira alinhada a objetivos estabelecidos, à coleta de dados para as medidas e à análise dos dados coletados; e uma *Measurement Task Ontology*, que descreve o processo de medição e a conceituação a ele associada. A *Core Ontology on Measurement* foi utilizada para fornecer o conhecimento sobre medição e o vocabulário adotado na abordagem proposta. A *Measurement Task Ontology*, por sua vez, foi utilizada para nortear a definição da abordagem, orientando sobre as atividades que deveriam ser contempladas. Buscando alcançar clareza conceitual e fidelidade à realidade, as ontologias foram definidas tendo como base a ontologia de fundamentação UFO (*Unified Foundational Ontology*). Para apoiar o uso da abordagem proposta, uma aplicação foi desenvolvida e, como avaliação preliminar da abordagem, a aplicação foi utilizada para obter indicadores a partir de bases de dados abertos disponibilizadas pela ANAC (Agência Nacional de Aviação Civil) e pela INFRAERO (Empresa Brasileira de Infraestrutura Aeroportuária).

Palavras-chave: Medição, Indicador, Ontologia, Ontologia de Medição, Dados Abertos.

ABSTRACT

In an organization, several decisions are made daily, from the strategic to operational level. Regardless of the level in which decisions are made, they must be based on data that provide information to guide decision-making when are analyzed. Nevertheless, although organizations collect and store data to support decision-making, data are not always appropriately used. This problem increases when one is dealing with open data. In this context, although a large amount of data is provided, identification and access to useful data are not trivial. Regardless of the domain, information is many times obtained from the results of a measurement process, in which measures are defined, and data collection and analysis are performed. Considering that scenario, this work proposes an approach for obtaining useful indicators for decision-making from open data sources. Although the approach has been developed considering the use of open data sources, it can be applied to other types of data sources. The approach was defined based on two ontologies also proposed in this work: a *Core Ontology on Measurement*, which presents a central domain-independent conceptualization about measurement, including aspects related to measures definition aligned to established goals, data collection and analysis; and a *Measurement Task Ontology*, which describes the measurement process and its conceptualization. The *Core Ontology on Measurement* was used to provide the knowledge about measurement and the vocabulary adopted in the proposed approach. The *Measurement Task Ontology*, in turn, was used to guide the approach definition through the activities that should be addressed. Aiming to obtain conceptual clarity and fidelity to reality, the ontologies were defined based on UFO (*Unified Foundational Ontology*). In order to support the use of the proposed approach, an application was developed and, as a preliminary evaluation of the approach, it was used to obtain indicators from open data sources provided by ANAC (Civil Aviation National Agency) and INFRAERO (Aviation Infrastructure Brazilian Company).

Keywords: Measurement, Indicator, Ontology, Measurement Ontology, Open Data.

SUMÁRIO

Capítulo 1 Introdução	1
1.1 Contexto	1
1.2 Motivação	3
1.3 Objetivos da Pesquisa	6
1.4 Método de Pesquisa	6
1.5 Organização da Dissertação	7
Capítulo 2 Revisão da Literatura	9
2.1 Medição	9
2.2 Ontologias.....	16
2.2.1 Tipos de Ontologias	17
2.2.2 Unified Foundational Ontology - UFO	20
2.3 Dados Abertos	26
Capítulo 3 Ontologias de Medição	31
3.1 <i>Core Ontology on Measurement</i> - COM	31
3.1.1 As Subontologias de COM	32
3.1.1.1 Subontologia Entidades Mensuráveis (<i>Measurable Entities</i>)	33
3.1.1.2 Subontologia Medida (<i>Measure</i>)	36
3.1.1.3 Subontologia Planejamento de Medição (<i>Measurement Planning</i>)	41
3.1.1.4 Subontologia Medição (<i>Measurement</i>)	45
3.1.1.5 Subontologia Análise de Medição (<i>Measurement Analysis</i>)	48
3.2 <i>Measurement Task Ontology</i> - MTO	52
3.2.1 Modelos Comportamentais da <i>Measurement Task Ontology</i>	52
3.2.1.1 Planejar Medição (<i>Plan Measurement</i>).....	55
3.2.1.2 Executar Medição (<i>Perform Measurement</i>).....	57
3.2.1.3 Analisar Medição (<i>Analyze Measurement</i>)	58
Capítulo 4 Abordagem para Obtenção de Indicadores a partir de Dados Abertos	60
4.1 Abordagem para Obtenção de Indicadores a partir de Dados Abertos	60
4.1.1 Planejar Obtenção de Indicadores	61
4.1.2 Obter Dados.....	66
4.1.3 Analisar Dados	66
4.1.4 Relações entre a Abordagem Proposta e a <i>Measurement Task Ontology</i>	69
4.2 Apoio Computacional para Uso da Abordagem Proposta	72
4.3 Uso da Abordagem Proposta para Obter Indicadores a partir de Bases de Dados Abertos disponibilizadas pela ANAC e INFRAERO	74
Capítulo 5 Conclusão	83
5.1 Considerações Finais	83
5.2 Contribuições.....	85
5.3 Perspectivas Futuras.....	86
Referências Bibliográficas	88
Apêndice 1 Mapeamento entre COM e RSMO	95

Capítulo 1

Introdução

Este capítulo apresenta o contexto, motivação e objetivos do trabalho, bem como o método de pesquisa adotado e a organização do texto desta dissertação.

1.1 Contexto

Organizações coletam e armazenam dados que podem ser utilizados para a tomada de decisão. Esses dados podem ser armazenados em sistemas de informação e outras aplicações computacionais utilizadas pela organização. Atualmente, diversas organizações têm publicado seus dados na Internet, seja por opção ou para atender a alguma regulamentação. Como consequência, tem-se uma grande quantidade de dados disponíveis na Internet que podem ser acessados, utilizados e distribuídos livremente por qualquer pessoa ou organização. Esses dados, conhecidos como Dados Abertos (OPEN KNOWLEDGE FOUNDATION, 2014), têm sido utilizados como uma nova forma para representação da informação pública, sendo base para a tomada de decisão política, econômica e social (HOXHA; BRAHAJ, 2011).

No contexto da tomada de decisão, a medição desempenha um papel fundamental (ROBERTS, 1985), uma vez que provê informações que podem embasar tomadas de decisão nas mais diversas áreas (ALTMAN; BLAND, 1983; LAORY *et al.*, 2012; MARSHALL; JOHNELL; WEDEL, 1996). Por exemplo, pesquisadores coletam dados de seus experimentos visando analisar algum fenômeno e chegar a alguma conclusão sobre ele; agências governamentais coletam dados sobre a população, a economia e o mercado para analisarem a situação de um país; médicos coletam dados de seus pacientes visando obterem informações para determinar os tratamentos adequados; gerentes de projetos coletam dados dos seus projetos para verificar seu desempenho e aderência aos planos.

Finkelstein e Leaning (1984) definem medição como o processo de atribuição de números ou símbolos a propriedades de entidades do mundo real, de acordo com regras bem definidas, a fim de descrevê-las. Em outras palavras, medição pode ser entendida como um processo que envolve um conjunto de ações com a finalidade de caracterizar entidades a partir da atribuição de valores às suas propriedades (BARCELLOS; FALBO; FRAUCHES, 2014). Por exemplo, pode-se medir o peso em quilos e a altura em metros de uma pessoa e a partir desses valores obter seu índice de massa corporal em quilos/metros²

e analisar se ela está na faixa de peso saudável ou não, a fim de tomar decisões caso a faixa não seja saudável.

É importante ressaltar que neste trabalho o termo medição é usado para denotar não apenas o ato de medir propriamente dito (por exemplo, o ato de usar uma fita métrica para medir a altura de uma pessoa em metros). Neste trabalho, o termo medição refere-se ao processo de medição como um todo, o qual consiste, basicamente, na definição de medidas para as quais se deseja coletar dados (por exemplo, peso em quilos, altura em metros e índice de massa corporal em quilos/metros²) e dos procedimentos necessários para realizar a coleta, na coleta de dados para essas medidas (por exemplo, a atribuição dos valores 90, 1.7 e 31.14 para essas medidas quando dizem respeito a João) e na análise dos dados coletados (por exemplo, a análise dos valores medidos para João considerando com base os valores recomendados pela Organização Mundial de Saúde, chegando-se à conclusão de que João está na faixa de peso que indica obesidade de grau I).

Conforme dito anteriormente, a partir da análise dos dados é possível obter informações úteis à tomada de decisão. Considerando o exemplo apresentado, a informação de que João está com obesidade de grau I poderia ser usada para a tomada de decisão sobre as ações que João deveria realizar (dieta, exercícios, medicamentos, etc.). Para que as informações obtidas a partir da análise de dados sejam realmente úteis, os dados devem ser coletados com algum propósito, ou seja, a medição deve ser orientada a objetivos (ISO/IEC, 2007).

Um objetivo se refere a um estado da realidade que um (ou mais) indivíduo ou organização deseja alcançar (SIMON, 1964). Por exemplo, se uma organização tem o objetivo de estar entre as 100 melhores organizações para se trabalhar, esse objetivo refere-se a um estado da realidade no qual a organização é uma das 100 melhores organizações para se trabalhar. Um objetivo também pode ser entendido como um conjunto de requisitos e restrições que se deseja satisfazer. Por exemplo, uma fábrica de ração para animais pode ter como objetivo minimizar o custo de seu produto mantendo os nutrientes mínimos necessários para a dieta do animal. Nesse caso, pode-se entender esse objetivo como um requisito (deve-se minimizar o custo do produto) associado a uma restrição (deve-se manter os nutrientes mínimos necessários) (MAEHR; ZUSHO, 2009).

Os objetivos que se deseja alcançar devem ser o ponto de partida para a identificação das medidas que devem ser utilizadas em um dado contexto, uma vez que o alcance dos objetivos poderá ser avaliado a partir das informações providas pela análise dos dados coletados para as medidas. Isso significa dizer que, como argumentado

anteriormente, a medição deve ser orientada a objetivos (BARCELLOS; FALBO; FRAUCHES, 2014).

Indicadores são cruciais para a análise do alcance a objetivos. Embora na literatura existam várias definições para indicadores, uma definição ampla e bem aceita é a de que um indicador é algo que indica ou monitora o estado de alguma coisa em algum assunto (HEINK; KOWARIK, 2010). No contexto deste trabalho, em linha com (BARCELLOS *et al.*, 2010; EPSTEIN; MANZONI, 1998; FORTUIN, 1988), dentre outros, entende-se que ‘algo’ é uma medida e que ‘alguma coisa’ é um objetivo, ou seja, um indicador é uma medida usada para monitorar o alcance de um objetivo. Vale ressaltar que, além de indicar o alcance ou não do objetivo, a natureza quantitativa de um indicador permite que se avalie o quão próximo o objetivo está de seu alcance e se há risco que o objetivo não seja alcançado (AMYOT *et al.*, 2011).

Considerando a importância da medição como provedora de informações de apoio à tomada de decisão nos mais diversos domínios e a crescente disponibilização de dados abertos, este trabalho encontra-se no contexto da obtenção de indicadores capazes de fornecer informações úteis à tomada de decisão a partir de fontes de dados (especialmente fontes de dados abertos).

1.2 Motivação

A tomada de decisão é uma atividade que está presente no dia a dia de todas as organizações. Idealmente, ela deve ser realizada a partir da análise de dados relacionados ao contexto no qual a decisão deve ser tomada. Nesse sentido, organizações coletam e armazenam dados que possam auxiliar a tomada de decisão. Porém, é comum o armazenamento de dados inúteis ou a utilização inadequada dos dados coletados, o que dificulta a obtenção das informações necessárias ou provê informações equivocadas (ALBRIGHT; WINSTON; ZAPPE, 2010).

A obtenção de informações úteis para a tomada de decisão pode não ser uma tarefa trivial. Organizações costumam armazenar seus dados em diversos sistemas e aplicações, muitas vezes, sem nenhuma integração entre eles. Em algumas situações, para a tomada de decisão, uma organização pode precisar usar dados próprios combinados com dados de clientes ou de outras organizações parceiras, o que pode dificultar, ainda mais, a identificação dos dados úteis para a obtenção das informações desejadas.

A dificuldade para se obter informações úteis a partir de dados armazenados torna-se ainda maior quando as fontes de dados a serem utilizadas são fontes de dados abertos,

pois esses dados estão disponíveis em grande volume, em diversas fontes, em diferentes formatos e com diferentes conceituações. Por exemplo, em relação ao formato, são encontrados dados disponibilizados em XML (*eXtensible Markup Language*), planilhas eletrônicas, RDF (*Resource Description Framework*) e até mesmo no formato de textos disponíveis em páginas Web, entre outros (HOXHA; BRAHAJ, 2011). Isso torna a identificação e acesso aos dados úteis uma atividade não trivial, sobretudo quando não há orientações sobre o que deve ser feito para identificar o que é útil em um determinado contexto.

Alguns trabalhos, como, por exemplo, (BRAUNSCHWEIG *et al.*, 2012; HIENERT *et al.*, 2011; FONSECA, 2014), têm abordado a extração e apresentação simplificada de dados abertos disponíveis na Web. No entanto, tipicamente, esses trabalhos consideram apenas dados estruturados como *Linked Open Data* (LOD), onde os dados são disponibilizados e compartilhados a partir de *links* que relacionam esses dados (BERNERS-LEE; HENDLER; LASSILA, 2001). Porém, boa parte da informação disponível na Web não é estruturada dessa forma, sobretudo os dados abertos brasileiros (BREITMAN *et al.*, 2012). Além disso, esses trabalhos não orientam sobre o que deve ser feito para identificar os dados realmente úteis para um determinado contexto e como utilizá-los adequadamente.

Observando-se os dados que podem ser utilizados como base para a tomada de decisão, é possível notar que eles tendem a ser dados coletados para medidas que caracterizam entidades. Por exemplo, dados disponibilizados sobre a presença de deputados nas sessões da Câmara dos Deputados permitem que sejam obtidos os valores para as medidas número de presenças e número de faltas de cada deputado nas sessões, número de deputados presentes em cada sessão, entre outras. Analisando-se, por exemplo, as medidas número de faltas e número de presenças dos deputados, seria possível caracterizar alguns deputados como faltosos e divulgar essa informação para os superiores tomarem as providências necessárias e também para a população considerar essa informação para decidir seu voto em uma eleição onde deputados faltosos sejam candidatos.

Assim, muitos dos dados disponibilizados, podem ser tratados quantitativamente sob o ponto de vista da medição, contribuindo para a objetividade das informações obtidas (ROBERTS, 1985), flexibilidade de análise (HAND; MANNILA; SMYTH, 2001) e, conseqüentemente, para a tomada de decisão. Além disso, é possível tratar dados sob a ótica do processo de medição independente do domínio ao qual os dados estejam relacionados.

Com respeito à medição, existem vários padrões que a tratam em domínios específicos (por exemplo:(ASTM, 2014; ISO/IEC, 2007; RICE *et.al.*, 2012)) e também alguns que tratam de aspectos relacionados à medição em geral, que independe do domínio de aplicação. Nesse contexto destaca-se o VIM (*International Vocabulary of Metrology*) (JCGM, 2012), que define um vocabulário sobre medição como uma tentativa de padronizar a terminologia usada nos diversos domínios. Analisando-se esses padrões é possível identificar alguns conceitos centrais presentes em todos os padrões ou na maioria deles. Porém, embora haja uma conceituação comum, muitas vezes um mesmo conceito é designado por diferentes termos em diferentes padrões e, outras vezes, um mesmo termo é usado para denotar diferentes conceitos.

Ontologias têm sido reconhecidas como um instrumento útil para reduzir ambiguidades conceituais e inconsistências, bem como para tornar as estruturas de conhecimento mais claras. Nesse sentido, ontologias podem ser usadas para promover o entendimento comum entre trabalhadores do conhecimento (USCHOLD; JASPER, 1999). A conceituação central sobre medição deve ser compartilhada em vários domínios, assim, essa conceituação pode ser representada como uma ontologia *core*, também chamada de ontologia central ou ontologia de núcleo. Ontologias centrais proveem uma definição precisa do conhecimento estrutural de uma área que é compartilhado para diversos domínios de aplicação naquela área (SCHERP *et al.*, 2011).

Visando à clareza conceitual e fidedignidade ao mundo real, ontologias centrais devem usar uma base de modelagem sólida, dada por uma ontologia de fundamentação. Conceitos e relações definidas em uma ontologia central devem estar alinhados a categorias básicas de uma ontologia de fundamentação (SCHERP *et al.*, 2011).

Além da conceituação central sobre medição, o conhecimento sobre o processo de medição, independente de domínio, também é relevante e pode ser representado como uma ontologia de tarefa. Ontologias de tarefa têm o objetivo de representar o conhecimento acerca de um processo, descrevendo a ordem em que suas tarefas são executadas, como uma tarefa é decomposta em outras subtarefas, quais papéis estão envolvidos na execução das tarefas e como objetos são envolvidos no processo (MARTINS, 2009).

Considerando esse cenário, neste trabalho é apresentada uma abordagem que visa auxiliar a identificação dos dados úteis presentes em fontes de dados, sobretudo de dados abertos, e a obtenção de indicadores que possam fornecer informações para a tomada de decisão. A abordagem proposta tem como base uma ontologia central de medição

desenvolvida à luz da *Unified Foundational Ontology* (UFO) (GUIZZARDI, 2005) e uma ontologia de tarefa de medição.

1.3 Objetivos da Pesquisa

Este trabalho tem como objetivo geral desenvolver uma abordagem baseada em ontologias de medição para obtenção de indicadores a partir de dados abertos. Esse objetivo geral pode ser decomposto nos seguintes objetivos específicos:

- (i) Desenvolver uma ontologia central de medição, para representar a conceituação central relacionada à medição, ou seja, a conceituação de medição que independe do domínio de aplicação;
- (ii) Desenvolver uma ontologia de tarefa de medição, para descrever o processo geral de medição, que é independente do domínio de aplicação;
- (iii) Desenvolver uma abordagem para obtenção de indicadores a partir de fontes de dados, usando como base as ontologias de medição definidas;
- (iv) Implementar um apoio computacional para o uso da abordagem proposta;
- (v) Avaliar o uso da abordagem proposta considerando uma fonte de dados abertos.

1.4 Método de Pesquisa

Este trabalho foi conduzido de acordo com as seguintes etapas:

- i) *Revisão da Literatura*: nesta etapa ocorreu a aquisição de conhecimento sobre os temas relacionados ao trabalho, destacando-se: dados abertos, medição e ontologias. A revisão da literatura ocorreu com a pesquisa por publicações relacionadas, através da leitura de artigos, livros, dissertações, teses e relatórios técnicos considerados relevantes ao trabalho.
- ii) *Desenvolvimento das ontologias de medição*: nesta etapa as informações obtidas a partir da revisão da literatura foram utilizadas como base para o desenvolvimento da *Core Ontology on Measurement* e da *Measurement Task Ontology*, que visam capturar, respectivamente, a conceituação central relacionada a medição e o processo geral de medição.
- iii) *Definição da Abordagem para Obtenção de Indicadores a partir de Dados Abertos*: nesta etapa o conhecimento acerca do processo de medição descrito na *Measurement Task Ontology* e a conceituação descrita na *Core Ontology on Measurement* foram utilizados como base para a definição de uma abordagem para a obtenção de

indicadores a partir de dados abertos. A abordagem propõe um conjunto de atividades que devem ser realizadas para, a partir de objetivos estabelecidos, organizar os dados de uma base de dados abertos e extrair indicadores capazes de prover informações úteis à tomada de decisão.

- iv) *Desenvolvimento de Apoio Computacional para a Abordagem*: nesta etapa foi desenvolvida uma aplicação Web para apoiar o uso da abordagem proposta.
- v) *Avaliação Preliminar da Proposta*: nesta etapa foi realizada uma avaliação preliminar da abordagem proposta. Para isso, a ferramenta desenvolvida foi utilizada em um estudo que explorou a obtenção de indicadores a partir de bases de dados abertos disponibilizadas pela ANAC (Agência Nacional de Aviação Civil) e pela INFRAERO (Empresa Brasileira de Infraestrutura Aeroportuária).
- vi) *Escrita da Dissertação*: os resultados obtidos durante a execução dos passos anteriores foram documentados nesta dissertação.

1.5 Organização da Dissertação

Neste capítulo inicial foram apresentadas as principais ideias desta dissertação, descrevendo o contexto de aplicação, motivação, objetivos e as etapas realizadas. Além desta introdução, este texto é composto pelos seguintes capítulos e apêndice:

- **Capítulo 2 (Medição, Ontologias e Dados Abertos)**: apresenta a fundamentação teórica relevante para este trabalho. Inclui conceitos e outros aspectos relacionados à medição; conceitos e classificação de ontologias, apresentação de UFO (*Unified Foundational Ontology*), a ontologia fundamentação utilizada neste trabalho; e, também, alguns aspectos relacionados a dados abertos.
- **Capítulo 3 (Ontologias de Medição)**: apresenta a *Core Ontology on Measurement* e a *Measurement Task Ontology* desenvolvidas nesse trabalho e que serviram como base para a abordagem proposta.
- **Capítulo 4 (Abordagem para Obtenção de Indicadores a partir de Dados Abertos)**: apresenta a abordagem proposta neste trabalho, um apoio computacional desenvolvido para apoiar seu uso e os resultados da utilização da abordagem em bases de dados abertos disponibilizadas pela ANAC e INFRAERO.

- **Capítulo 5 (Conclusões e Perspectivas Futuras):** apresenta as considerações finais do trabalho, as contribuições e propostas de trabalhos futuros para continuidade e aprimoramento do trabalho.
- **Apêndice 1 (Mapeamento entre COM e RSMO):** apresenta um mapeamento entre questões de competência e axiomas da *Core Ontology on Measurement* e da *Reference Software Measurement Ontology*, que foi usada como uma das fontes para desenvolver *Core Ontology on Measurement*.

Capítulo 2

Revisão da Literatura

Este capítulo apresenta a fundamentação teórica relevante para este trabalho. São discutidos aspectos relacionados a medição e ontologias. A ontologia de fundamentação UFO é apresentada. Alguns trabalhos que abordam dados abertos também são explorados.

2.1 Medição

Medir está presente nos mais diversos aspectos da vida humana. Mede-se o tamanho e o peso de pessoas e objetos, a área de um cômodo de uma casa ou a distância entre as cidades. Mas, para que medir?

Um pediatra mede o peso e a altura de uma criança para conhecer sua curva de crescimento e tomar as decisões pertinentes em caso de desvios. Um arquiteto mede a área de um cômodo de uma casa para tomar decisões sobre a disposição de móveis e iluminação. Um motorista mede a distância entre cidades para decidir por qual rota seguir em uma viagem. Ou seja, medir é essencial para o conhecimento, o controle e a tomada de decisão (ROCHA; SANTOS; BARCELLOS, 2012).

Na literatura existem diversas definições para medição (MARI, 2013). No passado, medição foi definida como o ato de atribuir valores numéricos a propriedades (STEVENS, 1951; TORGERSON, 1958). No entanto, medição de valores não numéricos também é legítima (por exemplo, é possível atribuir o valor ‘quente’ para a temperatura de um objeto quando ela é medida por alguém tocando o objeto) (FRIGERIO; GIORDANI; MARI, 2010; ROBERTS, 1985). Além disso, medição não é um único ato. Assim, medição pode ser definida como um processo que envolve um conjunto de ações com a finalidade de caracterizar entidades a partir da atribuição de valores às suas propriedades (BARCELLOS; FALBO; FRAUCHES, 2014).

Segundo Hempel (1952), a atribuição de valores não numéricos a propriedades de entidades permite que elas sejam classificadas. Por exemplo, objetos podem ser classificados em quentes ou frios. Por outro lado, a atribuição de magnitudes a propriedades provê um avanço à classificação, uma vez que proporciona uma caracterização mais precisa do que somente a classificação. Por exemplo, dizer que um objeto está a 100°C o caracteriza melhor em termos de sua temperatura do que dizer que ele está quente.

Medição vem sendo estudada há muitos anos por pesquisadores das mais diversas áreas e com diferentes focos. Alguns estudos dedicam-se a questões fundamentais da teoria da medição e tratam os conceitos da medição principalmente em termos matemáticos (FINKELSTEIN; LEANING, 1984; MODAVE, 2001; STEVENS, 1951). Outros estudos preocupam-se com aspectos relacionados ao ato de medir, como o uso de instrumentos para realizar a medição e a precisão dos valores obtidos (MESSICK, 1975; SPEITEL, 1992). Há, ainda, estudos que abordam aspectos mais práticos, como a aplicação da medição para resolver problemas em contextos específicos (ALTMAN; BLAND, 1983; ASSOCIATION; ASSOCIATION; FEDERATION, 2012; LAORY *et al.*, 2012; MARSHALL; JOHNELL; WEDEL, 1996; ZHU; WANG, 2013).

Considerando que o tema medição pode ser abordado sob diferentes perspectivas, Finkelstein (2009) afirma que, quando se fala em medição, é necessário esclarecer se o discurso refere-se à medição fortemente definida ou à medição fracamente definida. Segundo o autor, medição fortemente definida diz respeito às medições no sentido mais restrito da Física, onde a representação dos resultados se dá apenas por números e teorias fortemente embasadas em áreas como Física e Química, são necessárias para se obter os resultados. Por outro lado, medição fracamente definida diz respeito à medição de uma forma geral, que trata da representação de propriedades através de símbolos (por exemplo, uso do símbolo ‘quente’ para representar a temperatura de um indivíduo após tocá-lo ou uso do símbolo ‘38° C’ para representar a temperatura de um indivíduo medida utilizando-se um termômetro).

Um dos principais elementos da medição é a medida, que pode ser definida como a função ou mapeamento que relaciona uma propriedade com a escala que estrutura os valores que podem ser atribuídos àquela propriedade em uma medição (SASSOON, 2010). Medidas podem ser classificadas em dois tipos, dependendo da forma como são obtidas (DIEZ, 1996; JCGM, 2012; ROBERTS, 1985; STEVENS, 1946). O primeiro tipo envolve uma propriedade cuja quantificação é realizada diretamente, sem a necessidade de que outras propriedades tenham sido medidas anteriormente. Como exemplos, têm-se peso e altura. O segundo tipo envolve uma propriedade cuja quantificação depende da quantificação de outras propriedades. Em outras palavras, para medir essa propriedade, outras propriedades devem ser medidas antes. Como exemplo desse tipo tem-se o índice de massa corporal, que para ser medido precisa que peso e altura tenham sido medidos anteriormente.

Na maioria dos trabalhos que abordam esses tipos de medida o primeiro tipo é chamado de medida base ou fundamental e o segundo de medida derivada ou composta. Contudo, alguns autores tratam essa classificação de maneira diferente e a abordam como uma classificação da medição e não da medida. Por exemplo, para Andreas (2008), medição fundamental é aquela que envolve tão somente os conceitos metrológicos da propriedade medida durante a medição. Já a medição derivada envolve conceitos que vão além da propriedade medida. Por exemplo, quando se mede a temperatura utilizando um termômetro de mercúrio, a temperatura do objeto medido dilata o mercúrio dentro do cilindro do termômetro. Com isso, o topo do líquido do mercúrio sobe e utiliza-se a altura do topo do líquido para determinar a temperatura do objeto. Nesse exemplo, segundo Andreas (2008), a temperatura em graus Celsius medida através de um termômetro de mercúrio é uma medição derivada, pois utiliza conceitos de volume e comprimento para que seja obtido o valor da temperatura.

Quando se fala em medida, dois conceitos são fundamentais: escala e unidade de medida. A escala refere-se à estrutura que contém o conjunto de todos os valores possíveis de serem atribuídos a uma dada propriedade em uma medição. Por exemplo, a escala de uma medida que quantifica a propriedade altura (de pessoas) é o conjunto de todos os valores possíveis de ser obtidos quando a altura de uma pessoa for medida. Unidade de medida, por sua vez, denota a unidade, estabelecida por convenção, na qual uma medida é expressa (JCGM, 2012). A unidade de medida tem papel fundamental na medição, pois ela dá significado ao valor atribuído a uma propriedade. Por exemplo, dizer que a temperatura de um dado objeto é 100 não leva a nenhum entendimento, pois não se sabe o que 100 significa. No entanto, dizer que a temperatura de um dado objeto é 100 graus Celsius tem significado e permite que alguma conclusão seja inferida (por exemplo, que o objeto está quente).

Em Mari (1999) é apresentada uma proposta que trata medição utilizando três componentes nos quais os conceitos de medida, escala e unidade de medida são explorados. O primeiro componente, chamado *information-from-selection*, trata da identificação de todos os valores possíveis de serem atribuídos a uma propriedade durante uma medição. Segundo Mari (1999), a medição pode ser considerada como uma operação onde se seleciona um símbolo de um conjunto de todas as representações de resultados possíveis para essa medição. Nesse sentido, o componente *information-from-selection* materializa a teoria de que uma medição predispõe de um conjunto de valores que podem ser selecionados e

atribuídos à propriedade durante uma medição. Segundo o autor, nesse componente a semântica dos símbolos não é tratada.

O segundo componente, chamado *information-from-structure*, trata do estabelecimento de uma relação entre um estado de uma propriedade de uma entidade e os símbolos ou valores usados para representar essa propriedade. Este componente sugere que para cada conjunto de valores possíveis para uma propriedade definidos em *information-from-selection* existe uma semântica distinta. Por exemplo, o mesmo símbolo '3', sendo um valor possível de ser atribuído à propriedade tamanho, assume diferentes significados se relacionado a metros, centímetros ou milímetros. Assim, este componente aborda os conceitos de medida, escala e unidade de medida. A medida faz a relação funcional entre o resultado de uma medição e a propriedade medida, no sentido de estruturar a propriedade medida e sua representação. A escala, por sua vez, pode ser entendida como uma extensão da definição de conjunto de valores possíveis descrito em *information-from-selection*, na qual é adicionada semântica a esse conjunto, por meio da associação da escala com uma unidade de medida que permite que o símbolo selecionado em uma medição possa ser interpretado.

O último componente, chamado *information-from-connection*, trata da relação entre o valor atribuído a uma propriedade na medição de uma entidade e a caracterização dessa entidade a partir desse valor. Em outras palavras, esse componente se refere à relação da propriedade quantificada com as características que podem ser inferidas a partir dessa quantificação. Assim, é possível predicar sobre uma entidade através da medição de suas propriedades, ou seja, pode-se inferir características de uma entidade a partir da quantificação de suas propriedades. Por exemplo, a partir da medição das dimensões de um objeto, alguém poderia concluir se o objeto cabe ou não dentro de uma caixa, sem ter que inserir o objeto na caixa para chegar a essa conclusão.

Conforme mencionado anteriormente, é importante notar que medição não é um único ato, mas um processo que visa caracterizar entidades a partir da atribuição de valores às suas propriedades (BARCELLOS; FALBO; FRAUCHES, 2014). Pode-se perceber que as discussões conduzidas por Mari (1999) estão bem em linha com esse argumento, pois ela trata não só do ato de medir, mas também da definição do que se deseja medir e as medidas que devem ser usadas, bem como do uso dos valores medidos para caracterizar as entidades medidas.

Há diversas propostas para o processo de medição, seja no âmbito geral, seja para domínios específicos. De maneira geral, as atividades que compõem o processo de medição são planejamento, execução e análise (FRIGERIO; GIORDANI; MARI, 2010; ISO/IEC,

2007; LJUNGBERG, 2002; PAPADIMITRIOU et al., 2012). Durante o planejamento da medição, as entidades a serem medidas são identificadas (ex.: objetos, fenômenos e processos), bem como suas propriedades a serem medidas (ex.: tamanho, área e custo), quais medidas deverão ser utilizadas para quantificar essas propriedades (ex.: tamanho em metros e custo em reais) e que procedimentos devem ser realizados para coletar dados para as medidas definidas. A decisão sobre quais entidades e propriedades serão medidas e quais medidas serão usadas deve ser guiada por objetivos de medição. Por exemplo, se o objetivo de medição é verificar a altura dos membros de um time de basquete, seria necessário medir a propriedade altura das entidades membros do time, aplicando-se a medida altura em metros. Uma vez que a medição tenha sido planejada, ela pode ser executada. A execução da medição consiste em coletar dados para as medidas definidas aplicando-se os procedimentos de medição estabelecidos. Finalmente, na análise da medição os dados coletados são representados de forma a facilitar a análise e são analisados à luz dos objetivos estabelecidos a priori, sendo os resultados da análise comunicados às partes interessadas.

2.1.1 Métodos de Apoio à Medição em Organizações

Na literatura existem alguns métodos que visam apoiar aspectos relacionados ao processo de medição. Nesta seção são apresentados alguns deles.

Um dos métodos mais conhecidos no contexto da medição aplicada ao domínio de Engenharia de Software é o GQM (*Goal-Question-Metrics*) (BASILI; ROMBACH, 1994), que orienta a identificação de medidas a partir de objetivos organizacionais. O GQM baseia-se no entendimento de que para uma organização realizar medição de forma eficiente, ela deve ser guiada por propósitos bem definidos. Apesar de ter sido desenvolvido, a priori, para organizações de software, o GQM pode ser adaptado para outros domínios.

O resultado da aplicação do GQM é uma estrutura hierárquica com três níveis (Figura 2.1): (i) nível conceitual (Objetivo), (ii) nível operacional (Questão) e (iii) nível quantitativo (Medida). Através de uma abordagem *top-down* e orientada a objetivos, são definidos objetivos que são refinados em questões. Medidas são, então, definidas de forma que sejam adequadas para responder às questões. A análise das medidas permite verificar o grau de alcance dos objetivos (SOLINGEN e BERGHOUT, 1999 APUD ROCHA *et al.*, 2012).

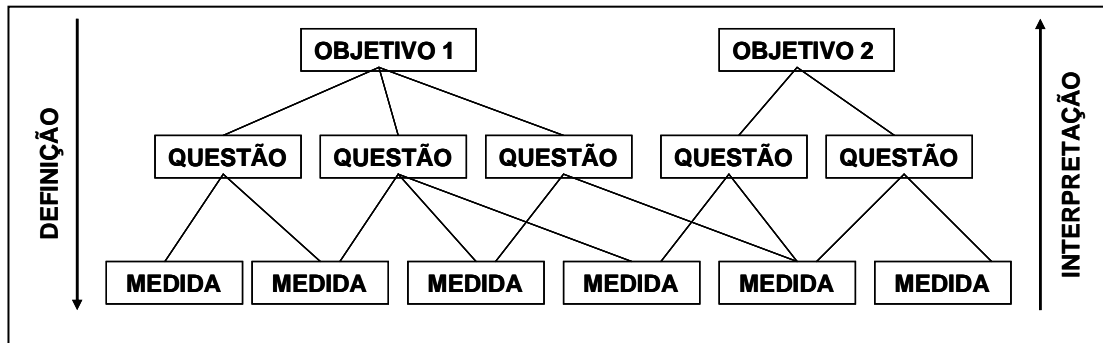


Figura 2.1 – GQM (ROCHA et al., 2012).

Na Figura 2.2 é apresentado um exemplo de uso do método GQM para definição de medidas. Nesse exemplo tem-se o objetivo de melhorar as estimativas de projeto. Esse objetivo pode ser refinado em várias questões como, por exemplo: “Qual a precisão das estimativas de cronograma do projeto?” e “Qual a precisão das estimativas de esforço do projeto?”. Essas questões podem ser respondidas através das medidas.

Objetivo
Propósito: Melhorar
Questão: precisão
Objeto: estimativas de projeto
Ponto de vista: analisado pelo ponto de vista dos gerentes de projeto

Questão 1
 Qual a precisão das estimativas de cronograma do projeto?
Medida 1a)
 Precisão Total de Cronograma = $\frac{\text{tempo real de todo o projeto}}{\text{tempo estimado do projeto}}$

Medida 1b)
 Precisão Cronograma por atividade = $\frac{\text{tempo real por atividade}}{\text{tempo estimado por atividade}}$

Questão 2
 Qual a precisão das estimativas de esforço do projeto?
Medida 2a)
 Precisão Total do Esforço = $\frac{\text{esforço real de todo o projeto}}{\text{esforço estimado para o projeto}}$

Medida 2b)
 Precisão esforço por atividade = $\frac{\text{esforço real por atividade}}{\text{esforço estimado por atividade}}$

Figura 2.2 – Exemplo de uso do método GQM.

Outro método que pode ser utilizado no contexto da medição é o BSC (Balanced ScoreCard) (KAPLAN; NORTON, 1996), que é um método usado para descrever, implementar e gerenciar a estratégia da organização. Nesse sentido, pode ser usado para

traduzir a missão da organização em objetivos estratégicos e estabelecer um conjunto de indicadores e medidas de desempenho. O BSC considera quatro perspectivas: (i) financeira; (ii) clientes; (iii) processos de negócios internos e (iv) aprendizado e crescimento..

Para cada perspectiva, com base em objetivos da organização, é selecionado um pequeno e conciso número de medidas. Para cada medida é estabelecida uma meta a ser alcançada. Entende-se por meta, um valor de medida, ao qual se prevê ou pretende alcançar. Aliado à medida e sua meta, devem-se identificar ações corretivas ou intervenções a serem realizadas caso a meta não seja alcançada. Os dados são coletados ao longo do tempo e analisados e, em caso de metas não serem alcançadas, ações corretivas devem ser colocadas em prática. A perspectiva financeira reúne medidas capazes de fornecer informações sobre como as finanças da organização estão. Essas medidas tendem a ser direcionadas aos acionistas da organização. A perspectiva de clientes envolve medidas que respondem como os clientes estão satisfeitos e veem a organização. A perspectiva de processos de negócios internos reúne medidas que auxiliam a identificar quais aspectos dos processos devem ser melhorados. Por fim, a perspectiva de aprendizado e crescimento reúne medidas capazes de fornecer informações que permitam a melhoria contínua, inovação e criação de valores.

Outro método relacionado à medição em organizações é o uso de KPIs (*Key Performance Indicators*) (PARMENTER, 2010). KPIs são medidas utilizadas para compreender se os objetivos da organização estão sendo atingidos. Alguns exemplos comuns de KPIs usados no contexto de negócios são: *Time to Market* (tempo de lançamento de um produto, que começa com a idealização do conceito e termina quando o produto está disponível para venda), *Lead Time* (duração de um determinado processo), *Stock Out* (tempo que um determinado produto leva para alcançar estoque zero) e *Market Share* (fatia do mercado que um determinado produto conquistou durante um determinado período de tempo).

Pidun e Felden (2011) argumentam que no âmbito organizacional alguns aspectos são mais difíceis de serem mensurados por KPIs, que são fundamentalmente numéricos. Nesses casos, os autores sugerem o uso de *soft goals*, fatores de sucesso, ontologias e modelagem de processos de negócios a fim de fazer uma avaliação qualitativa dos aspectos. Segundo os autores, essas alternativas devem ser utilizadas aliadas à medição.

Apesar de, atualmente, haver várias propostas que tratam a medição em organizações de uma forma abrangente, ou seja, considerando que praticamente qualquer aspecto de uma organização pode ser medido, esse pensamento é relativamente novo. Não

faz muito tempo que a avaliação quantitativa de uma organização se limitava a medidas relacionadas a valores financeiros. Ou seja, acreditava-se que para avaliar se uma organização estava atingindo bons resultados, e algumas vezes, no intuito de tentar prever o futuro, eram utilizadas apenas medidas tais como retorno de investimento (*ROI – Return of Investment*), despesas e custos.. À medida que o entendimento sobre medição e o reconhecimento de seus benefícios e aplicações aumentaram, aumentou, também, a percepção de que, para o sucesso de um negócio, outros aspectos deveriam ser medidos, como clientes e recursos (KAPLAN; NORTON, 1996; PIDUN; FELDEN, 2011; TONI; TONCHIA, 2001).

Mesmo havendo métodos de apoio à identificação de medidas úteis para as organizações, muitas ainda têm selecionado e utilizado medidas erroneamente. Uma das razões é que poucas organizações se preocupam com o que é um indicador, o que ele deve representar para a organização e como o processo de identificação dos indicadores, coleta e análise de dados deve ser conduzido (PARMENTER, 2010).

2.2 Ontologias

A disciplina denominada Ontologia é originária da Filosofia como um ramo da meta-física que pode ser definida como "A ciência do ser enquanto ser" e busca estabelecer um sistema de categorias para uma porção particular da realidade (GUIZZARDI, 2007). Segundo Gruber (1995), uma ontologia é uma especificação explícita de uma conceituação. Nessa definição Gruber utiliza o termo conceituação no sentido de uma visão abstrata e simplificada do mundo, ou seja, ontologias devem capturar uma visão específica a respeito da realidade, abstraindo os aspectos essenciais de um contexto. Assim, ontologias podem ter visões distintas sobre a mesma porção da realidade e, ainda assim, serem corretas semanticamente.

Ontologias têm sido reconhecidas nos mais diversos campos de sistemas de informação como um instrumento para representação do conhecimento de forma mais clara, explícita e sem ambiguidades. Com o objetivo de definir ontologias mais precisas, no início do século 20, foi empregado o termo Ontologia Formal, como uma analogia a Lógica Formal. Essa vertente buscava utilizar a teoria da lógica formal para restringir a interpretação pretendida a respeito de um domínio (GUIZZARDI, 2007).

Ontologia, como um artefato da engenharia, é constituída por um vocabulário específico, usado para descrever uma determinada realidade, adicionado a um conjunto de suposições da realidade. Normalmente, essas suposições são representadas por predicados

de primeira ordem, unários ou binários, chamados respectivamente de conceitos e relações. Em casos mais simples, uma ontologia descreve uma hierarquia de conceitos. Em casos mais sofisticados, axiomas são adicionados para expressar outras relações entre conceitos para restringir a interpretação pretendida (GUARINO, 1998; GUIZZARDI, 2007).

2.2.1 Tipos de Ontologias

Existem diversas classificações para ontologias. Uma das mais citadas na literatura e adotada neste trabalho é a proposta por Guarino (1998), mostrada na Figura 2.3, que classifica ontologias por seu o nível de generalidade.

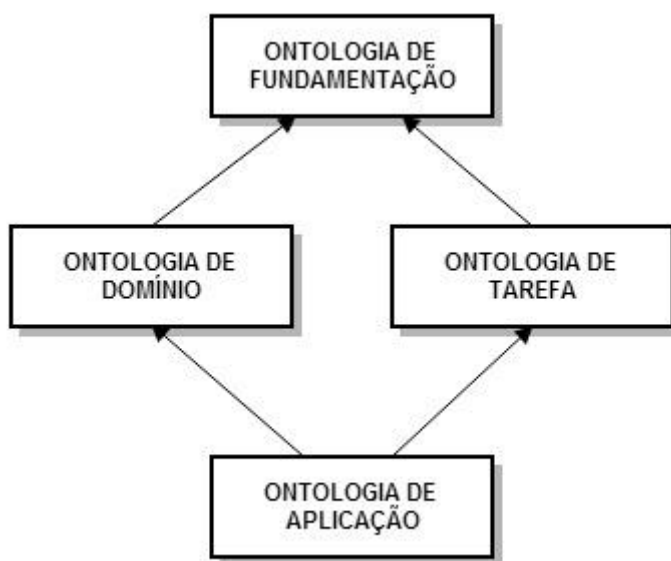


Figura 2.3 - Tipos de ontologias e relacionamentos entre eles (GUARINO, 1998).

Ontologias de fundamentação, também conhecidas como ontologias de alto nível, são ontologias que representam sistemas de categorias filosoficamente bem fundamentados e independentes de domínio. Essas ontologias modelam os conceitos e relações mais básicos e gerais que compõem o mundo, como objeto, evento e relações todo-parte (GUIZZARDI, 2005). As ontologias de fundamentação são desenvolvidas com a intenção de servirem como base para a construção de novas ontologias. Alguns exemplos de ontologias de fundamentação encontrados na literatura são *Bung-Wand-Weber* (BWW) (WAND; WEBER, 1990), *General Formal Ontology* (GFO) (HERRE *et al.*, 2006), *Descriptive Ontology for Language and Cognitive Engineering* (DOLCE) (BORGO; MASOLO, 2009) e *Unified Foundational Ontology* (UFO) (GUIZZARDI, 2005). Esta última foi utilizada neste trabalho e será apresentada mais adiante.

Com um nível menor de generalidade, **ontologias de domínio** descrevem o vocabulário relacionado a um domínio específico (GUARINO, 1998). Ontologias de domínio são, normalmente, desenvolvidas para compartilhar informação, tornar suposições do domínio explícitas ou analisar o conhecimento do domínio. Ontologias de domínio têm sido investigadas, principalmente, por comunidades da Inteligência Artificial e Engenharia de Domínio. Uma grande quantidade de ontologias de domínio tem sido desenvolvida para os mais diversos domínios, tais como medicina, direito, engenharia, organizações e química (GUIZZARDI, 2005).

Ontologias de tarefa proveem um vocabulário de termos usados para resolver problemas associados com uma tarefa, independente de domínio, através de um conjunto de primitivas de representação. O conhecimento de tarefa é associado à descrição de decomposição de tarefas em subtarefas, do controle de fluxo ao longo dessas subtarefas e dos papéis de conhecimento do domínio que são usados ou produzidos pelas subtarefas (MARTINS, 2009).

Por fim, **ontologias de aplicação** têm o objetivo de reduzir os *gaps* entre ontologias de domínio e de tarefa e permitir que especialistas de domínio usem a mesma linguagem adotada em uma aplicação (GENNARI *et al.*, 1994). Em outras palavras, uma ontologia de aplicação define os conceitos relevantes de um domínio no desenho de uma aplicação em particular. Normalmente, constitui-se de um conjunto de conceitos tomados de diversas ontologias de domínios e de tarefa (HEIJSTF; SCHREIBER; WIELINGA, 1997).

Além dos tipos de ontologias presentes na classificação de Guarino, **ontologias centrais** também são relevantes no contexto deste trabalho. Ontologias centrais ou ontologias de núcleo proveem uma definição precisa de um conhecimento estruturado em um campo específico que abrange diferentes domínios de aplicação (SCHERP *et al.*, 2011). Ontologias centrais são concebidas visando ao reuso, pois, uma vez que modelam a conceituação comum a diversos domínios, elas podem ser utilizadas como base para a construção de ontologias para esses domínios (KIM *et al.*, 2007). Tendo em vista que uma ontologia central pode usada em diversos domínios do conhecimento, é importante que ela seja definida com base em uma ontologia de fundamentação, a fim de prover maior clareza conceitual e ser fidedigna ao mundo real. Assim, ontologias centrais proveem um refinamento aos conceitos de uma ontologia de fundamentação, adicionando conceitos e relacionamentos específicos do campo que a ontologia central trata (SCHERP *et al.*, 2011).

Ontologias centrais estão em um nível intermediário de generalidade entre ontologias de fundamentação e domínio. Nesse sentido, não são genéricas o suficiente para representarem as definições mais básicas e genéricas do mundo como as ontologias de fundamentação, e nem descrevem um domínio específico do problema como ontologias de domínio. Falbo *et al.* (2013) defendem que não existe uma separação exata entre os tipos de ontologias, e sim uma gradação contínua que vai desde as ontologias de fundamentação até as ontologias de domínio, passando por ontologias centrais, conforme ilustra a Figura 2.4.

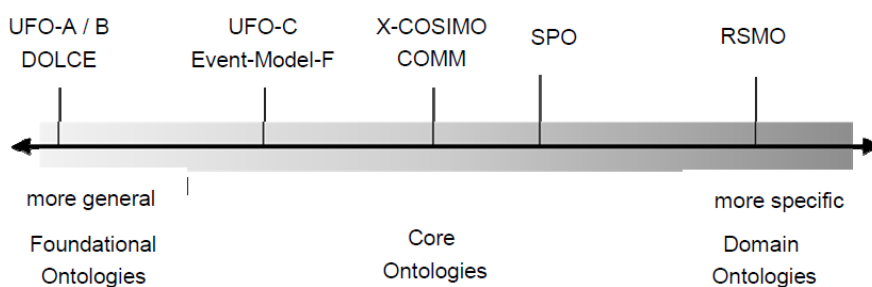


Figura 2.4 - Tipos de ontologias em uma visão contínua (FALBO et al., 2013).

Na Figura 2.4, as ontologias mais à esquerda são ontologias de fundamentação, descrevendo conceitos gerais do mundo e fundamentadas em teorias filosóficas, como é o caso de DOLCE (BORGO; MASOLO, 2009) e UFO (partes A e B) (GUIZZARDI, 2005). Ontologias mais à direita são ontologias de domínio, que abordam conceituações de domínios específicos, como é o caso de RSMO (*Reference Software Measurement Ontology*), que trata da conceituação relacionada à medição em Engenharia de Software (BARCELLOS; FALBO; DAL MORO, 2010).

À medida que se segue da esquerda para a direita na figura, o nível de generalidade diminui. Assim, há ontologias classificadas como ontologias centrais que são mais gerais que outras, também classificadas como tal. Na Figura 2.4 há cinco ontologias centrais, todas desenvolvidas com base em alguma ontologia de fundamentação e que podem ser usadas como base para o desenvolvimento de ontologias mais específicas: UFO-C (GUIZZARDI; FALBO; S.S.GUIZZARDI, 2008) e *Event-Model-F* (SCHERP *et al.*, 2009, 2012) são ontologias que tratam de conceitos sociais como objetivos e intenções; X-COSIMO (*Cross-Context Semantic Information Management Ontology*) (FRANZ; STAAB; ARNDT, 2007) aborda a comunicação entre pessoas e sistemas; COMM (*Core Ontology on Multimedia*) (ARNDT *et al.*, 2007, 2009; STAAB *et al.*, 2008) trata de conceitos relacionados à multimídia; e SPO (*Software Process Ontology*) (BRINGUENTE; FALBO; GUIZZARDI, 2011) trata aspectos relacionados a processos de software. Embora as cinco ontologias

sejam ontologias centrais, *Event-Model-F* e UFO-C são mais gerais que COMM e X-COSIMO, que tratam de conceituações mais próximas da conceituação de um domínio (multimídia e sistemas de informações) que as anteriores. COMM e X-COSIMO, por sua vez, são mais gerais que SPO, que trata de uma conceituação ainda mais próxima da conceituação de um domínio.

2.2.2 Unified Foundational Ontology - UFO

Conforme dito na seção anterior, idealmente, ontologias centrais devem ser desenvolvidas com base em uma ontologia de fundamentação. Para o desenvolvimento da ontologia central proposta neste trabalho foi utilizada a ontologia de fundamentação UFO (ALBUQUERQUE, 2013; GUIZZARDI; WAGNER; FALBO, 2013; GUIZZARDI, 2005), que tem sido desenvolvida com base em teorias na área de Ontologias Formais, Lógica Filosófica, Filosofia da Linguagem, Linguística e Psicologia Cognitiva. UFO é composta por três partes: UFO-A, uma ontologia de objetos (*Endurants*) e cerne de UFO, UFO-B, uma ontologia de eventos (*Perdurants*) e UFO-C, uma ontologia de entidades sociais (*endurants* e *perdurants*) construída com base em UFO-A e B. Na Figura 2.5 é apresentado um fragmento de UFO contendo conceitos de UFO-A, UFO-B e UFO-C que são relevantes para este trabalho. Os conceitos destacados em cinza são conceitos diretamente usados na ontologia central proposta neste trabalho e descrita no próximo capítulo. As descrições dos conceitos de UFO apresentadas nesta seção foram baseadas principalmente em (ALBUQUERQUE, 2013; GUIZZARDI; FALBO; S.S.GUIZZARDI, 2008; GUIZZARDI, 2005). Nas descrições, os conceitos de UFO são apresentados em *itálico* e exemplos dos conceitos são apresentados sublinhados.

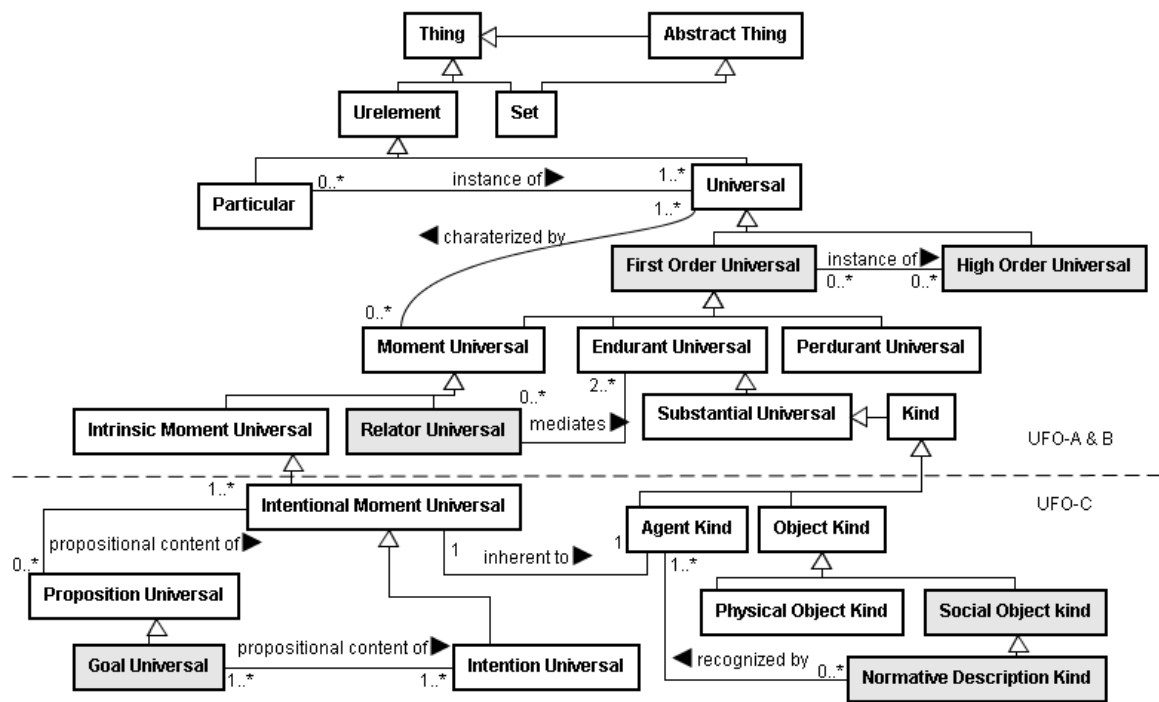


Figura 2.5 - Fragmento de UFO.

O conceito raiz de UFO é *Thing*, que é especializado em *Urelement* e *Set*. *Sets* são entidades abstratas (*Abstract Thing*), ou seja, entidades que simplesmente existem, sem ser explicitamente criadas ou destruídas. *Urelements* são todas as entidades que não são *Sets*. Em relação a *Urelements*, uma distinção fundamental que existe em UFO é entre as categorias *Particular* e *Universal*. *Universals* são padrões de características que podem ser instanciados em diferentes entidades (por exemplo, *Pessoa*). *Particulars*, por sua vez, são entidades que existem na realidade, possuindo uma identidade única (por exemplo, a pessoa *Maria*). O modelo apresentado na Figura 2.5 mostra como *Universals* são decompostos. Uma vez que os conceitos definidos a partir de *Particular* são análogos aos especializados a partir de *Universal*, eles não são apresentados na Figura 2.5.

Universals podem ser *First Order Universals*, ou seja, *universals* cujas instâncias são *particulars* (por exemplo, *Pessoa*, cujas instâncias são indivíduos, tais como a pessoa *Maria*), ou *High Order Universals*, que são *universals* cujas instâncias também são *universals* (por exemplo, *Mamífero*, cujas instâncias poderiam ser *Pessoa*, *Cachorro*, *Cavalo*, etc.). *First Order Universals* podem ser classificados em *Endurant Universals*, *Perdurant Universals* e *Moment Universals*. *Endurant Universals* persistem no tempo, mantendo sua identidade (por exemplo, *Pessoa*). *Perdurant Universals* (eventos), por sua vez, são constituídos por várias partes temporais que são distribuídas ao longo do tempo (por exemplo, *Processo*). *Moment Universals* (propriedades) são *universals* que caracterizam outros *universals*. Por exemplo, o *moment universal* *Cor* caracteriza o *universal* *Maçã*. *Moments* são existencialmente dependentes

de outra entidade, no sentido de que, por exemplo, a cor de uma maçã depende da maçã para existir. Dependência existencial pode ser utilizada, também, para diferenciar *intrinsic* e *relational moments*. *Intrinsic Moment Universals* são dependentes de uma única entidade para existir (por exemplo, Cor). *Relator Universals*, por sua vez, dependem de uma pluralidade de entidades (por exemplo, Emprego) e, por isso, proveem a relação material¹ entre eles.

Substantial Universals compreendem as entidades que não precisam de outra para existir. Enquanto persiste no tempo, um *substantial* pode instanciar diversos tipos de *substantial universals*. *Kind* é um tipo de *substantial* que é instanciado em todas as possíveis situações e define o que o *substantial* é. Por exemplo, Pessoa, pois uma pessoa sempre será uma pessoa enquanto ela existir.

Com relação a *Kinds*, uma importante distinção feita em UFO se dá entre agentes e objetos. Um *Object Kind* é um *substantial universal* não agente. Suas instâncias (*Objects*) não agem, apenas participam de ações. *Object Kinds* podem ser categorizados em *Physical Object Kind* (por exemplo, Livro) e *Social Object Kind* (por exemplo, Linguagem). Uma *Normative Description Kind* é um *social object kind* cujas instâncias definem uma ou mais regras/normas reconhecidas por pelo menos um agente (por exemplo, um método descrevendo um conjunto de diretivas sobre como executar uma ação em uma organização). Um *Agent Kind* é um *substantial universal* cujas instâncias (*Agents*) são capazes de executar ações com alguma intenção.

Intentional Moment Universal é um tipo especial de *intrinsic moment universal* cujas instâncias são inerentes a *Agents*. Um *Intentional Moment Universal* possui um conteúdo proposicional chamado *Proposition Universal*. *Intentional moment universals* em que a intencionalidade é "pretender algo" ("*intending something*") são chamados de *Intention Universals*. Uma *Intention* caracteriza uma situação desejada por um *Agent* (por exemplo, uma organização pode ter a intenção de ser bem sucedida). O conteúdo proposicional de uma *intention* é um *Goal* (por exemplo, estar entre as dez maiores empresas do mundo poderia ser o conteúdo proposicional da *intention* ser bem sucedida).

Na Figura 2.6 é apresentado um fragmento de UFO-A que trata conceitos relacionados a *Qualities*.

¹ Diferente das relações ditas formais, que acontecem diretamente, sem a necessidade de outra entidade existir (por exemplo, a relação entre o conhecimento x de um indivíduo e esse indivíduo existe tão logo x e o indivíduo existam), relações materiais dependem de uma entidade com o poder de mediar outras entidades (um *relator*). Por exemplo, a relação trabalhar entre um indivíduo e uma empresa depende de uma entidade (o *relator* emprego) para mediar a relação entre o indivíduo e empresa.

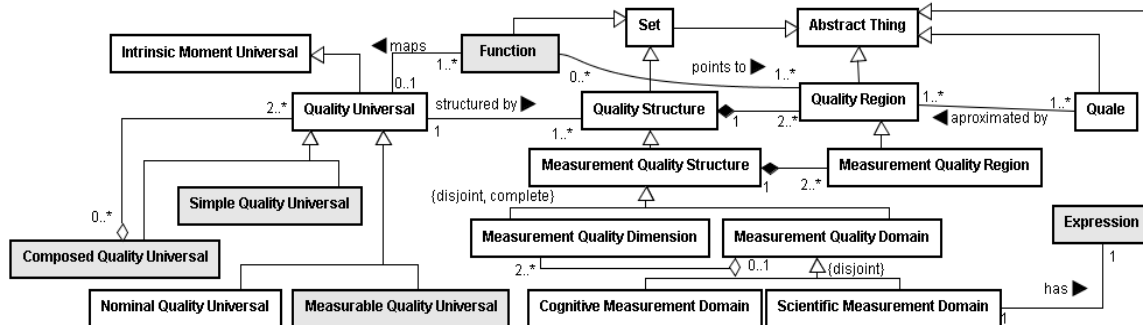


Figura 2.6 - Fragmento de UFO-A relacionado a *Qualities*.

Quality Universals referem-se a propriedades que caracterizam *Universals* (por exemplo, Peso e Altura podem caracterizar o universal Pessoa). *Quality Universals* são *Intrinsic Moment Universals* associados a *Quality Structures*, as quais podem ser entendidas como o conjunto de todas as possíveis regiões que delimitam o espaço de valores que podem ser associados a um dado *Quality Universal*. Por exemplo, o *quality universal* Peso é associado a uma *Quality Structure* dada por um espaço de valores que é uma estrutura linear isomórfica à porção positiva do eixo dos números reais. As regiões que compõem uma *Quality Structure* são chamadas de *Quality Regions* e são regiões que aproximam *qualia*². Um *Quale* é uma percepção de um *quality* em uma *quality structure*. Por exemplo, seja o *quality* altura que caracteriza uma dada pessoa, o ponto na *quality structure* que indica a altura da pessoa é o *quale*. Um *quale* é uma percepção e, dessa forma, é intrínseco a agentes cognitivos, não podendo ser diretamente compartilhado ou comunicado. A comunicação de um *quale* é feita por meio de símbolos (por exemplo, 1.86 pode ser o símbolo usado para comunicar a altura de certa pessoa).

Function é uma especialização de *Set* que mapeia instâncias de um *Quality Universal* para pontos em uma *Quality Structure*. De acordo com a *quality structure* a que estão associados, *Quality Universals* são classificados em *Simple Quality Universals* e *Composed Quality Universals*. *Simple Quality Universals* são associados a *quality structures* unidimensionais (por exemplo, Peso), enquanto que *Composed Quality Universals* são associados a *quality structures* multidimensionais (por exemplo, Índice de Massa Corporal).

Uma importante distinção a respeito de *quality universals* está relacionada com sua natureza. *Measurable Quality Universals* são *quality universals* que podem ser medidos objetivamente por agentes cognitivos ou dispositivos de medição, sendo possível estabelecer distâncias entre suas *quality regions*. Peso e Altura são exemplos de *Measurable Quality Universals*. Diferentemente, *Nominal Quality Universals*, como Nome e CEP, são

² Plural de *quale*.

normalmente baseados em convenções sociais e não podem ser medidos objetivamente. Este trabalho envolve apenas *Measurable Quality Universals*, portanto, aspectos relacionados a *Nominal Quality Universals* não serão discutidos.

Measurement Quality Structures são estruturas que permitem avaliar objetivamente distâncias entre dois valores e verificar se esses valores são iguais ou não. São classificadas de acordo com o número de dimensões em *Measurement Quality Dimension*, que representam *measurement quality structures* mais elementares (unidimensionais), e *Measurement Quality Domain*, que representam *quality structures* multidimensionais. *Measurement Quality Domains*, por sua vez, podem ser *Cognitive Measurement Quality Domain* ou *Scientific Measurement Quality Domain*. A diferença prática entre os dois tipos é que regiões de *scientific domains* podem ser avaliadas e ordenadas qualitativamente, enquanto regiões de *cognitive domains* não podem. *Scientific domains* são compostos de acordo com alguma álgebra e possuem uma *Expression* para determinar sua formação. Por exemplo, o *scientific measurement quality domain* do Índice de Massa Corporal (IMC) é formado utilizando as dimensões Peso e Altura e de acordo com a expressão $IMC = \text{Peso} / (\text{Altura} \times \text{Altura})$.

A Figura 2.7 apresenta um fragmento de UFO-A que trata conceitos relacionados a *Reference Structures*.

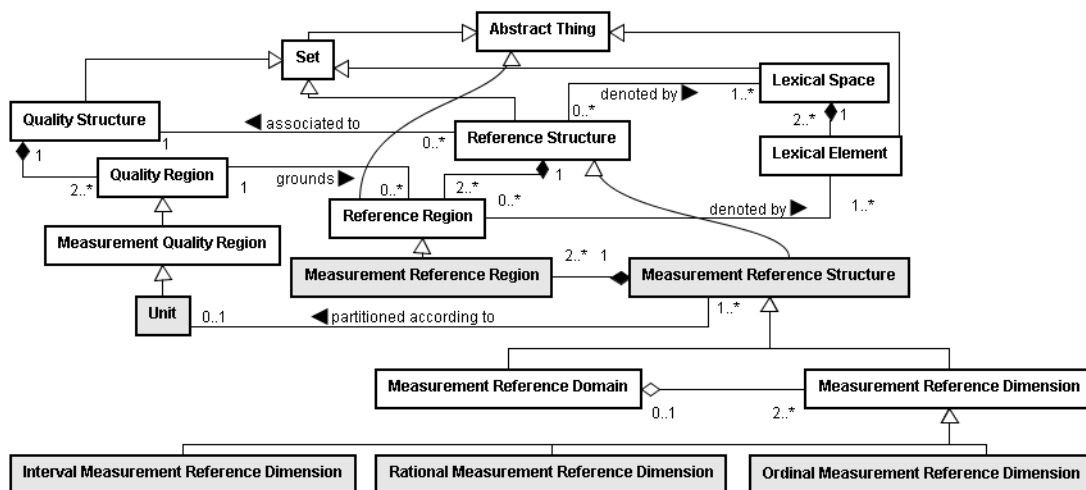


Figura 2.7 - Fragmento de UFO-A relacionado a *Reference Structures*.

Como dito anteriormente, uma *Quality Region* é uma região que aproxima um *quale*. Uma vez que, um *quale* não pode ser compartilhado ou comunicado, para permitir a sua comunicação, é necessário o uso de símbolos (*Lexical Elements*) associados a *Reference Regions* e *Reference Structures*. Uma *Reference Region* é uma entidade abstrata baseada em uma *Quality Region* que age como uma ponte entre essas regiões e os elementos léxicos usados para

comunicar o *quale* aproximado. Em outras palavras, uma *Reference Region* relaciona uma *Quality Region* que aproxima um *quale* ao símbolo usado para comunicar esse *quale*. Uma *Reference Structure*, por sua vez, é associada a uma *Quality Structure* e é um conjunto de *Reference Regions* baseadas em *Quality Regions* daquela *Quality Structure*. Dessa forma, quando o 'valor' de um *quality* é denotado por um elemento léxico (por exemplo, uso do elemento léxico 1.86 para denotar a altura de uma pessoa), o que está realmente sendo referenciado é a *quality region* que mais aproxima o *quale* comunicado através do símbolo 1.86.

Reference Structures são topologicamente isomórficas às *Quality Structures* às quais são associadas. Então, possuem o mesmo número de dimensões e suas *Reference Regions* são isomórficas às *Quality Regions* da *Quality Structure*. *Reference Structures* associadas a *Measurement Quality Structures* são chamadas *Measurement Reference Structures* e agem como escalas baseadas nas *quality structures*. Elas são compostas por *Measurement Reference Regions*. *Measurement Reference Structures* podem ser particionadas em espaços com a mesma magnitude de acordo com uma *Unit*.

De acordo com o número de dimensões, *Measurement Reference Structures* podem ser classificadas em *Measurement Reference Domain* e *Measurement Reference Dimension*. *Measurement Reference Dimensions*, por sua vez, podem ser classificadas em *Ordinal Measurement Reference Dimension*, *Interval Measurement Reference Dimension* e *Rational Measurement Reference Dimension*. Como os nomes sugerem, *Ordinal*, *Interval* e *Rational Reference Dimensions* representam *reference structures* com propriedades de escala ordinal, intervalar e de razão, respectivamente.

Para ilustrar os conceitos *Quale*, *Quality Structure*, *Quality Region*, *Reference Structure*, *Reference Region* e as relações entre eles, na Figura 2.8 tem-se a representação dos conceitos ilustrando o que ocorre quando se diz “o peso dessa maçã é $2a$ ”. O peso p da maçã m é um *quale* que é aproximado por uma *Quality Region* em uma *Quality Structure*. A *Quality Region* está associada a uma *Reference Region* de uma *Reference Structure* e a *Reference Region* é denotada pelo símbolo léxico 2 , que é usado para comunicar o *quale*. A *Reference Structure* é particionada de acordo com a unidade a , sendo que na figura a é um valor arbitrário. Por exemplo, se $a = 100\text{g}$, o peso comunicado da maçã seria 200g .

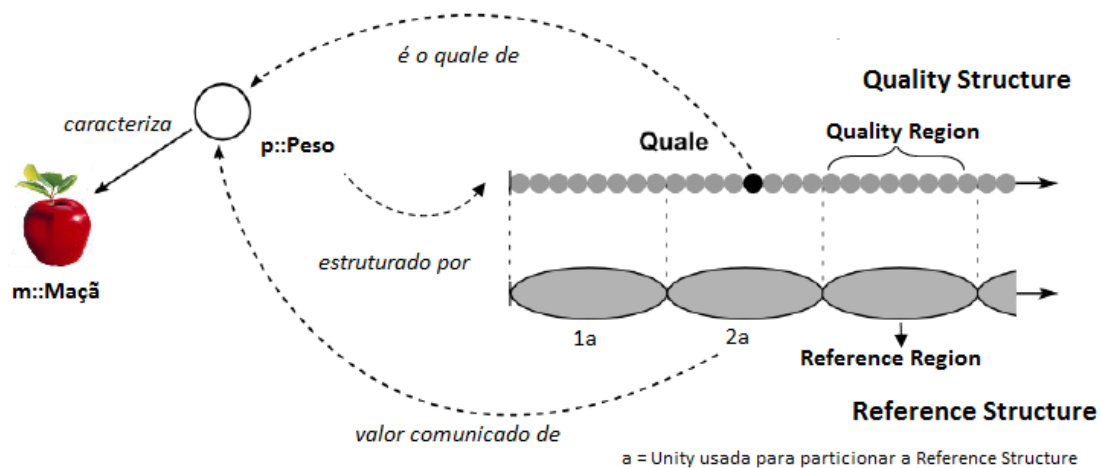


Figura 2.8 – Relação entre *Quale*, *Reference Region* e *Quality Region* (adaptado de ALBUQUERQUE, 2013).

2.3 Dados Abertos

O termo Dados Abertos surgiu em 2000 para indicar a acessibilidade da informação que é pública e disponível a qualquer indivíduo, sem a necessidade de alguma licença ou patente (AUER *et al.*, 2007). Em outras palavras, o conceito Dados Abertos faz referência ao conjunto de dados produzido e disponibilizado por uma organização³, buscando transparência. Esses dados, eventualmente, serão consumidos e redistribuídos pela sociedade.

Diversas organizações publicam seus dados gratuitamente na Web. Com isso, é cada vez maior o volume de dados disponíveis. Embora o objetivo da publicação dos dados seja permitir que eles sejam acessados por qualquer indivíduo e prover transparência, a forma como os dados são publicados, muitas vezes, não permite que se obtenham informações úteis a partir deles. Geralmente, os dados são disponibilizados em estruturas e formatos heterogêneos, variando desde XML (*eXtensible Markup Language*), RDF (*Resource Description Framework*), OWL (*Web Ontology Language*) e *Web Services*, até planilhas, arquivos CSV (*Comma-Separated Values*) e páginas HTML (*HyperText Markup Language*) (BÖHM *et al.*, 2010; HOXHA; BRAHAJ, 2011).

Além da variedade de formatos, também há diversidade de fontes e tipos de dados disponíveis (desde informações sobre o consumo de combustível de uma frota até dados com localizações geoespaciais) que podem ser livremente obtidos utilizando-se mecanismos

³ O termo organização é usado aqui para denotar qualquer entidade social que disponibilize dados abertos, tais como: órgãos governamentais, empresas e organizações não governamentais, entre outros.

diversos, tais como APIs (*Application Programming Interface*), documentos, *Web Services* e bancos de dados abertos (HEATH; BIZER, 2011).

Considerando essa heterogeneidade de formatos e tipos de dados e a necessidade de acessá-los mais facilmente, foi desenvolvida uma proposta, chamada *Linked Open Data* (LOD), que busca unificar a forma de disponibilização de dados abertos (BERNERS-LEE; HENDLER; LASSILA, 2001). *Linked Open Data* utiliza a mesma ideia da *World Wide Web*, onde são compartilhados documentos relacionados por *links*. Dessa forma, em *Linked Open Data* os dados são compartilhados a partir de *links* que relacionam esses dados. Assim, em analogia a um documento fazer referência a outro documento através de uma URL (*Uniform Resource Locator*), um objeto (dado) faz referência a outro objeto através de URIs (*Uniform Resource Identifier*). Por exemplo, um documento RDF contendo dados a respeito do carro Fusca pode ter um relacionamento denominado *isProducedBy* que faz referência a uma URI que, quando acessada através de uma requisição HTTP (*HyperText Transfer Protocol*), fornece um RDF contendo dados da empresa Volkswagen.

A proposta de *Linked Open Data* consiste em um conjunto de princípios para relacionar e anotar dados utilizando vocabulários ou esquemas, com o intuito de integrar os dados abertos. Nessa proposta (i) instâncias devem ser representadas por URIs, de forma que (ii) todos possam acessá-los através de uma requisição HTTP; (iii) cada URI, quando acessada, deve prover informações úteis sobre o objeto em um formato triplicado (RDF), e (iv) incluir *links* para outras URIs quando em sua descrição (esquema RDF) houver alguma referência. Assim, o *Linked Open Data* tem como objetivo relacionar dados de diferentes fontes e disponibilizados por diferentes organizações de forma que seja possível consultar os dados de maneira simples e integrada (BERNERS-LEE, 2006).

Apesar de ser uma tecnologia que vem sendo cada vez mais utilizada e que apresenta vantagens para representação e compartilhamento de dados, em *Linked Open Data* os dados são armazenados em triplas, não sendo adequadas para agregações em grande volume de dados, sobretudo para dados quantitativos. Dados estatísticos e quantitativos são mais adequadamente armazenados em formatos tabulares como, por exemplo, em banco de dados relacionais, *datawarehousing* e cubos (KIMBALL; ROSS, 2002). Além disso, aplicações cujos dados são baseadas em *Linked Open Data*, onde usuários consomem dados de proveniências desconhecidas, podem gerar semânticas conflitantes com a cognição do usuário (BIZER; HEATH; BERNERS-LEE, 2009).

Outra forma de acessar dados abertos é através de APIs provedoras de dados. Nesse caso, é necessário um desenvolvedor que conheça as funções/métodos da API

provedora de dados para criar uma nova aplicação que consuma os serviços disponíveis nessa API. Através de APIs é possível combinar dados de diferentes provedores e obter aplicações, ou *mashups*, especializadas. O site *programmableweb*⁴ reúne diversos Web APIs. Exemplos bem conhecidos são *Amazon Product Advertising API*⁵ e *facebook API*⁶ (HEATH; BIZER, 2011).

Associando o uso de *Linked Open Data* e APIs, Hoxha e Brahaj (2011) utilizaram *data cube vocabulary* (W3C, 2014), que é uma representação dimensional em RDF na qual a agregação de dados não é suportada, para estruturar dados estatísticos governamentais. Para melhorar a manipulação dos dados, os autores usaram o *Google Visualization API*⁷, que é baseado em uma tabela tipada, como um cubo, para representações gráficas, através de conceitos de *datawarehousing*. Dessa forma, os dados são armazenados em RDF, mas durante sua manipulação são transportados para uma estrutura mais adequada.

O tema Dados Abertos é consideravelmente recente e ainda há muito a ser explorado. Analisando-se a literatura percebe-se que uma das preocupações iniciais no contexto de Dados Abertos foi permitir que os dados fossem publicados e acessados, tendo sido desenvolvidas algumas propostas de tecnologias para facilitar o acesso aos dados disponibilizados, como o *Linked Open Data* e APIs, discutidos anteriormente.

Para serem publicados, os dados precisam ser obtidos. Nesse contexto, tem-se, entre outros, o trabalho de Cordeiro *et al.*(2011), no qual é proposta uma arquitetura baseada em uma ontologia para o domínio de emergências para capturar e disponibilizar no formato RDF dados relacionados a emergências. Os dados são providos principalmente pelas equipes responsáveis por emergências e, adicionalmente, por cidadãos e jornalistas. Informações a respeito da proveniência dos dados são obtidas e utilizadas para avaliar a veracidade dos dados (MENDONÇA *et al.*, 2013). Os dados capturados são triplicados, relacionados, transformados (filtrados e convertidos, quando necessário) e armazenados em um repositório de triplas. Por fim, são explorados e consumidos a partir de um sistema.

A preocupação com a obtenção dos dados para publicação e acesso se estendeu para permitir não apenas acessar os dados, mas explorá-los, ou seja, selecionar dados que sejam úteis em um determinado contexto e analisá-los para obter informações úteis. Buscando facilitar a exploração de dados disponibilizados em formato RDF, Araújo e Schwabe (2009) apresentam uma ferramenta para usuários não familiarizados com a tecnologia RDF. A ferramenta permite a navegação, seleção e exploração de dados

⁴ <http://programmableweb.com/>

⁵ <http://docs.amazonwebservices.com/AWSECommerceService/latest/DG/>

⁶ <https://developers.facebook.com/>

⁷ <http://code.google.com/apis/visualization/>

armazenados em um repositório de dados RDF através de uma interface amigável na qual é possível realizar consultas aos dados.

Braunschweig *et al.*, (2012) também propõem uma ferramenta de apoio à exploração de dados abertos armazenados no formato RDF. A ferramenta requer que alguns conjuntos de dados do *Linked Open Data*, ou seja, triplas em RDF, sejam registrados manualmente para que seus dados e metadados sejam armazenados em um banco de dados relacional a partir do qual o usuário pode fazer consultas estruturadas ou através de palavras-chave para explorar os dados armazenados.

Outro exemplo é o trabalho de Fonseca (2014), no qual é realizado um mapeamento e integração entre bases de dados abertos brasileiras (Servidores e Remuneração, Sistema de Informações Organizacionais do Governo Federal (SIORG) e Catálogos de Unidades Federativas) utilizando a ontologia ORG (REYNOLDS, 2014), que aborda conceitos governamentais. Os dados resultantes da integração foram publicados e disponibilizados de forma que consultas diversas pudessem ser feitas (por exemplo, pode-se investigar a remuneração e lotação de um dado servidor público).

Ainda no contexto de exploração dos dados para obtenção de informações úteis, tem-se a ferramenta proposta por Hienert *et al.*, (2011), que possibilita o *upload* de arquivos em formato RDF, CSV e planilhas Excel, para integrar dados e gerar gráficos que facilitem sua interpretação.

Embora sejam encontradas na literatura diversas iniciativas relacionadas à exploração de dados abertos, percebe-se que a maioria foca em dados publicados no formato RDF, relacionamentos de dados por meio de *Linked Open Data* e apresentação dos dados em ferramentas computacionais para consultas *ad hoc* (JANSSEN; ZUIDERWIJK, 2014). Apesar de muitos dados serem publicados no formato RDF, principalmente devido às vantagens providas pelo *Linked Open Data*, ainda há muitos dados publicados em outros formatos. Por exemplo, o Brasil tem realizado um trabalho para disponibilização de dados governamentais⁸ pelo Comitê de Organização da Informação. O projeto DadosGov provê um catálogo com cerca de 1.300 séries históricas de dados representando 8 anos de ações do governo. Esses dados foram disponibilizados em planilhas e nos formatos XML e JSON (*JavaScript Object Notation*) por mais de 40 órgãos diferentes. O volume total é de aproximadamente 2,5 milhões de registros (BREITMAN *et al.*, 2012).

Além de muitos trabalhos se limitarem a dados armazenados no formato RDF, apesar de haver propostas de ferramentas que permitem a exploração dos dados, os

⁸ Decreto N° 7.724, DE 16 DE MAIO DE 2012, regulamenta a garantia do acesso à informação (http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/Decreto/D7724.htm)

trabalhos não orientam sobre como identificar os dados relevantes para a obtenção de informações úteis em um dado contexto. Considerando essa lacuna, este trabalho propõe uma abordagem que apoia a obtenção de indicadores capazes de fornecer informações úteis à tomada de decisão a partir de dados abertos. Conforme dito no capítulo de Introdução desta dissertação, a abordagem proposta é definida com base em uma ontologia central de medição e em uma ontologia da tarefa de medição, as quais são apresentadas no próximo capítulo.

Capítulo 3

Ontologias de Medição

Neste capítulo são apresentadas a Core Ontology on Measurement e a Measurement Task Ontology desenvolvidas no contexto deste trabalho e nas quais a abordagem proposta se baseia.

3.1 Core Ontology on Measurement - COM

A *Core Ontology on Measurement* (COM) tem como propósito representar a conceituação central relacionada à medição, ou seja, a conceituação que é independente do domínio no qual a medição é aplicada. As principais fontes de conhecimento para o desenvolvimento de COM foram o VIM (*International Vocabulary of Metrology*) (JCGM, 2012) e a *Reference Software Measurement Ontology* (RSMO) (BARCELLOS; FALBO; ROCHA, 2013; BARCELLOS, 2009).

O VIM, conforme mencionado no Capítulo 1, define uma terminologia relacionada à metrologia⁹, compreendendo princípios básicos sobre quantidades, unidades de medidas, procedimentos de medição, entre outros. A terminologia proposta no VIM foi desenvolvida como uma iniciativa de padronização do vocabulário relacionado à metrologia. Seus conceitos baseiam-se, principalmente, na norma ISO/IEC 31 (ISO, 1992). No contexto do desenvolvimento de COM, VIM forneceu conhecimento relacionado aos conceitos básicos da medição, tais como medida, medição e valor medido, e auxiliou na identificação do que faz parte do cerne da conceituação sobre medição e é independente do domínio de aplicação.

A *Reference Software Measurement Ontology* (RSMO) é uma ontologia que descreve o domínio de medição de software e foi desenvolvida com base em UFO. Embora RSMO trate do domínio de medição de software, alguns de seus conceitos são centrais à medição e são independentes de domínio. Para o desenvolvimento de COM, RSMO contribuiu com a nomenclatura e com o conhecimento relacionado a alguns aspectos que não são muito explorados no VIM, como, por exemplo, a análise de medição.

Embora a RSMO tenha sido fundamentada em UFO, recentemente foram realizadas algumas evoluções em UFO (ALBUQUERQUE, 2013) que impactam diretamente na fundamentação dos conceitos centrais de medição. Assim, embora a fundamentação de

⁹ Metrologia pode ser entendida como a ciência da medição e sua aplicação (JCGM, 2012).

alguns conceitos de RSMO tenha servido de inspiração para a fundamentação dos conceitos de COM, a maior parte da fundamentação precisou ser revista visando à adequação às evoluções de UFO. Além disso, foram realizadas alterações em alguns aspectos da conceituação apresentada, para tratar a medição de forma geral. No Apêndice 1 são apresentados os mapeamentos entre COM e RSMO.

COM foi desenvolvida seguindo a abordagem SABiO (*Systematic Approach for Building Ontologies*) (FALBO, 2004), cujas principais atividades são: (i) identificação do propósito e requisitos a serem atendidos, utilizando-se questões de competência que a ontologia deve ser capaz de responder; (ii) captura da conceituação com base nas questões de competência, identificando-se e organizando-se conceitos, relações, propriedades e axiomas relevantes; (iii) formalização da ontologia em uma linguagem formal; (iv) integração com ontologias existentes, visando aproveitar conceituações existentes; (v) avaliação da ontologia para verificar se satisfaz aos requisitos estabelecidos; e (vi) documentação da ontologia.

Conforme discutido no Capítulo 2, ontologias centrais podem ter diferentes graus de generalidade. Considerando a representação contínua mostrada na Figura 2.5 (Capítulo 2), COM situa-se em um ponto mais próximo das ontologias de domínio do que das ontologias de fundamentação.

3.1.1 As Subontologias de COM

A *Core Ontology on Measurement* é formada por cinco subontologias, como mostra a Figura 3.1. Na figura, é utilizado um diagrama de pacotes UML (*Unified Modelling Language*) para representar as subontologias e seus relacionamentos. O relacionamento de dependência entre uma subontologia e outra (por exemplo, entre *Measure* e *Measurable Entities*), indica que a primeira subontologia utiliza conceitos da ontologia indicada.

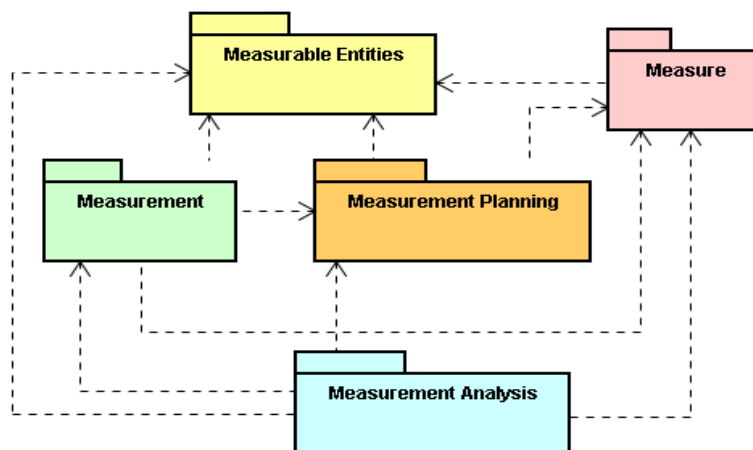


Figura 3.1 – As subontologias de COM.

A seguir, as subontologias de COM são descritas, sendo apresentados para cada subontologia as questões de competência, o modelo conceitual, a descrição do modelo, a fundamentação dos conceitos à luz de UFO e os axiomas definidos. Nas descrições dos modelos das subontologias, usa-se **negrito** para os conceitos de COM em Português (quando aparecem pela primeira vez), *itálico* para os termos de UFO e sublinhado para os exemplos (instâncias dos conceitos). Os conceitos de COM em Português são seguidos dos termos equivalentes em Inglês (os termos usados nos modelos), usando-se *itálico e fonte diferenciada*. Nos modelos, os conceitos são apresentados em cores diferentes para indicar suas subontologias de origem. Conceitos oriundos das subontologias de COM são apresentados nas cores dos pacotes correspondentes às subontologias na Figura 3.1. Conceitos de UFO são apresentados na cor branca.

3.1.1.1 Subontologia Entidades Mensuráveis (*Measurable Entities*)

Esta subontologia aborda a conceituação relacionada a entidades que podem ser medidas, seus tipos e propriedades.

Questões de Competência:

As questões de competência que esta subontologia visa responder são:

QC1. Quais tipos de entidade podem ser medidos?

QC2. Qual é o tipo de uma determinada entidade mensurável?

QC3. Quais elementos mensuráveis caracterizam um tipo de entidade mensurável?

QC4. Quais elementos mensuráveis caracterizam uma entidade mensurável?

QC5. Quais elementos mensuráveis podem ser diretamente medidos e quais não podem?

QC6. A partir de quais elementos mensuráveis um elemento indiretamente mensurável pode ser medido?

Modelo Conceitual:

O modelo conceitual da subontologia *Measurable Entities* é apresentado na Figura 3.2.

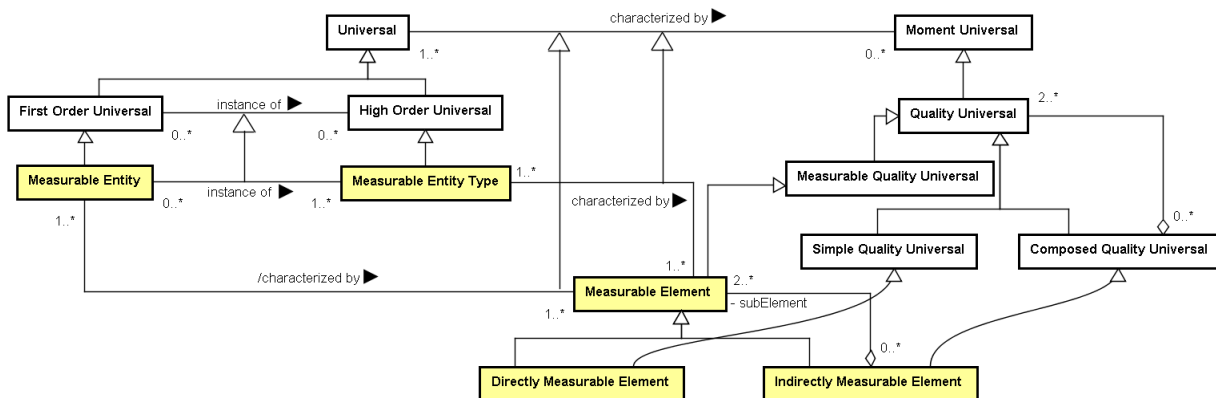


Figura 3.2 – Modelo conceitual da subontologia *Measurable Entities*.

Descrição do Modelo:

Uma instância de **Entidade Mensurável** (*Measurable Entity*) é algo que pode ser medido, como, por exemplo, uma pessoa, um projeto e uma organização. Entidades mensuráveis são classificadas em **Tipos de Entidade Mensurável** (*Measurable Entity Type*). Por exemplo, pessoas são entidades mensuráveis do tipo de entidade mensurável Pessoa. Entidades mensuráveis são caracterizadas por **Elementos Mensuráveis** (*Measurable Element*s) que são propriedades que podem ser medidas. Por exemplo, uma pessoa pode ser caracterizada por seu peso e altura. Elementos Mensuráveis podem ser **Elementos Diretamente Mensuráveis** (*Directly Measurable Element*), os quais não dependem de outros elementos para serem medidos (por exemplo, peso), ou **Elementos Indiretamente Mensuráveis** (*Indirectly Measurable Element*), cuja medição depende de outros elementos, ditos seus **subelementos** (*subElement*). Índice de massa corporal é um exemplo de elemento indiretamente mensurável, pois depende dos elementos peso e altura para ser medido.

Fundamentação:

O conceito **Entidade Mensurável** (*Measurable Entity*) é um tipo de *First Order Universal* que é caracterizado por *Measurable Quality Universals* (propriedades que podem ser medidas), designando as coisas que podem ser medidas. **Tipo de Entidade Mensurável** (*Measurable Entity Type*), por sua vez, corresponde, em UFO, a *High Order Universals* caracterizados por *Measurable Quality Universals*. Em linha com UFO, onde *First Order Universal* é instância de *High Order Universal*, Entidade Mensurável é instância de Tipo de Entidade Mensurável.

Entidades mensuráveis são caracterizadas por **Elementos Mensuráveis** (*Measurable Element*), os quais são *Measurable Quality Universals* em UFO, denotando propriedades que podem ser medidas. Em UFO, *Quality Universals* denotam propriedades presentes em *Universals* e estão associados a *Quality Structures*, que podem ser entendidas como o conjunto de todas as possíveis regiões que delimitam o espaço de valores que pode ser associado a um dado *quality universal*. *Quality Measurable Universals* são *quality universals* que podem ser medidos, ou seja, cujas instâncias estão localizadas em uma *quality structure* e podem ser identificadas por agentes ou dispositivos de medição.

Elementos Diretamente Mensuráveis (*Directly Measurable Element*) correspondem em UFO a *Simple Quality Universals* (*qualities* associados a estruturas de qualidade unidimensionais) que podem ser medidos; **Elementos Indiretamente Mensuráveis** (*Indirectly Measurable Element*), por sua vez, dizem respeito a *Composed Quality Universals* (*qualities* associados a estruturas de qualidade multidimensionais) passíveis de medição.

Axiomas:

Uma vez que os modelos não são capazes de capturar sozinhos algumas restrições que envolvem os conceitos, para cada subontologia foi definido um conjunto de axiomas em lógica de primeira ordem, a fim de tornar essas restrições explícitas. Para a subontologia *Measurable Entities* foram definidos os axiomas apresentados a seguir.

A1. Se um elemento mensurável *elm* é subelemento de um elemento indiretamente mensurável *elm-im*, então *elm* e *elm-im* devem caracterizar o mesmo tipo de entidade mensurável *tp-ems*.

$$(\forall elm \in \text{Measurable Element}, elm\text{-im} \in \text{Indirectly Measurable Element}) (\text{subElement}(elm, elm\text{-im}) \rightarrow (\exists tp\text{-ems} \in \text{Measurable Entity Type}) \text{characterizedBy}(tp\text{-ems}, elm) \wedge \text{characterizedBy}(tp\text{-ems}, elm\text{-im}))$$

A2. Se uma entidade mensurável *em* é instância de um tipo de entidade mensurável *tem* e um elemento mensurável *elm* caracteriza esse *tem*, então *elm* caracteriza *em*.

$$(\forall em \in \text{Measurable Entity}, tem \in \text{Measurable Entity Type}, elm \in \text{Measurable Element}) (\text{instanceOf}(em, tem) \wedge \text{characterizedBy}(tem, elm) \rightarrow \text{characterizedBy}(em, elm))$$

A3. Se um elemento mensurável $elm1$ é subelemento de um elemento indiretamente mensurável $elm2$ e este é subelemento de outro elemento indiretamente mensurável $elm3$, então $elm1$ é subelemento de $elm3$.

$$(\forall elm1 \in Measurable\ Element, elm2, elm3 \in Indirectly\ Measurable\ Element) \\ (subElement(elm1, elm2) \wedge subElement(elm2, elm3)) \rightarrow subElement(elm1, elm3))$$

3.1.1.2 Subontologia Medida (*Measure*)

Esta subontologia aborda os aspectos relacionados à definição de medidas.

Questões de Competência:

As questões de competência que esta subontologia visa responder são:

- QC1. Quais medidas podem ser usadas para quantificar um elemento mensurável?
- QC2. Quanto à dependência de uma medida em relação a outras, qual é a natureza de uma medida?
- QC3. Qual é a unidade de medida de uma medida?
- QC4. Qual é a escala de uma medida?
- QC5. Qual é o tipo de uma escala?
- QC6. Quais são os valores de escala e que, por conseguinte, podem ser atribuídos a uma medida?
- QC7. Como medir uma medida?
- QC8. Quais são as medidas correlatas a uma medida?

Descrição do Modelo:

Medida (*Measure*) é uma função utilizada para quantificar um Elemento Mensurável (*Measurable Element*), permitindo associá-lo a **Valores de Escala** (*Scale Value*), contidos em uma determinada **Escala** (*Scale*). Uma Escala é uma estrutura que identifica os possíveis valores para os quais uma medida pode ser mapeada. Cada valor ou região que forma uma escala é um Valor de Escala (*Scale Value*). Medidas podem ser expressas em **Unidades de Medida** (*Measure Unit*), que são unidades definidas e adotadas por convenção com a finalidade de particionar a escala em regiões iguais. Como exemplo, tem-se a medida altura em metros, expressa na unidade de medida metros e que pode ser usada para quantificar o elemento mensurável altura, permitindo associá-lo a algum valor da escala formada pelos valores equivalentes aos números reais positivos. Escalas podem ser dos seguintes tipos: **Escala Ordinal** (*Ordinal Scale*), **Escala Intervalar** (*Intervale Scale*) ou **Escala Razão** (*Rational Scale*). Uma medida tem Escala Ordinal se os valores para as quais pode ser mapeada representam uma certa ordem ou *ranking*. Por exemplo, a medida grau de satisfação do cliente, que pode ser mapeada para os valores alto, médio ou baixo, tem escala ordinal. Uma medida tem Escala Intervalar quando é possível quantificar a distância entre seus valores e não existe um valor representando nulidade. Por exemplo, a medida temperatura em graus Celsius, expressa na unidade de medida graus Celsius, possui escala intervalar, pois é possível quantificar a distância entre dois valores (por exemplo, há 10°C entre as temperaturas 20°C e 30°C), mas o valor zero não significa nulidade (0°C não significa ausência de temperatura). Em uma Escala Intervalar, embora seja possível calcular a distância entre dois valores, não é possível estabelecer outras relações entre eles (por exemplo, não é possível dizer que 50°C é duas vezes mais quente que 25°C). Por fim, uma medida tem Escala de Razão quando, além de ser possível quantificar a distância entre valores para os quais ela pode ser mapeada, o zero absoluto, representando um ponto de nulidade, é um valor possível. Além disso, é possível realizar operações matemáticas entre os valores, mesmo que sejam usadas unidades de medida diferentes. Por exemplo, a medida distância em quilômetros, expressa em quilômetros (km), tem escala razão, pois 0km significa a nulidade da distância (ou seja, nenhuma distância) e, além disso, é possível, por exemplo, relacionar valores de distância em quilômetros com valores de distância em metros e, também, estabelecer relações entre valores (a distância 100km é duas vezes a distância 50km).

Em relação à sua natureza, uma medida pode ser **Medida Base** (*Base Measure*) ou **Medida Derivada** (*Derived Measure*). Medidas base são medidas atômicas, que não

dependem de outras para serem obtidas. Dessa forma, elas são usadas para quantificar Elementos Diretamente Mensuráveis. Medidas derivadas, por sua vez, são obtidas a partir de outras e são usadas para quantificar Elementos Indiretamente Mensuráveis.

Medidas são obtidas aplicando-se **Procedimentos de Medição** (*Measurement Procedure*). Um Procedimento de Medição aplicável a uma medida derivada é chamado de **Procedimento de Medição Derivativo** (*Derivative Measurement Procedure*) e inclui **Fórmulas de Medição** (*Measurement Formula*) que quantificam as relações entre as medidas usadas para obtenção da medida derivada. Um Procedimento de Medição aplicável a uma medida base é chamado **Procedimento de Medição Básico** (*Basic Measurement Procedure*).

Por fim, uma medida pode se relacionar com outras, ditas **medidas correlatas** (*correlated measure*), podendo influenciar seus valores. Por exemplo, as medidas distância e tempo são medidas correlatas, uma vez que a distância a ser percorrida influencia no tempo necessário para percorrê-la. As medidas a partir das quais uma medida derivada é obtida são medidas correlatas a ela. Por exemplo, as medidas peso e altura são medidas correlatas à medida índice de massa corporal.

Fundamentação:

Uma **Medida** (*Measure*) é uma função utilizada para quantificar um Elemento Mensurável (*Measurable Element*), permitindo que um valor seja atribuído a ele, ou seja, uma Medida permite associar um valor¹⁰ a um Elemento Mensurável. Nesse sentido, Medida é uma *Function* em UFO.

Medidas Derivadas (*Derived Measure*) são calculadas por fórmulas (*Measurement Formula*), que são *Expression* em UFO. Medida possui **Escala** (*Scale*), que é uma *Measurement Reference Structure* em UFO. Uma *Measurement Reference Structure* é composta pelas regiões que representam os valores que podem ser atribuídos a uma dada propriedade. Essas regiões são ditas *Measurement Reference Regions*. Assim, uma vez que **Valor de Escala** (*Scale Value*) representa os possíveis valores a serem associados a um Elemento Mensurável, Valor de Escala é uma *Measurement Reference Region* em UFO. Uma Escala pode ser particionada de acordo com uma **Unidade de Medida** (*Measure Unit*). Em UFO, *Measurement Reference Structure* são particionadas por *Units*. Dessa forma, Unidade de Medida é uma *Unit*. Os tipos de escala **Ordinal**, **Intervalar** e de **Razão** (*Ordinal, Interval e Rational Scale*) correspondem em UFO aos conceitos *Ordinal, Interval e Rational Reference Dimensions*, os quais representam *reference structures* com propriedades de escala ordinal, intervalar e razão, respectivamente.

¹⁰ O conceito Valor Medido, que denota o valor que é associado a um Elemento Mensurável quando se aplica uma Medida em uma Medição, é tratado na subontologia *Measurement*, apresentada mais adiante.

Um **Procedimento de Medição** (*Measurement Procedure*) descreve o procedimento que deve ser realizado para que seja possível obter valores para uma medida, sendo uma *Normative Description* em UFO.

Axiomas:

A1. Se uma medida derivada $mdd-dv$ é usada para quantificar um elemento indiretamente mensurável $elm-im$ e $mdd-dv$ deriva de uma medida mdd , então existe um elemento mensurável elm que é quantificado por mdd e é subelemento de $elm-im$

$$\begin{aligned} & (\forall mdd-dv \in \text{Derived Measure}, elm-im \in \text{Indirectly Measurable Entity}, mdd \in \text{Measure}) \\ & \quad (\text{quantifies}(mdd-dv, elm-im) \wedge \text{derivedFrom}(mdd-dv, mdd) \rightarrow \\ & \quad (\exists elm \in \text{Measurable Element}) \text{quantifies}(mdd, elm) \wedge \text{subElement}(elm-im, elm)) \end{aligned}$$

A2. Se uma medida derivada $md-dv$ é calculada pela fórmula fm e esta fórmula usa a medida md , então $md-dv$ deriva da medida md .

$$\begin{aligned} & (\forall md-dv \in \text{Derived Measure}, fm \in \text{Formula}, md \in \text{Measure}) \\ & \quad (\text{calculatedBy}(md-dv, fm) \wedge \text{uses}(fm, md) \rightarrow \text{derivedFrom}(md-dv, md)) \end{aligned}$$

A3. Se uma medida md é expressa em uma unidade de medida $un-md$ e md tem uma escala esc , então esc deve ser particionada de acordo com $un-md$.

$$\begin{aligned} & (\forall md \in \text{Measure}, un-md \in \text{Measure Unit}, esc \in \text{Scale}) \\ & \quad (\text{isExpressedIn}(md, un-md) \wedge \text{has}(md, esc) \rightarrow \text{partitionedAccordingTo}(esc, un-md)) \end{aligned}$$

A4. Se uma medida derivada $md-dv$ deriva de uma medida md , então md é correlata à $md-dv$.

$$(\forall md-dv \in \text{Derived Measure}, md \in \text{Measure}) (\text{derivedFrom}(md-dv, md) \rightarrow \text{correlatedMeasure}(md, md-dv))$$

A5. Se uma medida md é usada para caracterizar um tipo de entidade mensurável tem e md é usada para quantificar um elemento mensurável elm , então elm deve caracterizar tem .

$$\begin{aligned} & (\forall md \in \text{Measure}, tem \in \text{Measurable Entity Type}, elm \in \text{Measurable Element}) \\ & \quad (\text{usedToCharacterize}(md, tem) \wedge \text{usedToQuantify}(md, elm) \rightarrow \text{characterizedBy}(tem, elm)) \end{aligned}$$

3.1.1.3 Subontologia Planejamento de Medição (*Measurement Planning*)

Esta subontologia aborda o planejamento da medição, que consiste em definir por que a medição deve ser realizada, o que se deseja medir e como medir. Assim, esta subontologia trata da definição de objetivos de medição, necessidades de informação e a identificação das medidas e procedimentos de medição que devem ser usados para satisfazer as necessidades de informação e apoiar o alcance dos objetivos de medição.

Questões de Competência:

As questões de competência que esta subontologia visa responder são:

- QC1. A partir de quais objetivos um objetivo de medição é definido?
- QC2. Quais são as necessidades de informação identificadas a partir dos objetivos de medição?
- QC3. A que elemento mensurável uma necessidade de informação se refere?
- QC4. A que entidade mensurável uma necessidade de informação se refere?
- QC5. Que medidas são selecionadas para satisfazer uma necessidade de informação?
- QC6. Que medidas são usadas como indicadores para monitorar o alcance de um objetivo de medição?
- QC7. Que procedimentos de medição são selecionados para as medidas selecionadas no planejamento de medição?

Modelo Conceitual

O modelo conceitual da subontologia *Measurement Planning* é apresentado na Figura 3.4. Visando diminuir a poluição visual, não são apresentados os conceitos de UFO a partir dos quais são especializados alguns dos conceitos discutidos em subontologias anteriores.

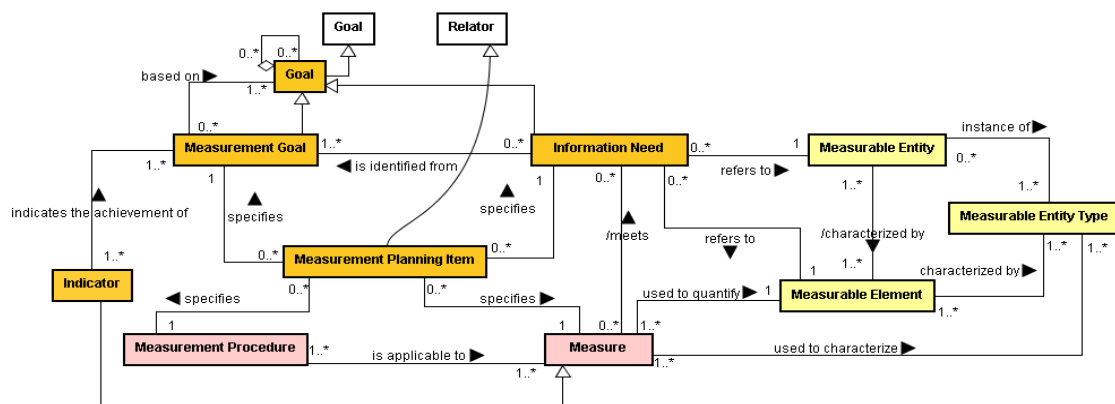


Figura 3.4 – Modelo conceitual da subontologia *Measurement Planning*.

Descrição do Modelo:

Um **Objetivo de Medição** (*Measurement Goal*) é um objetivo definido com base em um objetivo mais geral (*Goal*) e é utilizado como ponto de partida para guiar a identificação das medidas necessárias em um determinado contexto. Por exemplo, em um consultório, um médico pode ter como objetivo atender bem seus pacientes. A partir desse objetivo geral, podem-se estabelecer objetivos de medição, como, por exemplo, verificar o estado de saúde do paciente. Objetivos podem ser decompostos em outros. Por exemplo, verificar se o paciente tem um bom índice de massa corporal pode ser um subobjetivo de verificar o estado de saúde do paciente. A partir de um objetivo de medição podem ser identificadas **Necessidades de Informação** (*Information Need*), que determinam quais informações devem ser atendidas por medidas para que seja possível verificar se o objetivo de medição desejado é alcançado. Considerando o exemplo dado, conhecer o índice de massa corporal do paciente pode ser uma necessidade de informação identificada a partir de verificar o índice de massa corporal do paciente. Necessidades de informação também podem ser decompostas. Por exemplo, conhecer o peso do paciente e conhecer a altura do paciente são subnecessidades de conhecer o índice de massa corporal do paciente. Necessidades de informação referem-se a Elementos Mensuráveis (*Measurable Element*) e a Entidades Mensuráveis (*Measurable Entity*). Por exemplo, a necessidade de informação conhecer o índice de massa corporal do paciente refere-se ao elemento mensurável índice de massa corporal de uma entidade do tipo Pessoa (o paciente João, por exemplo). Medidas atendem necessidades de informação, levando-se em consideração os elementos mensuráveis e entidades mensuráveis relacionados a elas. Por exemplo, a medida Índice de Massa Corporal em quilos/metros² pode ser usada para atender à necessidade de informação conhecer o índice de massa corporal do paciente. Medidas que são usadas diretamente na análise do alcance de objetivos são chamadas **Indicadores** (*Indicator*). Por exemplo, uma vez que no planejamento de medição definiu-se que a medida Índice de Massa Corporal em quilos/metros² é usada para indicar o alcance ao objetivo de medição verificar se o paciente tem um bom índice de massa corporal, nesse contexto, ela desempenha o papel de Indicador. Por fim, um **Item de Planejamento de Medição** (*Measurement Planning Item*) conecta um Objetivo de Medição (*Measurement Goal*), uma Necessidade de Informação (*Information Need*), uma Medida (*Measure*) e um Procedimento de Medição (*Measurement Procedure*), indicando que a medida atende a necessidade de informação e que esta foi

identificada a partir do objetivo de medição. Dessa forma, faz-se o alinhamento da medição com os objetivos que se deseja alcançar.

Fundamentação:

Em UFO, *Intention* é alguma coisa que um agente intenciona. Por exemplo, um médico tem a intenção de cuidar da saúde dos seus pacientes. O conteúdo proposicional de uma *Intention* é um *Goal*. No exemplo citado, um conteúdo proposicional possível seria verificar o estado de saúde do paciente. Assim, **Objetivo de Medição** (*Measurement Goal*) é um *Goal* em UFO. Similarmente, um **Objetivo** (*Goal*) é um *Goal* em UFO. **Necessidade de Informação** (*Information Need*) é um *Goal* que visa à obtenção de alguma informação e, portanto, está sempre relacionado a um Elemento Mensurável (*Measurable Element*) e uma Entidade Mensurável (*Measurable Entity*). **Indicador** (*Indicator*) é um papel que uma medida desempenha quando indica o alcance a algum objetivo, sendo, assim, um *Role* em UFO. Por fim, **Item de Planejamento de Medição** (*Measurement Planning Item*) é um *Relator* que conecta as entidades envolvidas em um item de planejamento de medição.

Axiomas:

A1. Se um item de planejamento de medição *ipm* especifica o objetivo de medição *obj* e a necessidade de informação *inf*, então *inf* deve ser identificada a partir de *obj*.

$$(\forall ipm \in \text{Measurement Planning Item}, obj \in \text{Measurement Goal}, inf \in \text{Information Need}) \\ (specifies(ipm, obj) \wedge specifies(ipm, inf)) \rightarrow (isIdentifiedFrom(inf, obj))$$

A2. Se um item de planejamento de medição *ipm* especifica a medida *mdd* e o procedimento de medição *prc*, então *prc* deve ser aplicável a *mdd*.

$$(\forall ipm \in \text{Measurement Planning Item}, mdd \in \text{Measure}, prc \in \text{Measurement Procedure}) \\ (specifies(ipm, mdd) \wedge specifies(ipm, prc)) \rightarrow isApplicableTo(prc, mdd)$$

A3. Uma medida *mdd* satisfaz uma necessidade de informação *inf* se e somente se existe um item de planejamento de medição *ipm* que especifica a *inf* e a medida *mdd*.

$$(\forall mdd \in \text{Measure}, inf \in \text{Information Need}) \text{meets}(mdd, inf) \leftrightarrow (\exists ipm \in \text{Measurement Planning Item}) \\ specifies(ipm, inf) \wedge specifies(ipm, mdd)$$

A4. Se uma necessidade de informação *inf* refere-se a uma entidade mensurável *em* e a um elemento mensurável *elm*, então o elemento mensurável *elm* deve caracterizar a entidade mensurável *em*.

$$(\forall inf \in Information\ Need, em \in Measurable\ Entity, elm \in Measurable\ Element) \\ (refersTo(inf, em) \wedge refersTo(inf, elm)) \rightarrow characterizedBy(em, elm))$$

A5. Se um item de planejamento de medição *ipm* especifica uma medida *mdd* e uma necessidade de informação *inf* que se refere ao elemento mensurável *elm*, então a medida *mdd* deve quantificar o elemento mensurável *elm*.

$$(\forall ipm \in Measurement\ Planning\ Item, mdd \in Measure, inf \in Information\ Need, elm \in Measurable\ Element) \\ (specifies(ipm, mdd) \wedge specifies(ipm, inf) \wedge refersTo(inf, elm)) \rightarrow (usedToQuantify(mdd, elm))$$

A6. Se um item de planejamento de medição *ipm* especifica uma medida *mdd* e uma necessidade de informação *inf* que se refere à entidade mensurável *em* do tipo de entidade mensurável *tem*, então a medida *mdd* deve ser usada para caracterizar *tem*.

$$(\forall ipm \in Measurement\ Planning\ Item, mdd \in Measure, inf \in Information\ Need, em \in Measurable\ Entity, \\ tem \in Measurable\ Entity\ Type) (specifies(ipm, mdd) \wedge specifies(ipm, inf) \wedge refersTo(inf, em) \wedge \\ instanceOf(em, tem)) \rightarrow usedToCharacterize(mdd, tem))$$

A7. Para toda medida *mdd* que desempenha o papel de indicador deve existir um item de planejamento de medição *ipm* que especifica *mdd* e o objetivo de medição *obj*, cujo alcance é indicado por *mdd*.

$$(\forall mdd \in Measure) (mdd \in Indicator) \rightarrow (\exists ipm \in Measurement\ Planning\ Item, obj \in Measurement\ Goal) \\ (includes(ipm, mdd) \wedge includes(ipm, obj) \wedge indicatesTheAchievementOf(mdd, obj))$$

A8. Se uma medida *mdd* é um indicador do alcance de um objetivo de medição *obj*, então deve existir uma necessidade de informação *inf* identificada com base no objetivo de medição *obj* que é atendida pela medida *mdd*.

$$(\forall mdd \in Indicator, obj \in Measurement\ Goal) (indicatesTheAchievementOf(mdd, obj)) \rightarrow \\ (\exists inf \in Information\ Need) isIdentifiedBasedOn(inf, obj) \wedge meets(mdd, inf))$$

A9. Se um objetivo *obj2* é um subobjetivo do objetivo de medição *obj1*, então *obj2* é um subobjetivo de medição.

$$(\forall obj2 \in Goal, obj1 \in Measurement\ Goal) (subgoal(obj2, obj1) \rightarrow subMeasurementGoal(obj2))$$

A10. Se um objetivo *obj* é um subobjetivo de uma necessidade de informação *inf*, então *obj* é uma subnecessidade de informação.

$$(\forall obj \in Goal, inf \in Information\ Need) (subgoal(obj, inf) \rightarrow subneed(obj))$$

A11. Se uma necessidade de informação *inf2* é uma subnecessidade de informação de *inf1* e *inf1* é identificada a partir de *obj*, então *inf2* é identificada a partir de *obj*.

$$(\forall inf1, inf2 \in Information\ Need, obj \in Measurement\ Goal) \\ (subneedinf2, inf1) \wedge isIdentifiedFrom(inf1, obj) \rightarrow isIdentifiedFrom(inf2, obj))$$

A12. Se uma medida *mdd* atende uma necessidade de informação *inf2* que é subnecessidade da necessidade de informação *inf1*, então *mdd* atende à necessidade *inf1*.

$$(\forall inf1, inf2 \in Information\ Need, mdd \in Measure) \\ (meets(mdd, inf2) \wedge subneed(inf2, inf1) \rightarrow meets(mdd, inf1))$$

3.1.1.4 Subontologia Medição (*Measurement*)

Esta subontologia trata de aspectos relacionados à medição propriamente dita, ou seja, a coleta de dados para as medidas.

Questões de Competência:

As questões de competência que esta subontologia visa responder são:

QC1. Baseado em qual item de planejamento de medição é realizada uma medição?

QC2. Qual entidade mensurável é medida em uma medição?

QC3. Qual elemento mensurável da entidade mensurável é medido em uma medição?

QC4. Que medida é aplicada em uma medição?

QC5. Qual procedimento de medição é adotado em uma medição?

QC6. Qual é o valor medido em uma medição?

Modelo Conceitual:

O modelo conceitual da subontologia *Measurement* é apresentado na Figura 3.5. Alguns relacionamentos apresentados nos modelos anteriores foram omitidos para diminuir a poluição visual do modelo.

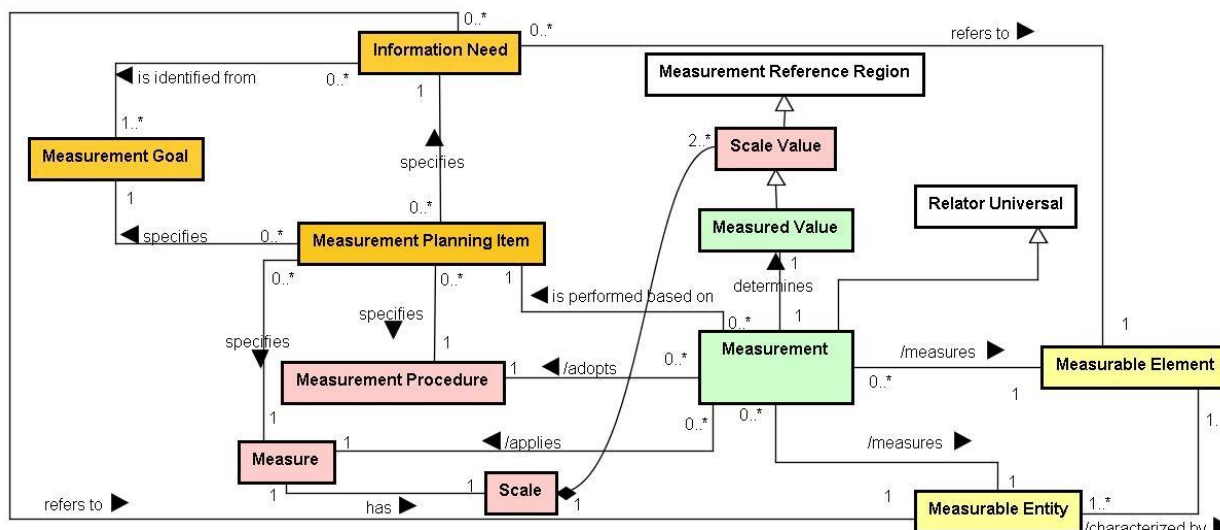


Figura 3.5 – Modelo conceitual da subontologia *Measurement*.

Descrição do Modelo:

Medição (*Measurement*) é uma ação, realizada com base em um Item de Planejamento de Medição (*Measurement Planning Item*), que mede um Elemento Mensurável (*Measurable Element*) de uma Entidade Mensurável (*Measurable Entity*), aplicando-se uma Medida (*Measure*) e adotando-se um Procedimento de Medição (*Measurement Procedure*) para obter um **Valor Medido** (*Measured Value*). Por exemplo, a medição do elemento mensurável peso da entidade mensurável João aplicando-se a medida peso em quilos, adotando-se um procedimento de medição e produzindo o valor medido 75 kg. A medida aplicada e o procedimento de medição adotado em uma medição são a medida e o procedimento de medição incluídos no item de planejamento de medição usado. Analogamente, o elemento mensurável e a entidade mensurável medidos são aqueles referidos pela necessidade de informação especificada no item de planejamento usado.

Fundamentação:

Medição (*Measurement*) é uma *Action* em UFO. A ação de medir é realizada para associar um Valor Medido (*Measured Value*) a um Item de Planejamento de Medição (*Measurement Planning Item*). À luz de UFO, significa dizer que a ação leva a uma relação material que conecta as entidades envolvidas na medição. De acordo com UFO, relações materiais requerem *Relators* para que sejam estabelecidas. Assim, há o conceito Medição, como uma *Action* e há, também, o conceito Medição, como um *Relator* em UFO. Uma vez que há equivalência um para um entre a *Action* e o *Relator* Medição (uma *action* Medição leva a um *relator* Medição), no modelo representa-se apenas o último.

Valor Medido (*Measured Value*) é um valor da escala da medida (*Scale Value*) e, dessa forma, é uma *Measurement Quality Region*. Pode-se dizer, ainda, que um Valor Medido é um papel (*Role* em UFO) que um Valor de Escala desempenha, quando associado a um Item de Planejamento de Medição (*Measurement Planning Item*) (e, conseqüentemente, a um Elemento Mensurável (*Measurable Element*)) em uma Medição (*Measurement*).

Axiomas:

A1. Se uma medição é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica a medida *mdd*, então *mdc* aplica *mdd*.

$$(\forall mdc \in Measurement, ipm \in Measurement\ Planning\ Item, mdd \in Measure) \\ (isPerformedBasedOn(mdc, ipm) \wedge specifies(ipm, mdd)) \rightarrow applies(mdc, mdd)$$

A2. Se uma medição *mdc* é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica o procedimento de medição *prc*, então *mdc* adota *prc*.

$$(\forall mdc \in Measurement, ipm \in Measurement\ Planning\ Item, prc \in Measurement\ Procedure) \\ (isPerformedBasedOn(mdc, ipm) \wedge specifies(ipm, prc)) \rightarrow adopts(mdc, prc)$$

A3. Se uma medição *mdc* é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica a necessidade de informação *inf* que refere-se ao elemento mensurável *elm*, então *mdc* mede *elm*.

$$(\forall mdc \in Measurement, ipm \in Measurement\ Planning\ Item, inf \in Information\ Need, elm \in Measurable\ Element) \\ (isPerformedBasedOn(mdc, ipm) \wedge specifies(ipm, inf) \wedge refersTo(inf, elm)) \rightarrow measures(mdc, elm)$$

A4. Se uma medição *mdc* é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica a necessidade de informação *inf* que refere-se à entidade mensurável *em*, então *mdc* mede *em*.

$(\forall mdc \in Measurement, ipm \in Measurement\ Planning\ Item, inf \in Information\ Need, em \in Measurable\ Entity)$
 $(isPerformedBasedOn(mdc, ipm) \wedge specifies(ipm, inf) \wedge refersTo(inf, em)) \rightarrow measures(mdc, em)$

A5. Se uma medição *mdc* aplica uma medida *mdd* que tem escala *esc* e *mdc* determina um valor medido *vl*, então *vl* deve ser um valor da escala *esc*.

$(\forall mdc \in Measurement, mdd \in Measure, esc \in Scale, vl \in Measured\ Value)$
 $(applies(mdc, mdd) \wedge has(mdd, esc) \wedge determines(mdc, vl) \rightarrow isPartOf(vl, esc))$

3.1.1.5 Subontologia Análise de Medição (*Measurement Analysis*)

Esta subontologia trata de aspectos relacionados à análise da medição, envolvendo a análise de valores medidos e a obtenção de resultados a partir dessa análise.

Questões de Competência:

As questões de competência que esta subontologia visa responder são:

- QC1. Baseado em quais itens de planejamento da medição é realizada uma análise?
- QC2. Quais medidas são consideradas em uma análise de medição?
- QC3. Quais valores medidos são analisados em uma análise de medição?
- QC4. Qual o resultado de uma análise de medição?
- QC5. Qual entidade mensurável é caracterizada por um resultado de análise?
- QC6. Um resultado de medição fornece informações para quais objetivos de medição?

Modelo Conceitual:

O modelo conceitual da subontologia *Measurement Analysis* é apresentado na Figura 3.6. Alguns relacionamentos apresentados nos modelos anteriores foram omitidos para diminuir a poluição visual do modelo.

um *Relator* em UFO, sendo que no modelo representa-se apenas o último. **Resultado da Análise** (*Analysis Result*), por sua vez, é um *Social Object* em UFO.

Axiomas:

A1. Se uma análise de medição *an-md* considera uma medida *mdd*, então deve existir um item de planejamento *ipm* no qual *an-md* é baseada, que especifique *mdd*.

$$(\forall an-md \in \text{Measurement Analysis}, mdd \in \text{Measure}) \text{considers}(an-md, mdd) \rightarrow (\exists ipm \in \text{Measurement Planning Item}) \text{isPerformedBasedOn}(an-md, ipm) \wedge \text{specifies}(ipm, mdd)$$

A2. Se uma análise de medição *an-md* é realizada com base em um item de planejamento de medição *ipm* e *an-md* analisa um valor medido *vl* determinado em uma medição *mdc*, então *mdc* deve ter sido realizada com base em *ipm*.

$$(\forall an-md \in \text{Measurement Analysis}, ipm \in \text{Measurement Planning Item}, vl \in \text{Measured Value}, mdc \in \text{Measurement}) (\text{isPerformedBasedOn}(an-md, ipm) \wedge \text{analyzes}(an-md, vl) \wedge \text{determines}(mdc, vl) \rightarrow \text{isPerformedBasedOn}(mdc, ipm))$$

A3. Se uma análise de medição *an-md* é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica o objetivo de medição *obj*, então *an-md* considera *obj*.

$$(\forall an-md \in \text{Measurement Analysis}, ipm \in \text{Measurement Planning Item}, obj \in \text{Measurement Goal}) (\text{isPerformedBasedOn}(an-md, ipm) \wedge \text{specifies}(ipm, obj)) \rightarrow \text{considers}(an-md, obj)$$

A4. Se uma análise de medição *an-md* é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica a medida *mdd*, então *an-md* considera *mdd*.

$$(\forall an-md \in \text{Measurement Analysis}, ipm \in \text{Measurement Planning Item}, mdd \in \text{Measure}) (\text{isPerformedBasedOn}(an-md, ipm) \wedge \text{specifies}(ipm, mdd)) \rightarrow \text{considers}(an-md, mdd)$$

A5. Se uma análise de medição *an-md* é realizada com base em um item de planejamento de medição *ipm* e *ipm* especifica a necessidade de informação *inf* que se refere à entidade mensurável *em*, então *an-md* caracteriza *em*.

$$(\forall an-md \in \text{Measurement Analysis}, ipm \in \text{Measurement Planning Item}, inf \in \text{Information Need}, em \in \text{Measurable Entity}) (\text{isPerformedBasedOn}(an-md, ipm) \wedge \text{specifies}(ipm, inf) \wedge \text{refersTo}(inf, em)) \rightarrow \text{characterizes}(an-md, em)$$

A6. Se uma análise de medição *an-md* produz um resultado de análise *rst* e caracteriza uma entidade mensurável *em*, então *rst* caracteriza a entidade *em*.

$$(\forall an-md \in \text{Measurement Analysis}, rst \in \text{Analysis Result}, em \in \text{Measurable Entity}) \\ (\text{produces}(an-md, rst) \wedge \text{characterizes}(an-md, em) \rightarrow \text{characterizes}(rst, em))$$

A7. Se uma análise de medição *an-md* considera o objetivo de medição *obj* e produz um resultado de análise *rst*, o resultado *rst* provê informação para o objetivo de medição *obj*.

$$(\forall an-md \in \text{Measurement Analysis}, obj \in \text{Measurement Goal}, rst \in \text{Analysis Result}) \\ (\text{considers}(an-md, obj) \wedge \text{produces}(an-md, rst) \rightarrow \text{providesInformationTo}(rst, obj))$$

A8. Se um resultado de análise *rst* produzido em uma análise de medição *an-md* provê informações para o objetivo de medição *obj*, então existe uma medida *mdd* que é considerada na análise de medição *an-md* e indica o alcance de *obj*.

$$(\forall rst \in \text{Analysis Result}, an-md \in \text{Measurement Analysis}, obj \in \text{Measurement Goal}, mdd \in \text{Measure}) \\ (\text{produces}(an-md, rst) \wedge \text{providesInformationTo}(rst, obj) \rightarrow (\exists mdd \in \text{Measure}) \text{considers}(an-md, mdd) \wedge \\ \text{indicatesTheAchievementOf}(mdd, obj))$$

A9. Se uma análise de medição *an-md* considera uma medida *mdd* e analisa um valor medido *vl*, então *vl* deve ter sido determinado em uma medição *mdc* que aplica a medida *mdd*.

$$(\forall an-md \in \text{Measurement Analysis}, vl \in \text{Measured Value}, mdc \in \text{Measurement}, mdd \in \text{Measure}) \\ (\text{analyzes}(an-md, vl) \wedge \text{considers}(an-md, mdd) \wedge \text{determines}(mdc, vl) \rightarrow \text{applies}(mdc, mdd))$$

A10. Se uma análise de medição *an-md* caracteriza uma entidade mensurável *em* e analisa um valor medido *vl* determinado em uma medição *mc*, então a medição *mc* deve ter medido a entidade mensurável *em*.

$$(\forall an-md \in \text{Measurement Analysis}, vl \in \text{Measured Value}, em \in \text{Measurable Entity}, mc \in \text{Measurement}) \\ (\text{characterizes}(an-md, em) \wedge \text{analyzes}(an-md, vl) \wedge \text{determines}(mc, vl) \rightarrow \text{measures}(mc, em))$$

A11. Se uma análise de medição *an-mdc* analisa um valor medido *vl* determinado em uma medição *mdc* que medem a entidade mensurável *em*, então *an-mdc* caracteriza *em*.

$$(\forall an\text{-}mdc \in \textit{Measurement Analysis}, vl \in \textit{Measured Value}, mdc \in \textit{Measurement}, em \in \textit{Measurable Entity})$$

$$(\textit{analyzes}(an\text{-}mdc, vl) \wedge \textit{determines}(mdc, vl) \wedge \textit{measures}(mdc, em)) \rightarrow \textit{characterizes}(an\text{-}mdc, em)$$

3.2 Measurement Task Ontology - MTO

A ontologia de tarefa de medição (*Measurement Task Ontology* – MTO) definida neste trabalho é uma representação conceitual do processo de medição e, dessa forma, não trata de uma única tarefa. O termo ontologia de tarefa foi adotado para ficar em consonância com a classificação de ontologias definida por Guarino (1998), que é bastante utilizada na área de ontologias. Sendo uma ontologia de tarefa, MTO deve responder as seguintes questões de competência: (i) Quais são as atividades do processo de medição? (ii) Quem são os responsáveis por executá-las? (iii) Como as atividades são decompostas em subatividades? (iv) Como é fluxo que guia essas atividades? (v) Quais são os insumos e produtos de cada atividade?

Como proposto em (MARTINS; FALBO, 2008), ontologias de tarefa devem ser representadas através de dois tipos de modelos: modelo comportamental, cuja função é capturar a decomposição de tarefas, seus insumos e produtos, e como os papéis atuam em sua execução; e modelo estrutural, cuja função é representar as propriedades e relações dos insumos e produtos das atividades e outras entidades necessárias para caracterizar as tarefas. MTO é representada através desses dois tipos de modelos, tendo sido utilizados diagramas de atividades e diagramas de classes UML para modelar, respectivamente, a visão comportamental e estrutural.

O modelo estrutural de MTO é o modelo de classes integrado dos modelos conceituais das subontologias de COM. Na próxima seção são apresentados os modelos comportamentais de MTO.

3.2.1 Modelos Comportamentais da *Measurement Task Ontology*

Conforme discutido no Capítulo 2, o processo de medição é composto por três atividades principais: planejar medição, executar medição e analisar medição. A Figura 3.7

apresenta o modelo comportamental geral de MTO, onde é possível identificar essas atividades, os papéis envolvidos e os principais insumos e produtos. Todas as atividades presentes no diagrama são ações complexas (*complex actions*) em UFO, uma vez que envolvem a participação de diferentes objetos e agentes. Em UML, essas atividades correspondem a *Call Behavior Actions* e suas decomposições serão apresentadas mais adiante.

Nos modelos comportamentais de MTO, os estereótipos dos fluxos de objetos visam capturar algumas das distinções feitas em UFO-C a respeito da participação de objetos em ações, a saber: *creation*, que indica que um objeto é criado pela ação; *change*, que indica que alguma propriedade do objeto é alterada; e *usage* que indica que o objeto é usado, sem que haja mudança em suas propriedades. Quando o fluxo não é estereotipado, significa que o objeto é apenas usado durante a atividade.

Nas descrições dos modelos comportamentais de MTO usa-se **negrito** para os conceitos do modelo estrutural em Português, os quais são seguidos dos termos equivalentes em Inglês (os termos usados nos modelos) em *itálico e com fonte diferenciada*. Os nomes de atividades e subatividades são apresentados em Português “entre aspas” seguidos do nome equivalente em Inglês em *itálico*.

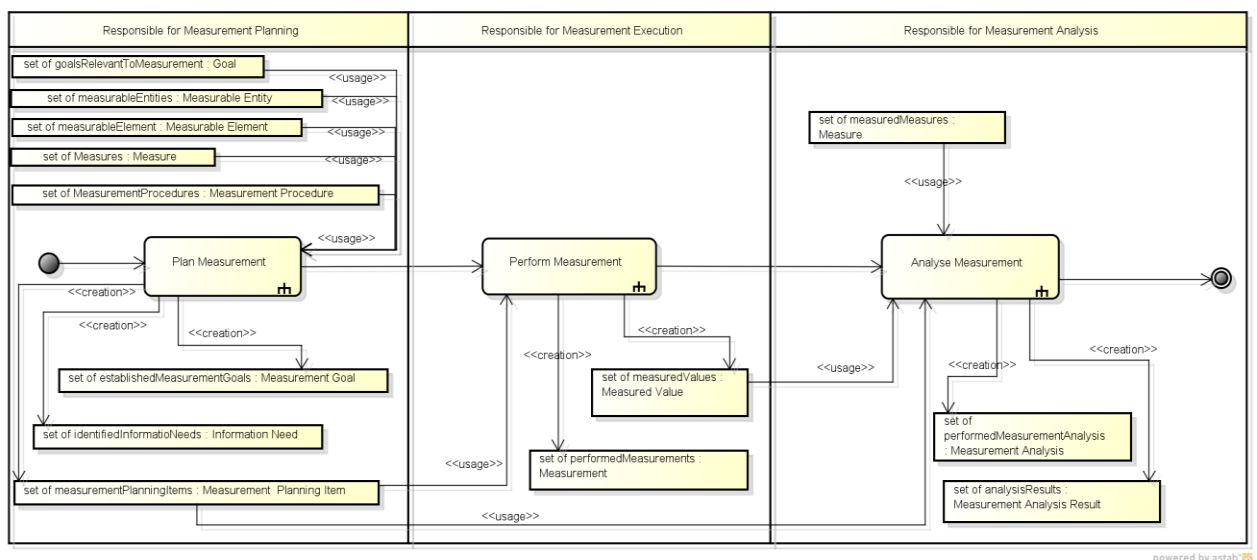


Figura 3.7 – Modelo comportamental geral de MTO.

O processo de medição tem início com o responsável pelo planejamento da medição realizando a atividade "Planejar Medição" (*Plan Measurement*). Nesta atividade, **Objetivos** (*Goals*) relevantes para a medição são utilizados como base para a identificação das **Necessidades de Informação** (*Information Need*) que devem ser atendidas e identificação das **Medidas** (*Measures*) necessárias. Também é nesta atividade que **Procedimentos de**

Medição (*Measurement Procedures*), que orientam a coleta de dados para as medidas, são estabelecidos. Cada combinação de **Objetivo de Medição**, **Necessidade de Informação**, **Medida** e **Procedimento de Medição** é organizada em um **Item de Planejamento de Medição** (*Measurement Planning Item*). Na atividade seguinte, "Executar Medição" (*Perform Measurement*), o responsável pela execução da medição seleciona, dentre os itens de planejamento de medição, aquele que será considerado e realiza a medição, obtendo **Valores Medidos** (*Measured Value*) e registrando as **Medições** (*Measurement*) realizadas. Por fim, na atividade "Analisar Medição" (*Measurement Analysis*), o responsável pela análise de medição seleciona as **Medidas** (*Measures*) para as quais foram realizadas medições e os **Valores Medidos** (*Measured Values*) que deseja analisar. Realiza as análises dos dados considerando os **Objetivos de Medição** (*Measurement Goals*) estabelecidos nos itens selecionados, registra as **Análises de Medição** (*Measurement Analysis*), em **Resultados de Análise** (*Analysis Results*) e relata os resultados para as partes interessadas. A seguir, as atividades presentes na Figura 3.7 são detalhadas.

3.2.1.1 Planejar Medição (*Plan Measurement*)

Na Figura 3.8 é apresentado o modelo comportamental de “Planejar Medição” (*Plan Measurement*).

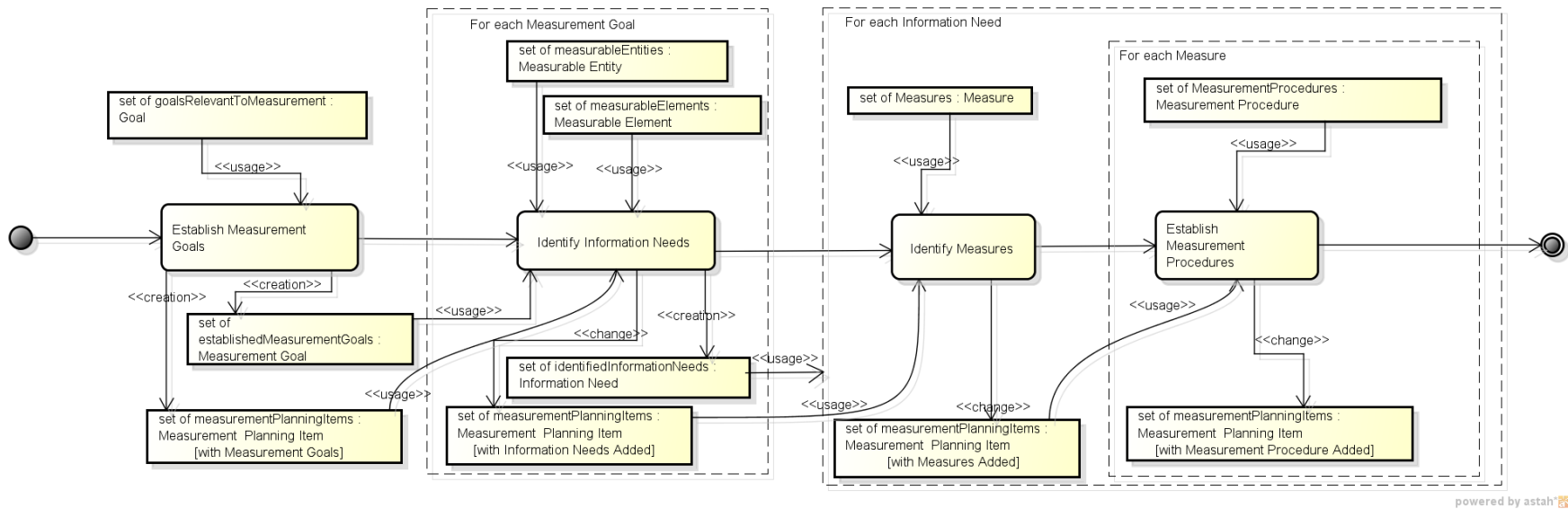


Figura 3.8 - Modelo comportamental de “Planejar Medição”.

O planejamento da medição tem início na subatividade "Estabelecer Objetivos de Medição" (*Establish Measurement Goals*), na qual, a partir de **Objetivos** (*Goals*) relevantes para a medição, são estabelecidos os **Objetivos de Medição** (*Measurement Goals*). Para cada **Objetivo de Medição** (*Measurement Goal*) estabelecido, na subatividade "Identificar Necessidades de Informação" (*Identify Information Needs*), são identificadas as **Necessidades de Informação** (*Information Needs*) que deverão ser atendidas, bem como a **Entidade Mensurável** (*Measurable Entity*) e o **Elemento Mensurável** (*Measurable Element*) que devem ser medidos para atender cada **Necessidade de Informação** (*Information Need*). Para cada **Necessidade de Informação** (*Information Need*) identificada, o responsável pelo planejamento da medição deve "Identificar Medidas" (*Identify Measures*) que deverão ser usadas para medir o **Elemento Mensurável** (*Measurable Element*) e a **Entidade Mensurável** (*Measurable Entity*) necessários para atender a **Necessidade de Informação** (*Information Need*) selecionada. Por fim, para cada **Medida** (*Measure*) identificada, é necessário "Estabelecer Procedimentos de Medição" (*Establishing Measurement Procedures*).

O principal resultado da atividade "Planejar Medição" (*Plan Measurement*) consiste em um conjunto de **Objetivos de Medição** (*Measurement Goals*), **Necessidades de Informação** (*Information Needs*), **Medidas** (*Measures*) e **Procedimentos de Medição** (*Measurement Procedures*) relacionados. Cada combinação de **Objetivo de Medição** (*Measurement Goal*), **Necessidade de Informação** (*Information Need*), **Medida** (*Measure*) e **Procedimento de Medição** (*Measurement Procedure*) é conectada em um **Item de Planejamento de Medição** (*Measurement Planning Item*). Os elementos relacionados em um **Item de Planejamento de Medição** (*Measurement Planning Item*) são conectados ao longo das subatividades do planejamento da medição. Os **Itens de Planejamento de Medição** (*Measurement Planning Items*) são criados na subatividade "Estabelecer Objetivos de Medição" (*Establish Measurement Goals*) e os **Objetivos de Medição** (*Measurement Goals*) estabelecidos nessa subatividade são os únicos elementos conectados aos **Itens de Planejamento de Medição** (*Measurement Planning Items*) até então (*Measurement Planning [with Measurement Goal]*). A cada subatividade realizada, novos elementos são conectados. Dessa forma, na próxima subatividade ("Identificar Necessidades de Informação" - *Identify Information Needs*), as **Necessidades de Informação** (*Information Needs*) identificadas são adicionadas aos **Itens de Planejamento de Medição** (*Measurement Planning Items*) definidos (*Measurement Planning [with Information Need Added]*). Analogamente, **Medidas**

(*Measures*) e **Procedimentos de Medição** (*Measurement Procedures*) são conectados nas subatividades seguintes.

3.2.1.2 Executar Medição (*Perform Measurement*)

Na Figura 3.9 é apresentado o modelo comportamental de “Executar Medição” (*Perform Measurement*).

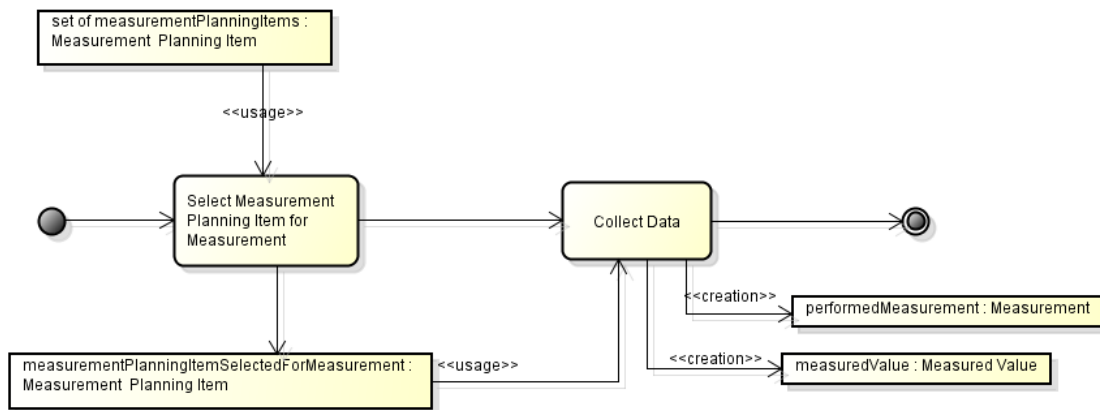


Figura 3.9 - Modelo comportamental de “Realizar Medição”.

Para iniciar a execução da medição, na subatividade "Selecionar Item de Planejamento para Medição" (*Select Measurement Planning Item for Measurement*), o responsável pela execução da medição seleciona o **Item de Planejamento de Medição** (*Measurement Planning Item*) que deverá ser considerado como base para a **Medição** (*Measurement*). Em seguida, durante a subatividade "Coletar Dados" (*Collect Data*), a **Medida** (*Measure*) especificada no **Item de Planejamento de Medição** (*Measurement Planning Item*) selecionado é usada para medir a **Entidade Mensurável** (*Measurable Entity*) referida pela **Necessidade de Informação** (*Information Need*) especificada no **Item de Planejamento de Medição** (*Measurement Planning Item*), aplicando-se o **Procedimento de Medição** (*Measurement Procedure*) determinado neste mesmo **Item de Planejamento de Medição** (*Measurement Planning Item*) e obtendo-se um **Valor Medido** (*Measured Value*).

3.2.1.3 Analisar Medição (*Analyze Measurement*)

Na Figura 3.10 é apresentado o modelo comportamental de “Analisar Medição” (*Analyze Measurement*).

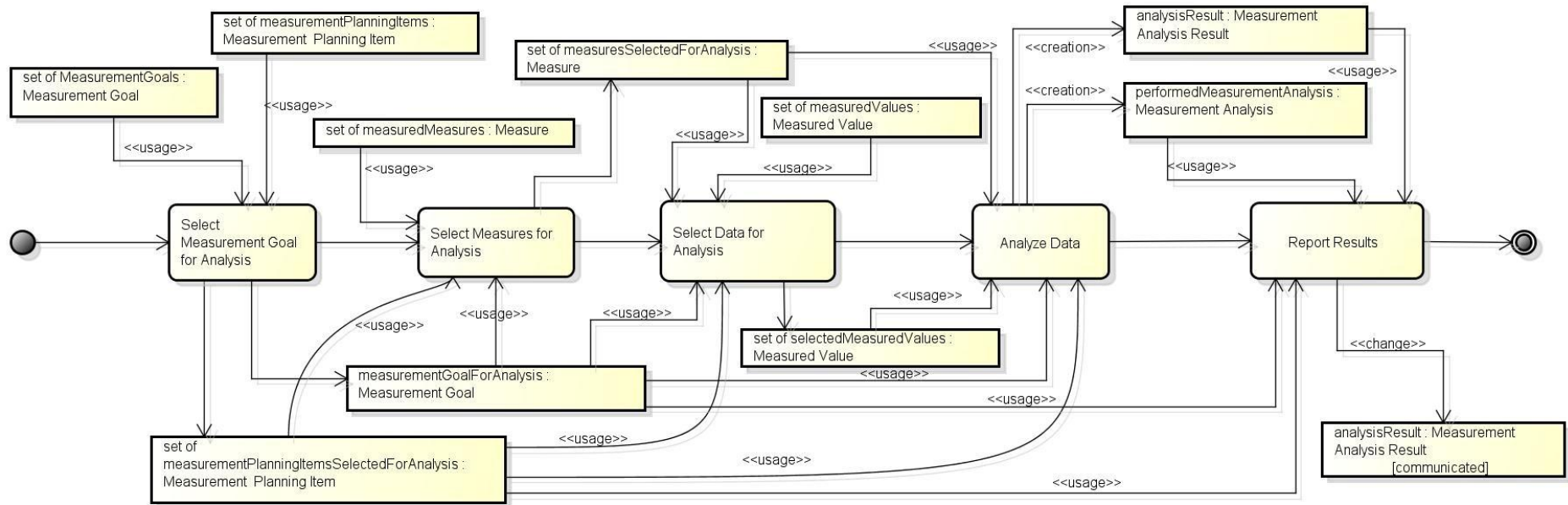


Figura 3.10 - Modelo comportamental de “Analisar Medição”.

Uma vez que dados tenham sido coletados, é possível analisá-los. Para isso, inicialmente, o responsável pela análise da medição deve "Selecionar Objetivo de Medição para Análise" (*Select Measurement Goal for Analysis*), identificando, dentre **Objetivos de Medição** (*Measurement Goals*) incluídos nos **Itens de Planejamento de Medição** (*Measurement Planning Items*), aquele para o qual se deseja obter informações com a análise dos dados. Em seguida, o responsável deve "Selecionar Medidas para Análise" (*Select Measures for Analysis*), indicando quais **Medidas** (*Measures*) serão consideradas na **Análise de Medição** (*Measurement Analysis*). As medidas selecionadas devem estar inclusas nos **Itens de Planejamento de Medição** (*Measurement Planning Items*) definidos, devem estar relacionadas ao **Objetivo de Medição** (*Measurement Goal*) selecionado e devem ter medições registradas. Mais de uma **Medida** (*Measure*) pode ser selecionada para análise. A terceira subatividade consiste em "Selecionar Dados para Análise" (*Select Data for Analysis*). Nela, o responsável pela análise seleciona quais **Valores Medidos** (*Measured Values*) para as **Medidas** (*Measures*) selecionadas serão utilizados na **Análise de Medição** (*Measurement Analysis*). Após a seleção, deve-se "Analisar Dados" (*Analyze Data*), registrando-se a **Análise de Medição** (*Measurement Analysis*) e produzindo-se um **Resultado da Análise** (*Measurement Analysis Result*). Por fim, na subatividade "Reportar Resultados" (*Report Results*), os **Resultados da Análise** (*Measurement Analysis Results*) são comunicados às partes interessadas.

Capítulo 4

Abordagem para Obtenção de Indicadores a partir de Dados Abertos

Neste capítulo é descrita a abordagem proposta neste trabalho para obtenção de indicadores a partir de dados abertos.. Para apoiar o uso da abordagem foi desenvolvida uma ferramenta, que é brevemente apresentada. Por fim, são apresentados alguns resultados do uso da abordagem considerando bases de dados abertos disponibilizadas pela ANAC e INFRAERO.

4.1 Abordagem para Obtenção de Indicadores a partir de Dados Abertos

A abordagem para obtenção de indicadores proposta neste trabalho foi definida tomando-se como base a *Core Ontology on Measurement* e a *Measurement Task Ontology* descritas no capítulo anterior. A *Core Ontology on Measurement* foi utilizada para fornecer o conhecimento sobre medição e o vocabulário adotado no procedimento. A *Measurement Task Ontology*, por sua vez, foi utilizada para nortear a definição do procedimento, orientando sobre as atividades que deveriam ser contempladas. Vale ressaltar que, embora a abordagem tenha sido definida considerando o uso de fontes de dados abertos, ela também pode ser utilizada para outras fontes de dados que não sejam de dados abertos.

Conforme mostra a Figura 4.1, o procedimento consiste em três atividades: Planejar Obtenção de Indicadores, Obter Dados e Analisar Dados. Essas atividades são compostas por subatividades e são descritas na sequência.

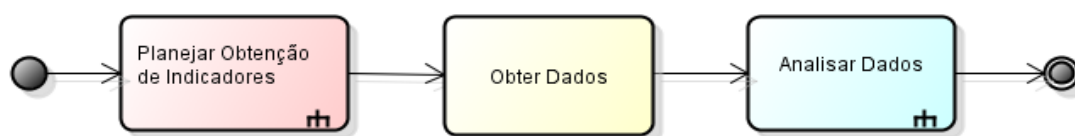


Figura 4.1 – Atividades da abordagem para obtenção de indicadores.

4.1.1 Planejar Obtenção de Indicadores

Nesta atividade é feito o planejamento para a obtenção dos indicadores. Para isso, as subatividades representadas na Figura 4.2 e descritas em seguida devem ser realizadas.

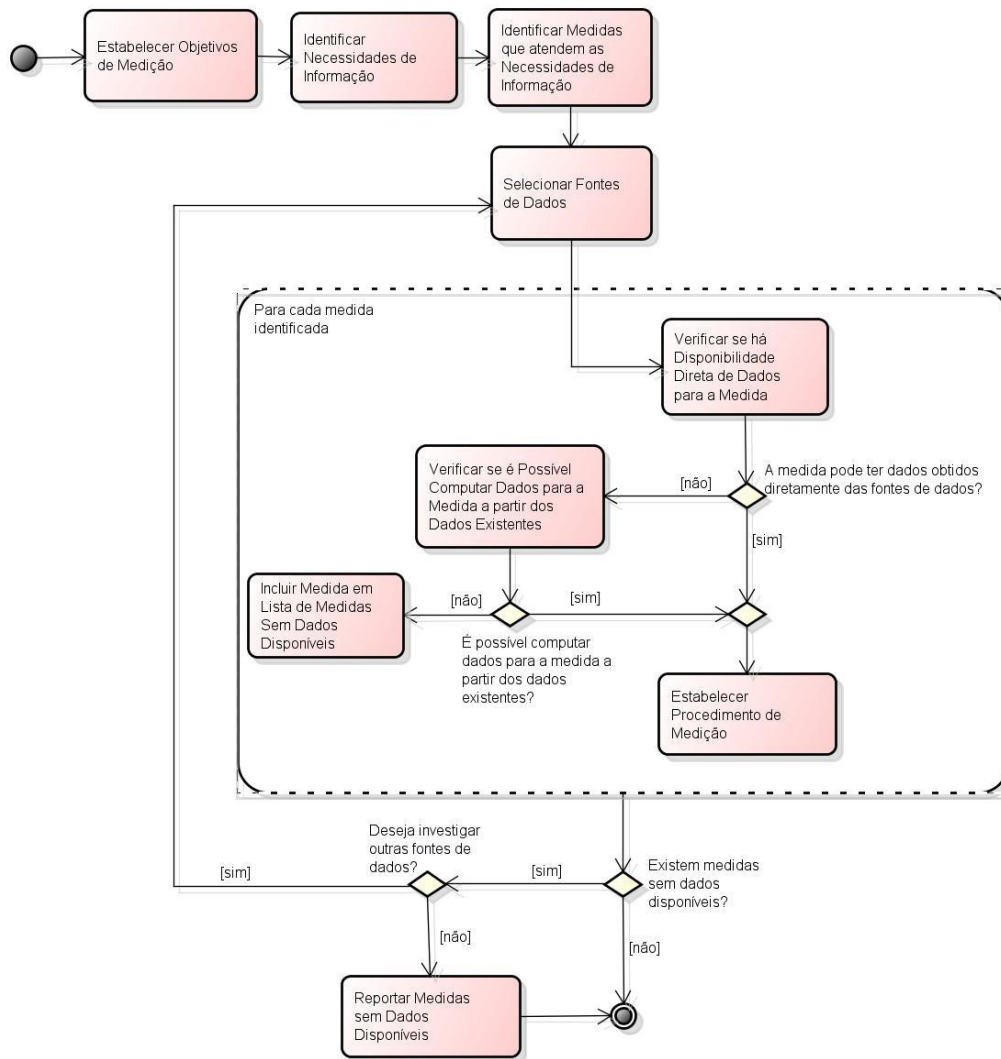


Figura 4.2 – Subatividades de “Planejar Obtenção de Indicadores”.

4.1.1.1 Estabelecer Objetivos de Medição

Nesta subatividade devem ser identificados os objetivos que se deseja alcançar a partir da análise de dados contidos nas fontes de dados utilizadas. Esses objetivos, ditos objetivos de medição, normalmente, são identificados com base em objetivos mais gerais (objetivos de negócio, por exemplo) e podem ser decompostos em subobjetivos.

Entrada: objetivos gerais.

Saída: objetivos de medição.

4.1.1.2 Identificar Necessidades de Informação

A partir dos objetivos de medição definidos, devem ser identificadas as necessidades de informação que devem ser atendidas para que o objetivo de medição seja alcançado. Uma forma simples de identificar necessidades de informação é, para cada objetivo estabelecido, fazer perguntas como ‘o que é necessário saber para que seja possível alcançar/analisar o alcance deste objetivo?’ ou ‘que informações são necessárias para que seja possível alcançar/analisar o alcance este objetivo?’. Quando pertinente, necessidades de informação podem ser decompostas em subnecessidades.

Entrada: objetivos de medição.

Saída: necessidades de informação.

4.1.1.3 Identificar Medidas que Atendem as Necessidades de Informação

A partir das necessidades de informação, devem ser identificadas medidas que sejam capazes de atendê-las. Para definir uma medida, inicialmente, é necessário identificar que propriedade (elemento mensurável) de que tipo de entidade (tipo de entidade mensurável) deve ser quantificado para atender a necessidade de informação. A partir daí, definem-se as medidas que serão utilizadas para quantificar as propriedades. Ao definir uma medida, além do elemento mensurável e tipo de entidade mensurável, é necessário indicar a unidade de medida na qual a medida é expressa e a sua escala.

Entrada: necessidades de informação.

Saída: medidas que atendem as necessidades de informação.

4.1.1.4 Selecionar Fontes de Dados

Nesta subatividade devem ser selecionadas as fontes de dados que serão utilizadas para a obtenção dos indicadores. Deve-se buscar por fontes de dados que possuam dados relativos às medidas identificadas na subatividade anterior, ou seja, dados capazes de atender as necessidades de informação identificadas. Para isso, as fontes de dados selecionadas devem tratar do mesmo domínio de discurso relacionado às medidas e às necessidades de informação. Uma forma de iniciar a verificação se uma fonte de dados é capaz de fornecer dados para as medidas identificadas é observar os tipos de entidades mensuráveis indicados nas medidas identificadas e verificar se a fonte de dados trata entidades desse tipo.

Entrada: necessidades de informação e medidas identificadas.

Saída: fontes de dados a serem utilizadas.

As subatividades 4.1.1.5, 4.1.1.6, 4.1.1.7 e 4.1.1.8 descritas a seguir são realizadas iterativamente para cada medida identificada.

4.1.1.5 Verificar se há Disponibilidade Direta de Dados para a Medida

Para cada medida identificada, deve-se verificar se há dados diretamente relacionados à medida disponíveis nas fontes selecionadas. É importante notar que na subatividade anterior as fontes são selecionadas com base em uma análise mais geral, onde se verifica se as fontes são capazes de fornecer dados úteis para atender as necessidades de informação identificadas. Já nesta subatividade é feita uma análise detalhada das fontes de dados selecionadas. Para isso, cada fonte de dados deve ser analisada, verificando, se para cada medida identificada, há dados nas fontes de dados. Nesse momento é importante considerar que nomes distintos podem ser usados para se referir a uma mesma medida e que um mesmo nome pode denotar medidas diferentes. Além disso, também é possível que o nome da medida não esteja explícito na fonte de dados. Dessa forma, a investigação deve ser feita com base na semântica da medida (o que a medida significa), ao invés de na sua sintaxe (como a medida é representada). Nesse sentido, deve-se atentar, principalmente, ao tipo da entidade medida (tipo de entidade mensurável) e à propriedade medida (elemento mensurável), que devem ser consistentes entre a medida identificada e os dados na fonte de dados. Outro aspecto que deve ser considerado é a unidade de medida. Caso a medida identificada seja expressa em uma unidade de medida, é preciso verificar se os dados disponíveis na fonte de dados são expressos na mesma unidade de medida. Caso não sejam, deve-se analisar se é necessário fazer a conversão para a unidade da medida identificada e, em caso afirmativo, a conversão deve ser realizada.

Cabe ressaltar que é possível que os dados encontrados na fonte de dados sejam referentes às medidas identificadas, mas com algumas diferenças que não comprometem seu uso. Nesse caso, é possível alterar a medida identificada para ficar consistente com os dados presentes na fonte de dados. Por exemplo, suponha que na subatividade “Identificar Medidas que Atendem as Necessidades de Informação” tenha sido identificada a medida tamanho em metros e que na fonte de dados haja dados para tamanho, mas expressos em centímetros. Caso seja possível utilizar os dados de tamanho em centímetros, pode-se, ao invés de fazer a conversão de centímetros para metros, alterar a medida inicialmente identificada, mudando sua unidade de medida de metros para centímetros.

Entrada: medidas identificadas e fontes de dados selecionadas.

Saída: medidas que possuem dados disponíveis diretamente na fonte de dados, medidas sem dados disponíveis diretamente na fonte de dados, alterações nas medidas identificadas (quando pertinente).

4.1.1.6 Verificar se é Possível Computar Dados para a Medida a partir dos Dados Existentes

Para cada medida identificada que não tenha dados diretamente disponíveis nas fontes de dados, deve ser feita uma investigação a fim de verificar se é possível obter dados para a medida a partir de dados existentes nas fontes de dados. Algumas formas de computar dados para medidas a partir de dados existentes são:

- Derivação: algumas medidas podem ser derivadas de outras aplicando-se fórmulas de medição. Por exemplo, suponha que a medida índice de massa corporal em quilos/metros² tenha sido identificada e não esteja disponível diretamente nas fontes de dados. Suponha, ainda, que nas fontes de dados existam dados para as medidas peso em quilos e altura em metros. Dessa forma, os dados para a medida índice de massa corporal podem ser obtidos aplicando-se a fórmula de medição: índice de massa corporal = peso/altura².
- Contagem: também é possível computar dados para medidas a partir da contagem de instâncias registradas nas fontes de dados. Por exemplo, suponha que a medida número de consultas realizadas tenha sido identificada e não tenha dados disponíveis diretamente nas fontes de dados. Suponha, também, que na fonte de dados estejam registradas todas as consultas realizadas. Nesse caso, os dados para a medida número de consultas realizadas poderiam ser obtidos a partir da contagem de todas as consultas registradas.
- Contagem seletiva: dados também podem ser computados a partir da contagem seletiva de instâncias registradas na fonte de dados. Por exemplo, suponha que a partir das fontes de dados do exemplo anterior se deseje obter dados para a medida número de consultas realizadas com atraso. Suponha, ainda, que para cada consulta registrada nas fontes de dados estejam registrados o paciente, o médico, o horário agendado e o horário de atendimento. Nesse caso, os dados para a medida poderiam ser obtidos a partir da contagem das consultas cujo horário de atendimento foi superior ao horário agendado.

De maneira similar ao que ocorre na subatividade “Verificar se há Disponibilidade Direta de Dados para a Medida”, é possível que seja necessário alterar as medidas

inicialmente identificadas para ficarem consistentes com os dados que podem ser computados a partir de dados presentes nas fontes de dados.

Entrada: medidas sem dados disponíveis diretamente na fonte de dados e fontes de dados.

Saída: medidas cujos dados podem ser computados a partir de dados existentes, alterações nas medidas identificadas (quando pertinente), medidas sem dados disponíveis.

4.1.1.7 Estabelecer Procedimento de Medição

Nesta subatividade para cada medida para a qual podem ser obtidos dados (diretamente disponíveis ou computados) nas fontes de dados, deve-se definir o procedimento de medição que deve ser adotado para a obtenção de dados. O procedimento de medição deve descrever claramente a forma como a medida será obtida a partir das fontes de dados utilizadas.

Entrada: Medidas que possuem dados diretamente disponíveis diretamente na fonte de dados, medidas cujos dados podem ser computados a partir de dados existentes.

Saída: procedimentos de medição para as medidas.

4.1.1.8 Incluir Medida em Lista de Medidas sem Dados Disponíveis

Nesta subatividade, as medidas sem dados disponíveis nas fontes selecionadas até o momento são incluídas em uma lista.

Entrada: medidas sem dados disponíveis.

Saída: lista de medidas sem dados disponíveis.

4.1.1.9 Reportar Medidas sem Dados Disponíveis

Caso haja medidas sem dados disponíveis nas fontes de dados selecionadas é possível selecionar novas fontes de dados e repetir as subatividades “Verificar se há Disponibilidade Direta de Dados para a Medida”, “Verificar se é Possível Computar Dados para a Medida a partir dos Dados Existentes”, “Estabelecer Procedimento de Medição” e “Incluir Medida em Lista de Medidas sem Dados Disponíveis” até que haja dados disponíveis para as medidas identificadas. No entanto, mesmo selecionando outras fontes de dados, é possível não haver dados para algumas medidas. Nesse caso, nesta subatividade, deve-se comunicar ao responsável pela disponibilização dos dados, as medidas sem dados disponíveis.

Entrada: lista de medidas sem dados disponíveis.

Saída: reporte de medidas sem dados disponíveis.

4.1.2 Obter Dados

Nesta subatividade os dados associados às medidas devem ser extraídos das fontes de dados e armazenados em uma base de dados cuja estrutura seja adequada para permitir/facilitar a obtenção de informações a partir da análise dos dados armazenados. A escolha da tecnologia a ser adotada para a base de dados deve levar em conta aspectos como volume de dados, complexidade/flexibilidade das consultas desejadas e tecnologias disponíveis, podendo ser utilizadas desde planilhas eletrônicas simples até sistemas de bancos de dados mais sofisticados. Minimamente, deve ser possível, para cada valor medido, identificar a entidade medida, o elemento medido e a medida aplicada. Idealmente, a base de dados deve permitir o armazenamento de objetivos de medição, necessidades de informação, tipo de entidade mensurável, entidade mensurável, elemento mensurável, medidas, unidades de medida, escalas, procedimentos de medição, fórmulas de medição e medições, bem como dos relacionamentos entre eles.

Para cada medida, devem ser registrados os dados associados (valores medidos). A extração dos dados para as medidas deve ser realizada considerando-se os procedimentos de medição definidos na subatividade “Estabelecer Procedimento de Medição”.

É provável que alguns dos dados que devem ser armazenados não estejam presentes nas fontes de dados, como, por exemplo, a entidade medida e o elemento medido, dentre outros. Assim, a inserção desses dados na base de dados deve ser feita manualmente ou utilizando-se outro mecanismo.

Entrada: objetivos de medição, necessidades de informação, medidas que possuem dados diretamente disponíveis na fonte de dados, medidas cujos dados podem ser computados a partir de dados existentes, procedimentos de medição, fontes de dados.

Saída: base de dados contendo os dados relevantes para análises.

4.1.3 Analisar Dados

Nesta atividade é realizada a análise dos dados visando à obtenção das informações necessárias para os objetivos de medição estabelecidos. Na Figura 4.3 são apresentadas suas subatividades, as quais são descritas em seguida.



Figura 4.3 – Subatividades de “Analisar Dados”.

4.1.3.1 Selecionar Objetivo de Medição para Análise

Nesta subatividade é realizada a seleção do objetivo de medição para o qual se deseja obter informações a partir da análise de dados.

Entrada: objetivos de medição definidos no planejamento.

Saída: objetivo de medição selecionado para análise.

4.1.3.2 Selecionar Medidas para Análise

Nesta subatividade é realizada a seleção das medidas que serão analisadas, tomando-se como base o objetivo de medição selecionado na subatividade anterior. A partir do objetivo de medição, devem ser identificadas as necessidades de informação a ele relacionadas e as medidas que satisfazem essas necessidades. As medidas selecionadas para análise, tipicamente, são aquelas capazes de fornecer informações sobre o alcance do objetivo de medição. Nesse contexto, essas medidas são chamadas de indicadores.

Entrada: objetivo de medição selecionado para análise, necessidades de informação associadas a esse objetivo, medidas e indicadores.

Saída: medidas e indicadores selecionados para análise.

4.1.3.3 Selecionar Dados para Análise

Uma análise não necessariamente leva em conta todos os valores medidos para uma medida. Nesse sentido, nesta subatividade é realizada a seleção dos dados que serão analisados, levando-se em consideração o objetivo de medição selecionado anteriormente.

Entrada: objetivo de medição, medidas e indicadores selecionados para análise e valores medidos.

Saída: valores medidos selecionados para análise.

4.1.3.4 Analisar Dados Selecionados

Nesta subatividade as medidas e valores medidos são analisados para que sejam produzidos resultados que forneçam as informações necessárias ao alcance do objetivo de medição. Ferramentas para representação de dados, como, por exemplo, gráficos de barras, gráficos de controle, gráficos de dispersão, dentre outros, podem ser utilizadas para apoiar a representação e a análise dos dados.

Entrada: objetivo de medição, medidas, indicadores e valores medidos selecionados para análise.

Saída: resultado da análise.

4.1.3.5 Reportar Resultados

Nesta atividade os resultados das análises realizadas, incluindo sua relação com o objetivo de medição selecionado, devem ser reportados às partes interessadas.

Entrada: objetivo de medição selecionado, resultados de análises.

Saída: relato dos resultados de análises.

A Figura 4.4 apresenta uma visão integrada das subatividades da abordagem proposta. As subatividades em cor de rosa referem-se à atividade “Planejar Obtenção de Indicadores”, a subatividade em amarelo refere-se a “Obter Dados” e as em azul são subatividades de “Analisar Dados”.

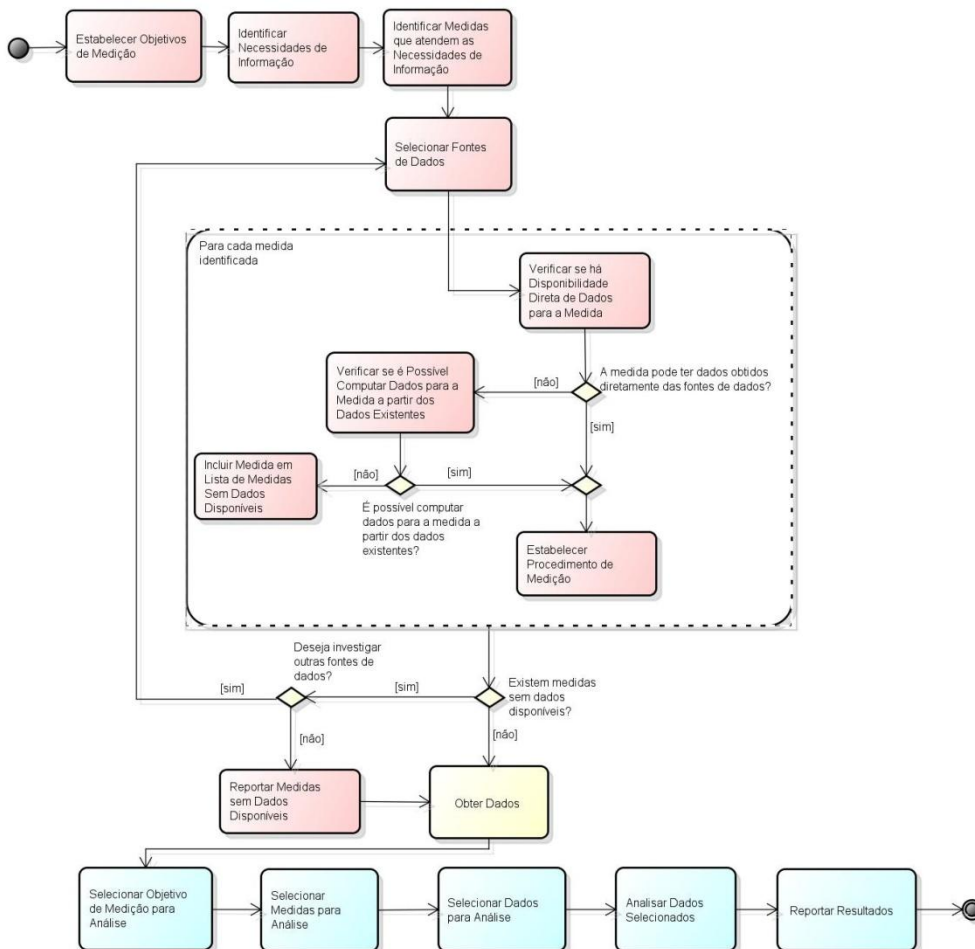


Figura 4.4 – Subatividades da abordagem proposta.

4.1.4 Relações entre a Abordagem Proposta e a *Measurement Task Ontology*

Conforme dito anteriormente, a abordagem proposta foi definida com base na *Core Ontology on Measurement* (COM), que forneceu o vocabulário a ser utilizado, e na *Measurement Task Ontology* (MTO) que guiou a definição das atividades e subatividades da abordagem. Para ilustrar as relações entre o procedimento proposto e MTO, na Tabela 4.1 são apresentadas as relações entre as atividades/subatividades da abordagem e as atividades/subatividades de MTO e na Tabela 4.2 são apresentadas as relações entre as entradas e saídas das atividades/subatividades da abordagem e os conceitos do modelo estrutural de MTO. Vale lembrar que, conforme dito no Capítulo 3, o modelo estrutural de MTO é equivalente ao modelo conceitual de COM.

Tabela 4.1 - Relações entre as atividades/subatividades do procedimento e MTO.

Atividade/Subatividade do Procedimento	Atividade/Subatividade de MTO
Planejar Obtenção de Indicadores	<i>Plan Measurement</i>
Estabelecer Objetivos de Medição	<i>Establish Measurement Goals</i>
Identificar Necessidades de Informação	<i>Identify Information Needs</i>
Identificar Medidas que Atendem as Necessidades de Informação	<i>Identify Measures</i>
Selecionar Fontes de Dados	-
Verificar se há Disponibilidade Direta de Dados para a Medida	-
Verificar se é possível Computar Dados para a Medida a partir dos Dados Existentes	-
Definir Procedimentos de Medição	<i>Establish Measurement Procedure</i>
Incluir Medida em Lista de Medidas Sem Dados Disponíveis	-
Reportar Medidas sem Dados Disponíveis	-
Obter Dados	<i>Perform Measurement</i>
Analisar Dados	<i>Analyse Measurement</i>
Selecionar Objetivo de Medição para Análise	<i>Select Measurement Goal for Analysis</i>
Selecionar Medidas para Análise	<i>Select Measures for Analysis</i>
Selecionar Dados para Análise	<i>Select Data for Analysis</i>
Analisar Dados Selecionados	<i>Analyse Data</i>
Reportar Resultados	<i>Report Results</i>

Tabela 4.2 - Relações entre as entradas e saídas das subatividades do procedimento e os conceitos do modelo estrutural de MTO.

Subatividade do Procedimento	Entradas e Saídas	Conceitos do modelo estrutural de MTO
Estabelecer Objetivos de Medição	Entradas	
	Objetivos gerais	<i>Goal</i>
	Saídas	
	Objetivos de medição	<i>Measurement Goal</i>

Tabela 4.3 - Relações entre as entradas e saídas das subatividades do procedimento e os conceitos do modelo estrutural de MTO (continuação).

Subatividade do Procedimento	Entradas e Saídas	Conceitos do modelo estrutural de MTO
Identificar Necessidades de Informação	Entradas	
	Objetivos de medição	<i>Measurement Goal</i>
	Saídas	
	Necessidade de informação	<i>Information Need</i>
Identificar Medidas que Atendem as Necessidades de Informação	Entradas	
	Necessidades de informação	<i>Information Need</i>
	Saídas	
	Medidas que atendem as necessidades de informação	<i>Measure</i>
Selecionar Fontes de Dados	Entradas	
	Necessidades de informação	<i>Information Need</i>
	Medidas identificadas	<i>Measure</i>
	Saídas	
	Fontes de dados a serem utilizadas	-
Verificar se há Disponibilidade Direta de Dados para a Medida	Entradas	
	Medidas identificadas	<i>Measure</i>
	Fontes de dados selecionadas	-
	Saídas	
	Medidas que possuem dados disponíveis diretamente na fonte de dados	<i>Measure</i>
	Medidas sem dados disponíveis diretamente na fonte de dados	<i>Measure</i>
	Alterações nas medidas identificadas	<i>Measure</i>
Verificar se é possível Computar Dados para a Medida a partir dos Dados Existentes	Entradas	
	Medidas sem dados disponíveis diretamente na fonte de dados	<i>Measure</i>
	Fonte de Dados	-
	Saídas	
	Medidas cujos dados podem ser computados a partir de dados existentes	<i>Measure</i>
	Medidas sem dados disponíveis	<i>Measure</i>
	Alterações nas medidas identificadas	<i>Measure</i>
Estabelecer Procedimento de Medição	Entradas	
	Medidas com dados disponíveis diretamente na fonte de dados	<i>Measure</i>
	Medidas cujos dados podem ser computados a partir de dados existentes	<i>Measure</i>
	Saídas	
	Procedimentos de medição para as medidas	<i>Measurement Procedure</i> <i>Measurement Formula</i>
Incluir Medida em Lista de Medidas Sem Dados Disponíveis	Entradas	
	Medidas sem dados disponíveis	<i>Measure</i>
	Saídas	
	Lista de medidas sem dados disponíveis	<i>Measure (lista de)</i>

Tabela 4.4 - Relações entre as entradas e saídas das subatividades do procedimento e os conceitos do modelo estrutural de MTO (continuação).

Subatividade do Procedimento	Entradas e Saídas	Conceitos do modelo estrutural de MTO
Reportar Medidas sem Dados Disponíveis	Entradas	
	Lista de medidas sem dados disponíveis	<i>Measure (lista de)</i>
	Saídas	
	Reporte de medidas sem dados disponíveis	-
Obter Dados	Entradas	
	Objetivos de medição	<i>Measurement Goal</i>
	Necessidades de informação	<i>Information Need</i>
	Medidas que possuem dados diretamente disponíveis na fonte de dados	<i>Measure</i>
	Medidas cujos dados podem ser computados a partir de dados existentes	<i>Measure</i>
	Procedimentos de medição	<i>Measurement Procedure</i>
	Fontes de dados	-
	Saídas	
Base de dados contendo os dados relevantes para análises	-	
Selecionar Objetivo de Medição para Análise	Entradas	
	Objetivos de medição definidos no planejamento	<i>Measurement Goal</i>
	Saídas	
Objetivo de medição selecionado para análise	<i>Measurement Goal</i>	
Selecionar Medidas para Análise	Entradas	
	Objetivo de medição selecionado para análise	<i>Measurement Goal</i>
	Necessidades de informação associadas ao objetivo	<i>Information Need</i>
	Medidas	<i>Measure</i>
	Indicadores	<i>Indicator</i>
	Saídas	
	Medidas selecionadas para análise	<i>Measure</i>
Indicadores selecionados para análise	<i>Indicator</i>	
Selecionar Dados para Análise	Entradas	
	Objetivo de medição selecionados	<i>Measurement Goal</i>
	Medidas selecionados para análise	<i>Measure</i>
	Indicadores selecionados para análise	<i>Indicator</i>
	Valores medidos	<i>Measured Value</i>
	Saídas	
	Valores medidos selecionados para análise	<i>Measured Value</i>

Tabela 4.5 - Relações entre as entradas e saídas das subatividades do procedimento e os conceitos do modelo estrutural de MTO (continuação).

Subatividade do Procedimento	Entradas e Saídas	Conceitos do modelo estrutural de MTO
Analisar Dados Selecionados	Entradas	
	Objetivo de medição	<i>Measurement Goal</i>
	Indicadores selecionados para análise	<i>Indicator</i>
	Medidas selecionadas para análise	<i>Measure</i>
	Valores medidos selecionados para análise	<i>Measured Values</i>
	Saídas	
	Resultado da Análise	<i>Measurement Analysis Result</i>
Reportar Resultados	Entradas	
	Objetivo de medição	<i>Measurement Goal</i>
	Resultado da análise	<i>Measurement Analysis Result</i>
	Saídas	
	Relato dos resultados de análise	-

4.2 Apoio Computacional para Uso da Abordagem Proposta

Com o propósito de apoiar o uso da abordagem proposta, foi desenvolvida uma ferramenta. Para o desenvolvimento das funcionalidades presentes na ferramenta, foi utilizado o *framework* OpenXava¹¹. O OpenXava é um framework escrito em Java para desenvolvimento ágil de aplicações *Web*. Sua arquitetura se utiliza de requisições AJAX (*Asynchronous Javascript And XML*) para apresentar uma interface dinâmica. Com base em classes Java desenvolvidas e anotadas por *annotations* providas pelo OpenXava e Hibernate¹², o *framework* produz funcionalidades CRUD (*Create, Read, Update e Delete*) e cria o banco de dados da ferramenta. As demais funcionalidades, bem como algumas lógicas do negócio, devem ser implementadas pelo desenvolvedor e são apoiadas por meio de *Actions*, *Views* e *Validators*. *Actions* permitem customizar e tratar ações disparadas pelo usuário, como, por exemplo, carregar arquivos ou apresentar uma janela a partir do clique de um botão. *Views* definem como e quais dados serão exibidos, sendo possível desenvolver *Views* mais complexas do que as geradas automaticamente pelo OpenXava. Por fim, *Validators* permitem validar um dado antes do salvamento.

Com base nas classes do modelo de classes definido para a ferramenta, foram criadas classes Java (com as devidas adequações), as quais foram anotadas com *annotations* inerentes ao OpenXava. O OpenXava tem uma integração com *JasperReports*, uma

¹¹ <http://openxava.org/>

¹² <http://hibernate.org/>

ferramenta de construção de relatórios, através de uma interface gráfica chamada *iReport*¹³. Isso permite que o desenvolvedor desenhe o relatório no *iReport*, que por sua vez, gera um arquivo *.jrxml, o qual é consumido pelo OpenXava para apresentar o relatório no escopo do *portlet* definido, com os dados inseridos na aplicação.

Neste trabalho, os *portlets* desenvolvidos pelo OpenXava foram organizados no portal *LifeRay*¹⁴. O *LifeRay* é um portal que dispõe aplicações integradas com um sistema de gerenciamento de usuários e permissões de acesso, além de prover menus e navegabilidade para os *portlets* adicionados.

Em relação aos tipos de fontes de dados, atualmente são suportados arquivos CSV e planilhas Excel. Para acrescentar novos tipos de arquivo, novas classes estendendo uma determinada *interface*¹⁵ devem ser implementadas. Na Figura 4.5 é apresentada uma tela da ferramenta, onde é possível observar (no menu) as funcionalidades de apoio à abordagem proposta.

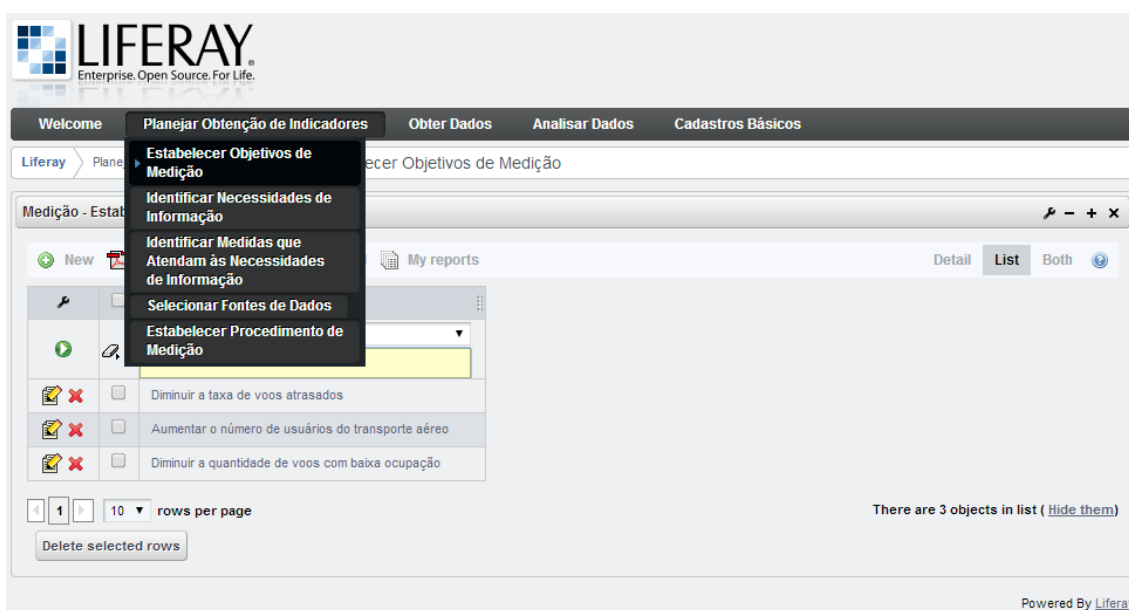


Figura 4.5 – Tela da ferramenta desenvolvida, mostrando menu com funcionalidades relacionadas às atividades da abordagem proposta.

¹³ <http://community.jaspersoft.com/questions/529236/ireport-353-released>

¹⁴ <http://www.liferay.com>

¹⁵ <http://docs.oracle.com/javase/tutorial/java/concepts/interface.html>

4.3 Uso da Abordagem Proposta para Obter Indicadores a partir de Bases de Dados Abertos disponibilizadas pela ANAC e INFRAERO

Para avaliar a viabilidade da abordagem proposta, como prova de conceito, a abordagem foi usada para a obtenção de indicadores a partir de bases de dados abertos disponibilizadas pela ANAC e INFRAERO. Segundo Oates (2006), uma prova de conceito mostra que uma proposta é exequível, porém não permite afirmar se ela funciona em um contexto real. Nesta seção são apresentados alguns dos resultados obtidos no estudo realizado.

Para realizar o estudo, inicialmente foi necessário escolher o domínio no qual o estudo seria realizado. Após uma pesquisa na Internet para identificar domínios onde há boa disponibilização de dados abertos, decidiu-se por Transporte Aéreo Brasileiro. Conforme descrito anteriormente, para realizar a abordagem é preciso ter como entrada objetivos gerais, a partir dos quais serão definidos os objetivos de medição. Definiu-se, então, como objetivo geral: Melhorar o transporte aéreo brasileiro.

Seguindo a abordagem proposta, inicialmente foi realizado o planejamento dos indicadores que se desejava obter (atividade “Planejar Obtenção de Indicadores”). Para isso, foram estabelecidos os objetivos de medição (subatividade “Estabelecer Objetivos de Medição”), as necessidades de informação (subatividade “Identificar Necessidades de Informação”) e as medidas necessárias para atendê-las (subatividade “Identificar Medidas que Atendem Necessidades de Informação”). Na Tabela 4.6 são apresentados os objetivos de medição, necessidades de medição e medidas identificados. Cabe destacar que vários outros objetivos de medição, necessidades de informação e medidas poderiam ser identificados a partir do objetivo geral considerado, no entanto, no contexto do estudo realizado, foi definido um conjunto limitado, que permitisse avaliar a exequidade da abordagem.

Tabela 4.6 – Objetivos de medição, necessidades de informação e medidas relacionadas.

<i>Objetivo de Medição</i>	<i>Necessidade de Informação</i>	<i>Medida</i>
Diminuir a taxa de voos atrasados	Qual a quantidade de voos realizados?	Número de Voos Realizados
	Qual a quantidade de voos atrasados?	Número de Voos Atrasados
	Qual a taxa de voos atrasados?	Taxa de Voos Atrasados
Diminuir a quantidade de voos com baixa ocupação	Qual a quantidade de assentos disponibilizados?	Número de Assentos Disponibilizados
	Qual a quantidade de assentos ocupados?	Número de Assentos Ocupados
	Qual a taxa de ocupação de assentos?	Taxa de Ocupação de Assentos

Tabela 4.7 – Objetivos de medição, necessidades de informação e medidas relacionadas(continuação).

<i>Objetivo de Medição</i>	<i>Necessidade de Informação</i>	<i>Medida</i>
Aumentar o número de usuários do transporte aéreo	Qual a quantidade de passageiros desembarcados?	Número de Passageiros Desembarcados
	Qual a quantidade de passageiros embarcados?	Número de Passageiros Embarcados
	Qual a quantidade total de passageiros movimentados?	Número Total de Passageiros Movimentados

Para ilustrar o uso da ferramenta no estudo, a seguir ,na Figura 4.6 é apresentada a tela com os objetivos de medição cadastrados.

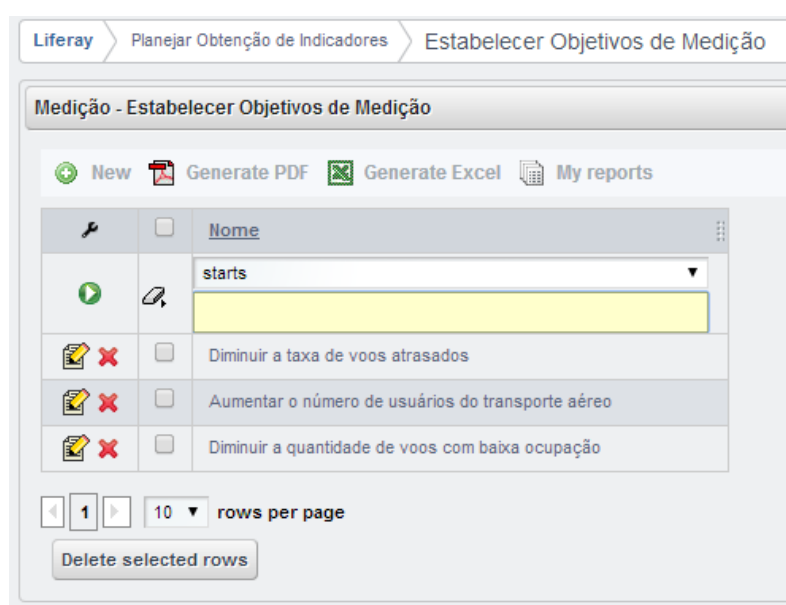


Figura 4.6 – Objetivos de medição estabelecidos a partir do objetivo geral.

Na Figura 4.7 é apresentada a tela com as necessidades de informação relacionadas ao objetivo de medição Diminuir a taxa de voos atrasados.

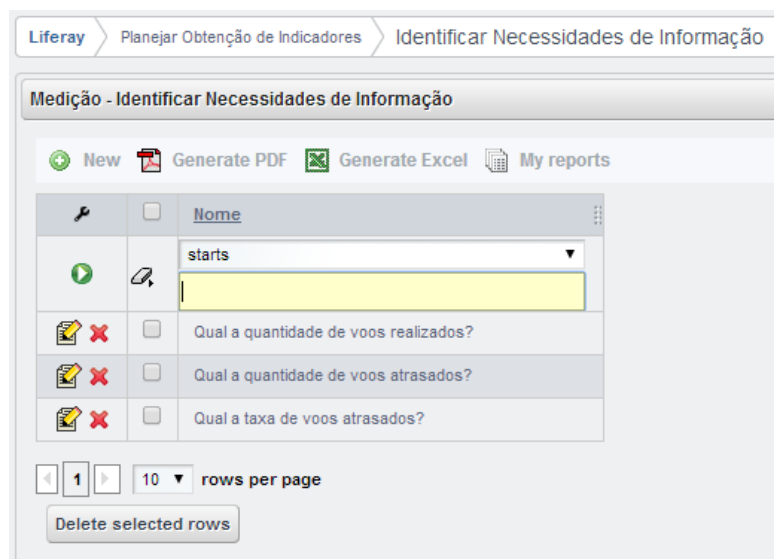


Figura 4.7 – Necessidades de informação relacionadas ao objetivo Diminuir a taxa de voos atrasados.

Para cada necessidade de informação, foram registradas as medidas identificadas. Na Figura 4.8 é apresentada a tela mostrando as medidas cadastradas para as necessidades de informação relacionadas ao objetivo de medição Diminuir a taxa de voos atrasados. Para cada objetivo de medição, foram identificadas as medidas que seriam usadas como indicadores para monitorar seu alcance. Para o objetivo de medição citado, a medida Taxa de Voos Atrasados foi identificada como indicador.

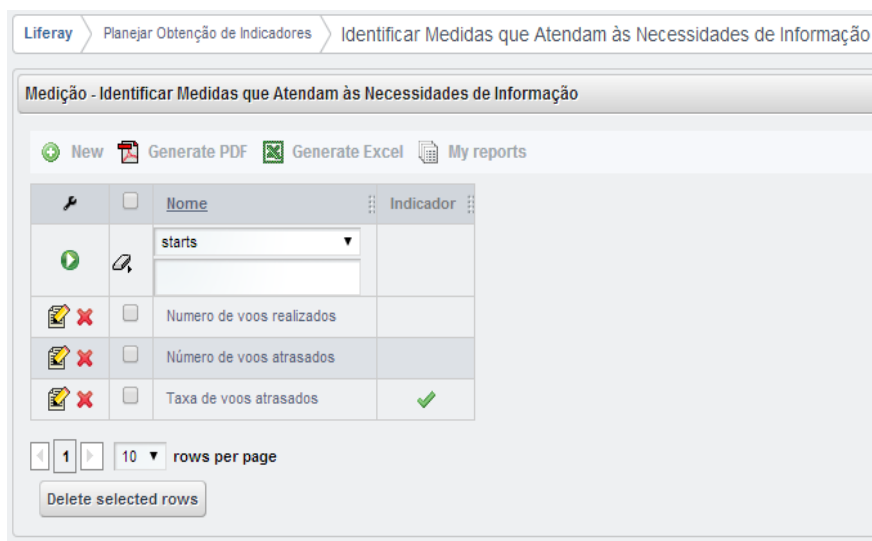


Figura 4.8 – Medidas cadastradas para as necessidades de informação.

A tela mostrada na figura anterior apresenta apenas a lista de medidas cadastradas. Clicando-se no nome da medida é possível visualizar em detalhes os dados de cada medida. Na Figura 4.9 é apresentada a tela onde é realizado o cadastro de medidas. Para ilustrar o

cadastro de uma medida, são apresentados os dados cadastrados para a medida Taxa de Voos Atrasados.

The screenshot shows the Liferay interface for registering a measure. The page title is "Identificar Medidas que Atendam às Necessidades de Informação". The main form is titled "Medição - Identificar Medidas que Atendam às Necessidades de Informação". It includes fields for "Nome" (Taxa de voos atrasados), "Escala" (Escala de Razão), "Unidade de Medida", and "Elemento Mensurável" (Taxa de voos atrasados). Below the form are three sections: "Tipos de Entidades Mensuráveis" with a table containing "Nome" and "Aeroporto"; "Necessidades de Informação" with a table containing "Nome" and "Qual a taxa de voos atrasados?"; and "Objetivos Indicados" with a table containing "Nome" and "Diminuir a taxa de voos atrasados". A "Save" button is at the bottom left.

Figura 4.9 – Cadastro da medida Taxa de Voos Atrasados.

Após a identificação das medidas, seguindo o procedimento, foram selecionadas as fontes de dados que seriam usadas (subatividade “Selecionar Fontes de Dados”). Para o estudo, foram utilizadas três fontes de dados. A primeira fonte usada é um dos arquivos disponibilizados pela ANAC conhecidos como Voo Regular Ativo, que contém dados de voos realizados em um determinado mês. Os arquivos são disponibilizados no formato CSV. A segunda fonte usada é um dos arquivos disponibilizados também pela ANAC como Demanda e Oferta de Aeroportos, que contém dados sobre a oferta e ocupação de assentos nas aeronaves. Os arquivos são disponibilizados no formato XLSX. A terceira fonte utilizada é um dos arquivos disponibilizados pela INFRAERO chamados Estatística

dos Aeroportos, que contêm dados sobre a movimentação de passageiros nos aeroportos. Os arquivos são disponibilizados no formato XLS. Na Figura 4.10 é apresentada a tela onde é feita a seleção e *upload* das fontes.



Figura 4.10 – Tela para realização de *upload* de fontes selecionadas.

O próximo passo foi, para cada medida identificada, verificar se havia dados diretamente disponíveis nas fontes de dados para a medida (subatividade “Verificar se há disponibilidade direta de dados para a medida”). Notou-se que as medidas Número de Assentos Disponibilizados, Número de Assentos Ocupados e Taxa de Ocupação de Assentos estavam diretamente disponíveis na fonte de dados provida pela INFRAERO. Uma vez que não havia dados diretamente disponíveis para as demais medidas, foi necessário, para cada uma delas, verificar se havia dados nas fontes de dados que permitiam computar as medidas (subatividade “Verificar se é possível computar a medida a partir de dados existentes”). Para cada medida identificada, foram definidos seus procedimentos de medição (subatividade “Estabelecer Procedimento de Medição”). Na Figura 4.11 é apresentada a tela para estabelecimento de procedimento de medição para a medida Número de Voos Atrasados. Para cada medida identificada, para definir o procedimento de medição devem ser informados a fonte de dados que contém os dados coletados para a medida, a coluna da fonte de dados da qual serão extraídas as entidades mensuráveis medidas, a coluna que será contada (em caso de contagem), o tipo de agregação (em caso de agregação) e a fórmula para seleção dos dados. A definição da fórmula usa sintaxe similar a SQL, sendo as primitivas "AND" e "OR" do SQL traduzidas para "E" e "OU".

Liferay > Planejar Obtenção de Indicadores > Estabelecer Procedimento de Medição

Medição - Estabelecer Procedimento de Medição

Search Refresh Save

Medida modified successfully

Nome Numero de voos atrasados

Fonte de dados VRA2013

Coluna referenciada para entidade [Aeroporto Origem] [Choose column](#)

Coluna referenciada para agregação [Número do Voo] [Choose column](#)

Tipo de Agregação Contagem

Seleção de Dados

< < = > <> () E OU Adicionar entidade Adicionar constante

{ [Tipo de Linha] = 'H' OU [Tipo de Linha] = 'N' OU [Tipo de Linha] = 'R' } E [Situação] = 'Realizado' E [Partida Prevista] < [Partida Real]

Save Voltar

Escolha uma coluna

VRA2013:

	<u>Sigla da Empresa</u>	<u>Número do Voo</u>	<u>D</u> <u>I</u>	<u>Tipo de Linha</u>	<u>Aeroporto Origem</u>	<u>Aeroporto Destino</u>	<u>Partida Prevista</u>	<u>Partida Real</u>	<u>Chegada Prevista</u>	<u>Chegada Real</u>	<u>Situação</u>
Dados	GLO	1631	0	N	SBVT	SBRJ	28/03/2013 05:50	28/03/2013 05:50	28/03/2013 07:08	28/03/2013 07:08	Realizado
	TUS	6550	0	C	SBGR	SBEG	22/03/2013 05:50	22/03/2013 03:41	22/03/2013 09:42	22/03/2013 07:27	Realizado
	TIB	5249	0	R	SBAU	SBGR	09/03/2013 05:50	09/03/2013 05:50	09/03/2013 07:05	09/03/2013 07:05	Realizado

Figura 4.11 – Tela para definição de procedimento de medição – medida Número de Voos Atrasados.

Após todas as medidas identificadas terem seus procedimentos de medição definidos, os dados a elas relacionados foram extraídos das fontes de dados abertos e armazenados na base de dados da ferramenta (atividade “Obter Dados”). Na Figura 4.12 é apresentada a tela onde é realizada a extração dos dados para as medidas.

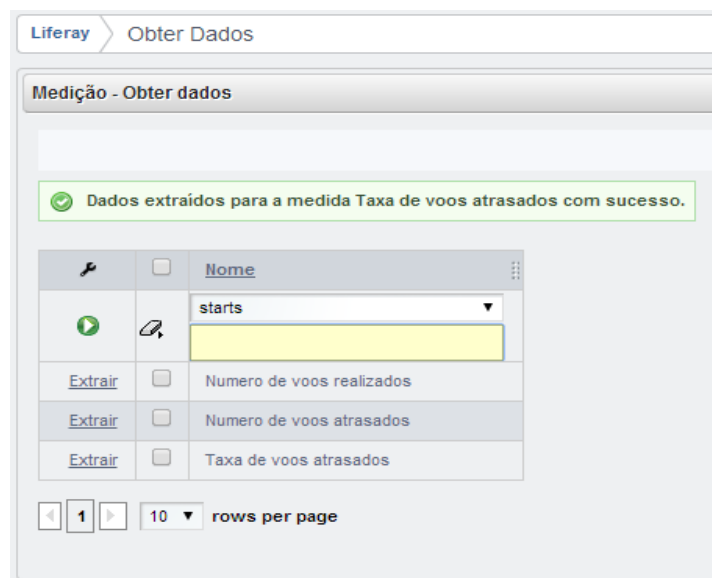


Figura 4.12 – Tela para extração de dados.

Após a extração dos dados, foram conduzidas as análises dos dados (atividade “Analisar Dados”). Para cada análise realizada, foi selecionado o objetivo de medição para o qual a análise deveria prover informações (subatividade “Selecionar Objetivo de Medição para Análise”), foram selecionadas as medidas que seriam analisadas (subatividade “Selecionar Medidas para Análise”) e os dados que seriam considerados (subatividade “Selecionar Dados para Análise”). O objetivo de medição, medidas e dados selecionados foram apresentados em relatórios, permitindo que análises fossem conduzidas, fornecendo informações para o monitoramento dos objetivos de medição.

Como exemplo, para o objetivo de medição Diminuir a taxa de voos atrasados, foi selecionada a medida (indicador) Taxa de Voos Atrasados e foram selecionados dados referentes aos meses de maio de 2013 e maio de 2014, para que pudesse ser analisado se houve acréscimo ou diminuição na taxa de voos atrasados em um mesmo mês de um ano para outro. O mês de maio foi escolhido por ser o último mês com dados disponíveis no momento do estudo. Nas Figura 4.13 e Figura 4.14 são apresentados fragmentos dos relatórios com os dados referentes à Taxa de Voos Atrasados, respectivamente, nos meses de maio de 2013 e maio de 2014, por aeroporto.

Objetivo: Diminuir a taxa de voos atrasados	
Medida: Taxa de voos atrasados	
Entidade	Valor
SAAR	0.1
SBAE	0.17105263
SBAR	0.088141024
SBAT	0.2
SBAU	0.22302158
SBAX	0.28735632
SBBE	0.2126367
SBBH	0.35732323
SBBR	0.24043308

Figura 4.13 – Fragmento de relatório com Taxa de Voos Atrasados por aeroporto em maio de 2013.

Objetivo: Diminuir a taxa de voos atrasados	
Medida: Taxa de voos atrasados	
Entidade	Valor
SBAE	0.17322835
SBAQ	0.10810811
SBAR	0.2589792
SBAT	0.06451613
SBAU	0.11016949
SBAX	0.04
SBBE	0.2379205
SBBH	0.21272728
SBBR	0.26840854
SBBV	0.18656716
SBCA	0.36082473
SBCB	0.5
SBCF	0.18256834
SBCG	0.1509716

Figura 4.14 – Fragmento de relatório com Taxa de Voos Atrasados por aeroporto em maio de 2014.

No final dos relatórios é apresentada a Taxa de Voos Atrasados considerando todos os aeroportos (vide Figura 4.15 e Figura 4.16)

SWIQ	0.5
SWKO	0.6363636
SWLB	0.75
SWLC	0.9230769
SWOB	0.44444445
SWPI	0.12903225
SWRD	0.17105263
SWSI	0.25
SWTP	0.11111111
TOTAL	0.30716879385483864

Figura 4.15 –Taxa de Voos Atrasados (total) em maio de 2013.

SSTL	0.34042552
SWBC	0.5
SWCA	1.0
SWEI	0.4
SWGN	0.48453608
SWHT	1.0
SWKO	0.7037037
SWLC	0.16
SWNK	1.0
SWOB	1.0
SWPI	0.15384616
SWRD	0.4893617
SWSI	0.35897437
SWVC	1.0
TOTAL	0.33449144665289243

Figura 4.16 –Taxa de Voos Atrasados (total) em maio de 2014.

Analisando-se os valores dos indicadores, percebe-se que houve um aumento na taxa de voos atrasados de maio de 2014 quando comparada com maio de 2013, indicando que o objetivo Diminuir a taxa de voos atrasados não foi alcançado.

Capítulo 5

Conclusão

Neste capítulo são apresentadas as considerações finais, as principais contribuições do trabalho e as perspectivas de trabalhos futuros.

5.1 Considerações Finais

Através da análise de dados é possível obter informações que auxiliam na percepção das necessidades de uma organização e permitem a tomada de decisões adequadas e alinhadas aos seus objetivos. Apesar de dados de diversas naturezas poderem ser analisados, dados quantitativos têm maior objetividade e flexibilidade, contribuindo para a obtenção de informações mais precisas e para a diversidade de análises (HAND; MANNILA; SMYTH, 2001; ROBERTS, 1985).

Atualmente, organizações disponibilizam na Internet vários dados que podem ser acessados publicamente. Considerando a diversidade e volume de dados abertos disponíveis, bem como de objetivos que podem levar ao acesso e análise desses dados, a identificação dos dados realmente úteis em um dado contexto pode ser ineficiente se realizada de maneira totalmente *ad hoc*.

A definição de quais dados se deseja coletar, a coleta dos dados e sua análise podem ser tratadas no contexto de um processo de medição, independente do domínio ao qual os dados se referem. O processo de medição pode ser visto, então, como um provedor dos dados a partir dos quais são obtidas informações úteis para a tomada de decisão.

Neste trabalho o processo de medição foi usado como base para o desenvolvimento de uma abordagem que visa auxiliar a obtenção de indicadores capazes de fornecer informações úteis. Para isso, foram desenvolvidas duas ontologias que representam o conhecimento central de medição e o conhecimento relacionado ao processo de medição.

A *Core Ontology on Measurement* (COM) (BARCELLOS; FALBO; FRAUCHES, 2014) foi desenvolvida com o propósito de representar a conceituação central relacionada à medição e aborda aspectos relacionados às entidades que podem ser medidas e suas propriedades mensuráveis, medidas, planejamento de medição, medição e análise. COM aborda conceitos que são independentes do domínio no qual a medição é aplicada, tais

como medida, medição e valor medido. Alguns conceitos, como, por exemplo, incerteza, precisão, amostra e instrumento, embora presentes em vários domínios onde a medição é aplicada e até em outras ontologias de medição, como em (KIM *et al.*, 2007), são significativos apenas aos domínios onde a medição apoia estudos mais experimentais (por exemplo, Física e Química), sendo irrelevantes em outros domínios. Por essa razão, conceitos como esses não foram abordados em COM.

A *Task Measurement Ontology* (TMO) representa o conhecimento relacionado ao processo de medição, descrevendo as atividades e subatividades que o compõem, os papéis envolvidos, as entradas e saídas e os fluxos entre as atividades e subatividades. Assim como em COM, em TMO são representadas apenas as atividades consideradas independentes do domínio de aplicação. Nesse sentido, atividades relacionadas à calibração de instrumentos e seleção de amostra, dentre outras, não foram incluídas por serem específicas da medição no âmbito mais experimental.

A abordagem proposta é baseada em COM, que forneceu o conhecimento sobre medição, e em OTM, que orientou sobre as atividades necessárias, suas entradas e saídas. Alinhada à OTM, a abordagem possui três atividades: Planejar Obtenção de Indicadores, Obter Dados e Analisar Dados. Na atividade Planejar Obtenção de Indicadores, são identificadas as medidas a serem utilizadas para monitorar o alcance a objetivos estabelecidos inicialmente. Na atividade Obter Dados os dados relacionados às medidas identificadas são extraídos das bases de dados abertas e, por fim, na atividade Analisar Dados, os dados são analisados para proverem informações úteis.

Foi desenvolvida uma ferramenta computacional para apoiar o uso da abordagem proposta. A ferramenta possui três módulos: Planejamento, Obtenção de Dados e Análise de Dados, conforme as atividades presentes na abordagem. Como prova de conceito da abordagem proposta, a ferramenta foi utilizada para obtenção de indicadores a partir de bases de dados abertas da ANAC (Agência Nacional de Aviação Civil) e INFRAERO (Empresa Brasileira de Infraestrutura Aeroportuária). A ferramenta desenvolvida é ainda uma versão inicial e precisa ser evoluída para apoiar mais eficientemente o uso da abordagem.

Cabe destacar que, embora tenha sido desenvolvida uma ferramenta específica para apoiar o uso da abordagem, a abordagem proposta é independente de tecnologia. Ela pode ser utilizada com apoio computacional desenvolvido especificamente para suportar o uso da abordagem, como a ferramenta construída neste trabalho, bem como com outros apoios mais gerais, como planilhas eletrônicas e outras ferramentas de manipulação de dados.

Diferente das abordagens para dados abertos encontradas na literatura (por exemplo, (BRAUNSCHWEIG *et al.*, 2012) e (HIENERT *et al.*, 2011), que propõem ferramentas para extração de dados), a abordagem não é uma tecnologia propriamente dita (um sistema, por exemplo), mas um procedimento que pode ser apoiado por tecnologia.

Embora a abordagem tenha sido desenvolvida para dados abertos, ela não é limitada a fontes de dados desse tipo, podendo ser usada, também, para obter indicadores a partir de outros tipos de fontes de dados. Por exemplo, uma organização que possui bases de dados alimentadas por seus sistemas de informação pode usar a abordagem para identificar indicadores a partir dos dados armazenados.

Por fim, é importante notar que, como o próprio nome diz, a abordagem proposta visa à obtenção de indicadores e não a exploração de informações em base de dados abertos. Dessa forma, se a informação que se quer obter ao se investigar, por exemplo, uma base de dados abertos sobre os servidores públicos federais brasileiros, for quais servidores estão lotados em um determinado órgão, a abordagem não se aplica. Porém, se o que se deseja saber é quantos funcionários estão lotados em um dado órgão ou quantos têm renda maior que um determinado valor, a abordagem se aplica. Ainda, convém perceber que a abordagem proposta pode ser associada a outros tipos de abordagens para permitir a obtenção mais eficiente e abrangente de informações. No exemplo mencionado, a abordagem proposta poderia ser utilizada associada a uma abordagem de exploração de informações, permitindo a obtenção de indicadores bem como de informações.

5.2 Contribuições

As principais contribuições desta dissertação são:

(i) A Abordagem Baseada em Ontologias para Obtenção de Indicadores a partir de Dados Abertos, descrita no Capítulo 4, que pode ser utilizada para guiar a obtenção de indicadores a partir de dados abertos.

(ii) A *Core Ontology on Measurement*, descrita no Capítulo 3, que apresenta a conceituação central de medição de software, independente do domínio de aplicação. Essa ontologia pode ser usada como base para a construção de ontologias para domínios específicos. Buscando facilitar o uso de COM como base para a construção de ontologias específicas de domínio, ela foi organizada em uma Linguagem de Padrões Ontológicos (FALBO *et al.*, 2013) e publicada no artigo *Towards a Measurement Ontology Pattern Language*, no *1st Joint Workshop on Ontologies in Conceptual Modeling and Information Systems Engineering* (Onto.Com/ODISE), realizado

no contexto da *8th International Conference on Formal Ontology in Information Systems* (FOIS 2014) (BARCELLOS; FALBO; FRAUCHES, 2014).

(iii) A *Measurement Task Ontology*, descrita no Capítulo 3, que apresenta o processo geral de medição, destacando suas atividades, papéis envolvidos, entradas e saídas.

(iv) A ferramenta apresentada no Capítulo 4, desenvolvida para apoiar o uso da abordagem.

5.3 Perspectivas Futuras

Considerando o estágio atual aqui apresentado, algumas das perspectivas de trabalhos futuros são destacadas a seguir:

- Explorar o uso da abordagem em outras bases de dados abertos.
- Conduzir estudos experimentais para avaliar o uso da abordagem por outras pessoas e com outras fontes de dados.
- A abordagem definida até o momento não trata de aspectos relacionados à integração de dados quando várias fontes de dados são utilizadas. A abordagem pode ser estendida para tratar esses aspectos seja pela inclusão de novas atividades, seja pela associação com outras abordagens que tratem da integração de dados.
- Avaliar o uso da abordagem proposta associada a uma abordagem que apoie a exploração de informações (vide último parágrafo da seção 5.1) para identificar os benefícios obtidos.
- Evoluir a ferramenta desenvolvida. A ferramenta foi desenvolvida com algumas limitações que podem ser exploradas em trabalhos futuros. Por exemplo, a versão atual da ferramenta contempla uma variedade limitada de cálculos e combinações para a obtenção de medidas que poderia ser expandida. Mecanismos para facilitar a seleção dos dados a serem considerados na análise também podem ser desenvolvidos e novos tipos de arquivos de fontes de dados podem ser suportados. Funcionalidades para apoiar a verificação da existência de dados associados às medidas identificadas podem ser desenvolvidas para potencializar o auxílio ao uso da abordagem. Por fim, a apresentação dos dados pode ser melhorada através de relatórios mais estruturados e gráficos.

- Utilização da *Core Ontology on Measurement* para o desenvolvimento de ontologias de medição para domínios específicos. *Feedbacks* obtidos a partir da utilização de COM podem ser utilizados para evoluí-la.
- Utilização da *Task Measurement Ontology* como base para o desenvolvimento de outras ontologias de tarefa para domínios específicos. *Feedbacks* obtidos a partir da utilização de OTM podem ser utilizados para evoluí-la.
- Utilização das ontologias propostas em iniciativas de interoperabilidade semântica de normas e padrões de medição, bem como de sistemas.

Referências Bibliográficas

ALBRIGHT, S.; WINSTON, W.; ZAPPE, C. **Data Analysis and Decision Making**. Mason, Ohio: Cengage Learning, 2010. p. 1080

ALBUQUERQUE, A. F. DE. **Ontological Foundations for Conceptual Modeling Datatypes**. Vitória: Universidade Federal do Espírito Santo, 2013.

ALTMAN, D. G.; BLAND, J. M. Measurement in medicine: the analysis of method comparison studies. **The statistician**, v. 32, p. 307–317, 1983.

AMYOT, D.; MUSSBACHER, G.; GHANAVATI, S.; KEALEY, J. **GRL Modeling and Analysis with jUCMNav** 5th International i* Workshop. **Anais...**2011

ANDREAS, H. Ontological Aspects of Measurement. **Science+Business Media**, v. 18, n. 3, p. 379–394, 2008.

ARAÚJO, S. F. C. DE; SCHWABE, D. **Explorator: a tool for exploring RDF data through direct manipulation** Linked data on the web WWW2009 workshop (LDOW2009). **Anais...**2009

ARNDT, R.; TRONCY, R.; STAAB, S.; HARDMAN, L. COMM: A Core Ontology for Multimedia Annotation. In: **Handbook on Ontologies**. [s.l.] Springer, 2009. p. 403–421.

ARNDT, R.; TRONCY, R.; STAAB, S.; HARDMAN, L.; VACURA, M. **COMM: Designing a Well-founded Multimedia Ontology for the Web**. [s.l.] Springer, 2007.

ASSOCIATION, A. P. H.; ASSOCIATION, A. W. W.; FEDERATION, W. E. **Standard methods for the examination of water and wastewater** (E. W. Rice, R. B. Baird, A. D. Eaton, L. S. Clesceri, Eds.) Washington DC American Public Health Association., , 2012.

ASTM. Standard Test Method for Determination of Biodiesel (Fatty Acid Methyl Esters) Content in Diesel Fuel Oil Using Mid Infrared Spectroscopy (FTIR-ATR-PLS Method) 1. p. 1–10, 2014.

AUER, S.; BIZER, C.; KOBILAROV, G.; LEHMANN, J.; CYGANIAK, R.; IVES, Z. DBpedia: A Nucleus for a Web of Open Data. In: CHOI, K.-S.; NIXON, L.; MAYNARD, D. (Eds.). **The Semantic Web**. [s.l.] Springer Berlin Heidelberg, 2007. p. 722–735.

BARCELLOS, M. P. **Uma Estratégia para Medição de Software e Avaliação de Bases de Medidas para Controle Estatístico de Processos de Software em Organizações de Alta Maturidade**. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2009.

BARCELLOS, M. P.; FALBO, R. D. A.; DAL MORO, R. **A well-founded software measurement ontology** Frontiers in Artificial Intelligence and Applications. **Anais...**2010

BARCELLOS, M. P.; FALBO, R. D. A.; ROCHA, A. R.; BRAZIL, V. **A Well-founded Software Process Behavior Ontology to Support Business Goals Monitoring in High Maturity Software Organizations**IEEE International EDOC Enterprise Computing Conference Workshops. **Anais...**Vitória - ES: IEEE, 2010

BARCELLOS, M. P.; FALBO, R. DE A.; FRAUCHES, V. G. V. **Towards a Measurement Ontology Pattern Language**ONTO.COM/ODISE, 2014.

BARCELLOS, M. P.; FALBO, R. DE A.; ROCHA, A. R. A strategy for preparing software organizations for statistical process control. **Journal of the Brazilian Computer Society**, v. 19, n. 4, p. 445–473, 2013.

BASILI, V. R.; ROMBACH, H. D. Goal Question Metrics Paradigm. In: **Encyclopedia of Software Engineering**. [s.l: s.n.].

BERNERS-LEE, T. **Linked Data - Design Issues**. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html>>.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific America**, v. 284, n. 5, p. 34–43, 2001.

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked Data - The Story So Far. **International Journal on Semantic Web and Information Systems**, v. 5, p. 1–22, 2009.

BÖHM, C.; NAUMANN, F.; FREITAG, M.; GEORGE, S.; HÖFLER, N.; KÖPPELMANN, M.; LEHMANN, C.; MASCHER, A.; SCHMIDT, T. Linking Open Government Data: What Journalists Wish They Had Known. **SEMANTICS**, 2010.

BORGO, S.; MASOLO, C. Foundational Choices in DOLCE. In: STAA, S.; STUDER, R. (Eds.). **Handbook on Ontologies**. [s.l.] Springer Berlin Heidelberg, 2009.

BRAUNSCHWEIG, K.; EBERIUS, J.; THIELE, M.; LEHNER, W. OPEN—Enabling Non-expert Users to Extract, Integrate, and Analyze Open Data . **Datenbank-Spektrum**, v. 12, n. 2, p. 121–130, 2012.

BREITMAN, K.; SALAS, P.; SARAIVA, D.; GAMA, V.; CASANOVA, M. A.; VITERBO, J.; PIRES, R.; FRANZOSI, E.; CHAVES, M. Open Government Data in Brazil. **IEEE Intelligent Systems**, v. 27, p. 45–49, 2012.

BRINGUENTE, A. C. DE O.; FALBO, R. DE A.; GUIZZARDI, G. Using a foundational ontology for reengineering a software process ontology. **Journal of Information and Data Management**, v. 2, n. 3, p. 511, 2011.

CORDEIRO, K. DE F.; MARINO, T.; CAMPOS, M. L. M.; BORGES, M. **Use of Linked Data in the design of information infrastructure for collaborative emergency management system**Computer Supported Cooperative Work in Design (CSCWD), 2011 15th International Conference on. **Anais...**IEEE, 2011

DIEZ, J. A. A Hundred Years of Numbers. An Historical Introduction to Measurement Theory 1887-1990 . **Studies in History and Philosophy of Science Part A**, v. 28, n. 1, p. 167–185, 1996.

EPSTEIN, M.; MANZONI, J.-F. Implementing corporate strategy:: From Tableaux de Bord to balanced scorecards. **European Management Journal**, v. 16, n. 2, p. 190–203, 1998.

FALBO, R. DE A. **Experiences in Using a Method for Building Domain Ontologies** (P. of the 16th I. C. on S. E. and K. Engineering, Ed.)International Workshop on Ontology In Action. **Anais...**Banff, Canada: 2004

FALBO, R. DE A.; BARCELLOS, M. P.; NARDI, J. C.; GUIZZARDI, G. Organizing Ontology Design Patterns as Ontology Pattern Languages. **The Semantic Web: Semantics and Big Data**, v. 7882, p. 61–75, 2013.

FINKELSTEIN, L. Widely-defined measurement – An analysis of challenges. **Measurement**, v. 42, p. 1270–1277, 2009.

FINKELSTEIN, L.; LEANING, M. S. A review of the fundamental concepts of measurement. **Measurement**, v. 2, n. 1, p. 25–34, 1984.

FONSECA, L. B. R. DA. **Mapeando Dados Governamentais com uma Ontologia de Organizações**. [s.l.] UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO, 2014.

FORTUIN, L. Performance indicators — Why, where and how? **European Journal of Operational Research**, v. 34, n. 1, p. 1–9, 1988.

FRANZ, T.; STAAB, S.; ARNDT, R. **The X-COSIM integration framework for a seamless semantic desktop**Proceedings of the 4th international conference on Knowledge capture. **Anais...**ACM, 2007

FRIGERIO, A.; GIORDANI, A.; MARI, L. Outline of a general model of measurement. **Synthese**, v. 175, n. 2, p. 123–149, 2010.

GENNARI, J. H.; TU, S. W.; ROTHENFLUH, T. E.; MUSEN, M. A. Mapping domains to methods in support of reuse. **International Journal of Human-Computer Studies**, v. 41, n. 3, p. 399–424, 1994.

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.

GUARINO, N. Formal Ontology and Information Systems. In: PRESS, I. O. S. (Ed.). **Formal Ontologies in Information Systems**. Trento, Italy: IOS Press, 1998. p. 3–15.

GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Models**. Enschede, The Netherlands: Telematica Instituut, 2005.

GUIZZARDI, G. On Ontology, ontologies, Conceptualizations, Modeling Languages and (Meta)Models. **Databases and Information Systems IV**, v. 155, 2007.

GUIZZARDI, G.; FALBO, R. DE A.; S.S.GUIZZARDI, R. **Grounding Software Domain Ontologies in the Unified Foundational Ontology (UFO): The case of the ODE Software Process Ontology**CIbSE, 2008.

GUIZZARDI, G.; WAGNER, G.; FALBO, R. D. A. Towards Ontological Foundations for the Conceptual Modeling of Events Background: The Unified Foundational Ontology (UFO). 2013.

HAND, D. J.; MANNILA, H.; SMYTH, P. **Principles of data mining**. [s.l.] MIT press, 2001.

HEATH, T.; BIZER, C. **Linked Data: Evolving the Web into a Global Data Space**. [s.l.] Morgan & Claypool, 2011.

HEIJSTF, G. VAN; SCHREIBER, A. T.; WIELINGA, B. J. Using explicit ontologies in KBS development. **International Journal of Human-Computer Studies**, v. 46, n. 2-3, p. 183–292, 1997.

HEINK, U.; KOWARIK, I. What are indicators? On the definition of indicators in ecology and environmental planning. **Ecological Indicators**, p. 584–593, 2010.

HEMPEL, C. G. **Fundamentals of Concept Formation in Empirical Science**. Chicago: University of Chicago Press, 1952. v. 2

HERRE, H.; BUREK, P.; HOEHNDORF, R.; LOEBE, F.; MICHALEK, H. General Formal Ontology (GFO), a foundational ontology integrating objects and processes. **Onto-Med Report**, v. 8, 2006.

HIENERT, D.; ZAPILKO, B.; SCHAER, P.; MATHIAK, B. **Web-based multi-view visualizations for aggregated statistics** Proceedings of the 5th International Workshop on Web APIs and Service Mashups. **Anais...ACM**, 2011

HOXHA, J.; BRAHAJ, A. **Open Government Data on the Web: A Semantic Approach** Emerging Intelligent Data and Web Technologies (EIDWT), **International Conference on IEEE**, , 2011.

ISO. **Standards Handbook: Quantities and units**. 3rd. ed. Geneva: ISO, 1992.

ISO/IEC. ISO/IEC 15939:2007 - Systems and software engineering - Measurement process. **Information Technology–Software Measurement ...**, v. 2007, 2007.

JANSSEN, M.; ZUIDERWIJK, A. Infomediary Business Models for Connecting Open Data Providers and Users. **Social Science Computer Review**, 2014.

JCGM. **International vocabulary of metrology – Basic and general concepts and associated terms** VIM, 2012.

KAPLAN, R. S.; NORTON, D. P. **The Balanced Scorecard Translating Strategy In Action** Proceedings of the IEEE, 1996. Disponível em:
<<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=628729>>

KIM, H. M.; SENGUPTA, A.; FOX, M. S.; DALKILIC, M. A Measurement Ontology Generalizable for Emerging Domain Applications on the Semantic Web. **Journal of Database Management**, v. 18, n. 1, 2007.

KIMBALL, R.; ROSS, M. **The Data Warehouse Toolkit**. Canada: John Wiley and Sons, Inc, 2002.

LAORY, I.; HADJ ALI, N. B.; TRINH, T. N.; SMITH, I. F. C. Measurement system configuration for damage identification of continuously monitored structures. **Journal of Bridge Engineering**, v. 17, n. 6, p. 857–866, 2012.

LJUNGBERG, A. **Process measurement****International Journal of Physical Distribution & Logistics Management**, 2002.

MAEHR, M. L.; ZUSHO, A. Achievement goal theory. **Handbook of motivation at school**, p. 77, 2009.

MARI, L. Notes towards a qualitative analysis of information in measurement results. **Measurement**, v. 25, p. 183–192, 1999.

MARI, L. A quest for the definition of measurement. **Measurement**, v. 46, p. 2890–2895, 2013.

MARSHALL, D.; JOHNNELL, O.; WEDEL, H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. **Bmj**, v. 312, n. 7041, p. 1254–1259, 1996.

MARTINS, A. F. **Construção de Ontologias de Tarefa e sua Reutilização na Engenharia de Requisitos**. Vitória, Brasil: Universidade Federal do Espírito Santo, 2009.

MARTINS, A. F.; FALBO, R. DE A. **Models for representing task ontologies**CEUR Workshop Proceedings. **Anais...2008**

MENDONÇA, R. R. DE; DA CRUZ, S. M. S.; CERDA, J. F. S. M. D. LA; CAVALCANTI, M. C.; CORDEIRO, K. F.; CAMPOS, M. L. M. **LOP: capturing and linking open provenance on LOD cycle**Proceedings of the Fifth Workshop on Semantic Web Information Management. **Anais...ACM**, 2013

MESSICK, S. The standard problem: Meaning and values in measurement and evaluation. **American Psychologist**, v. 30, n. 10, p. 955, 1975.

MODAVE, F. E. **A measurement theory perspective for mcdm**Fuzzy Systems, 2001. The 10th IEEE International Conference on, 2001

OATES, B. J. **Researching Information Systems and Computing**. [s.l.] Sage, 2006. v. 37p. 341

OPEN KNOWLEDGE FOUNDATION. **Open Definition**. Disponível em: <<http://opendefinition.org/od/>>.

PAPADIMITRIOU, D.; FÀBREGA, L.; VILÀ, P.; CAREGLIO, D.; DEMEESTER, P. Measurement-based Research: Methodology, Experiments, and Tools. **Computer Communication Review**, v. 42, n. 5, 2012.

PARMENTER, D. Key Performance Indicators (KPI). In: **Key Performance Indicators (KPI): Developing , Implementing , and Using Winning KPIs Second Edition**. [s.l.: s.n.]. p. 9.

PIDUN, T.; FELDEN, C. **Limitations of Performance Measurement Systems based on Key Performance Indicators** Proceedings of the Seventeenth Americas Conference on Information Systems Detroit, Michigan, 2011.

REYNOLDS, D. **The Organization Ontology**. Disponível em: <<http://www.w3.org/TR/vocab-org/>>. Acesso em: 29 abr. 2014.

ROBERTS, F. S. **Measurement Theory with Applications to Decisionmaking, Utility and the Social Sciences**. New York: Cambridge University Press, 1985.

ROCHA, A. R.; SANTOS, G.; BARCELLOS, M. P. **Medição de Software e Controle Estatístico de Processos**. 8. ed. [s.l.] Elsevier, 2012.

SASSOON, G. W. Measurement theory in linguistics. **Synthese**, v. 174, n. 1, p. 151–180, 2010.

SCHERP, A.; FRANZ, T.; SAATHOFF, C.; STAAB, S. **F--a model of events based on the foundational ontology dolce+ DnS ultralight** Proceedings of the fifth international conference on Knowledge capture. **Anais...ACM**, 2009

SCHERP, A.; FRANZ, T.; SAATHOFF, C.; STAAB, S. A core ontology on events for representing occurrences in the real world. **Multimedia Tools and Applications**, v. 58, n. 2, p. 293–331, 2012.

SCHERP, A.; SAATHOFF, C.; FRANZ, T.; STAAB, S. Designing core ontologies. **Applied Ontology**, v. 6, p. 177–221, 2011.

SIMON, H. A. On the concept of organizational goal. **Administrative Science Quarterly**, p. 1–22, 1964.

SOLINGEN, R.; BERGHOUT, E., 1999, The Goal/Question/Metric Method: a Practical Guide for Quality Improvement of Software Development, **McGraw-Hill** In Rocha, A.R., Santos, G. and Barcellos, M.P. *Medição de Software e Controle Estatístico de Processos*. Série de Livros do PBQP Software, ISSN 1679-1878. Vol. 8. 2012: **Ministério da Ciência, Tecnologia e Inovação (MCTI), SEPIN/PBQP Software - Brasília - DF, Brasil**.

SPEITEL, K. F. Measurement assurance. **Handbook of Industrial Engineering, 2nd ed., John Wiley & Sons, New York, NY**, 1992.

STAAB, S.; SCHERP, A.; ARNDT, R.; TRONCY, R.; GRZEGORZEK, M.; SAATHOFF, C.; SCHENK, S.; HARDMAN, L. Semantic multimedia. In: **Reasoning Web**. [s.l.] Springer, 2008. p. 125–170.

STEVENS, S. S. On the Theory of Scales of Measurement. **Science**, v. 103, n. 2684, p. 677–680, 1946.

STEVENS, S. S. Mathematics, measurement, and psychophysics. 1951.

TONI, A. DE; TONCHIA, S. **Performance measurement systems - Models, characteristics and measures** *International Journal of Operations & Production Management*, 2001.

TORGERSON, W. S. *Theory and methods of scaling*. 1958.

USCHOLD, M.; JASPER, R. A Framework for Understanding and Classifying Ontology Applications. *Methods*, v. 18, p. 1–12, 1999.

W3C. **The RDF Data Cube Vocabulary** *W3C Recommendation*, 2014. Disponível em: <http://travesia.mcu.es/portaln/jspui/bitstream/10421/7472/1/RDF_data_cube_vocabulary.pdf>

WAND, Y.; WEBER, R. Mario Bunge's Ontology as a formal foundation for information systems concepts. *Studies on Mario Bunge's Treatise, Rodopi, Atlanta*, p. 123–149, 1990.

ZHU, Y.; WANG, Z. Measurement of soil resistivity based on FFT DC component. *Guangxue Jingmi Gongcheng/Optics and Precision Engineering*, v. 21, n. 1, p. 115–123, 2013.

Apêndice 1

Mapeamento entre COM e RSMO

Neste apêndice são apresentados os mapeamentos entre as questões de competência e axiomas de COM de RSMO.

Conforme mencionado no Capítulo 3, a *Reference Software Measurement Ontology* (RSMO) (BARCELLOS; FALBO; ROCHA, 2013; BARCELLOS, 2009) foi uma das fontes utilizadas para o desenvolvimento da *Core Ontology on Measurement* proposta neste trabalho. Com exceção dos conceitos *Basic Measurement Procedure*, *Derivative Measurement Procedure* e *Measurement Planning Item*, todos os conceitos de COM estão presentes em RSMO, embora haja diferenças nas relações e restrições envolvidas. Algumas questões de competência e axiomas definidos em RSMO foram reutilizados em COM, sendo que em alguns deles, adaptações foram necessárias.

Na Tabela A1.1 são apresentados os mapeamentos entre as questões de competência (QC) e axiomas (A) de COM e RSMO.

Tabela A1.1 - Mapeamentos entre as questões de competência e axiomas de COM de RSMO.

COM	RSMO
<i>Subontologia Measurable Entities</i>	<i>Subontologia Entidades Mensuráveis</i>
QC1 a QC6	QC1 a QC6
A1 a A3	A1 a A3
<i>Subontologia Measure</i>	<i>Subontologia Medidas</i>
QC1	QC1
QC2	QC4
QC3	QC6
QC4	QC7
QC5	QC9 (Adaptada)
QC6	QC10
QC7	QC12 (Adaptada)
QC8	QC16
A1	A1 (Adaptado)
A2	A5 (Adaptado)

Tabela A1.1 - Mapeamentos entre as questões de competência e axiomas de COM de RSMO
(Continuação).

COM	RSMO
<i>Subontologia Measure</i>	<i>Subontologia Medidas</i>
A3	-
A4	-
A5	-
<i>Subontologia Measurement Planning</i>	<i>Subontologia Objetivos de Medição</i>
QC1	QC3
QC2	QC4 (Adaptada)
QC3	QC7 (Adaptada)
QC4	-
QC5	QC6 (Adaptada)
QC6	QC5
QC7	-
A1 a A7	-
A8	A2
A9 e A10	-
A11	A4 (Adaptado)
A12	A6
<i>Subontologia Measurement</i>	<i>Subontologia Medição</i>
QC1	-
QC2	QC1
QC3 a QC6	QC3 a QC5, e QC6 (Adaptada)
A1	-
A2	-
A3	-
A4	-
A5	A10
<i>Subontologia Measurement Analysis</i>	<i>Subontologia Análise de Medição</i>
QC1	-
QC2	QC9 (Adaptada)
QC3	QC10

Tabela A1.1 - Mapeamentos entre as questões de competência e axiomas de COM de RSMO
(Continuação).

COM	RSMO
<i>Subontologia Measurement Analysis</i>	<i>Subontologia Análise de Medição</i>
QC4	QC12
QC5	QC8 (Adaptada)
QC6	-
A1	-
A2	-
A3	-
A4	-
A5 a A8	-
A9	A4 (Adaptado)
A10 e A11	-