

Uma Estratégia para Medição de Software e Avaliação de Bases de Medidas para Controle Estatístico de Processos de Software em Organizações de Alta Maturidade

Monalessa Perini Barcellos^{1,2}, Ana Regina Rocha¹, Ricardo de Almeida Falbo²

¹ COPPE/UFRJ - Universidade Federal do Rio de Janeiro
Caixa Postal: 68511 – CEP: 21945-970 – Rio de Janeiro, RJ, Brasil

² UFES - Universidade Federal do Espírito Santo – Centro Tecnológico
Departamento de Informática, CEP: 29075-910 – Vitória, ES, Brasil

monalessa@inf.ufes.br; darocha@cos.ufrj.br; falbo@inf.ufes.br

Resumo. *As organizações de software têm aumentado seu interesse em melhoria de processos de software. Nos níveis mais elevados de maturidade, a melhoria de processos envolve a implementação do controle estatístico de processos, que requer medidas e dados adequados a esse contexto. Entretanto, a inadequação das medidas e dados coletados tem sido destacada na literatura como um dos principais obstáculos para a implementação bem sucedida do controle estatístico. Este artigo apresenta uma estratégia, proposta em uma tese de doutorado, para auxiliar as organizações de software na obtenção e manutenção de bases de medidas adequadas ao controle estatístico de processos, bem como na realização de medições apropriadas a esse contexto.*

Abstract. *Software organizations have increased their interest on software process improvement (SPI). In high maturity levels, SPI involves implementing statistical process control (SPC), which requires measures and data that are suitable for this context. However, this has been pointed in the literature as one of the main obstacles for a successful implementation of SPC in SPI efforts. This paper presents a strategy, proposed in a doctorate thesis, to support organizations to obtain and maintain measurement repository suitable for SPC, as well as to perform measurements appropriately in this context.*

1. Introdução

A crescente exigência do mercado por produtos e serviços de software cada vez melhores tem aumentado o interesse das organizações pela melhoria de processos. Atualmente, há vários *frameworks* de apoio à definição e institucionalização de programas com esse objetivo nos quais a medição ocupa papel fundamental. Exemplos notáveis são o MR MPS [Softex, 2009], o CMMI [Chrissis *et al.*, 2006], a ISO/IEC 15504 [ISO/IEC, 2003] e a ISO/IEC 12207 [ISO/IEC, 2008]. Alguns desses *frameworks* propõem a implementação da melhoria de processos em níveis, nos quais a maturidade e a capacidade dos processos evoluem gradativamente.

Nos níveis mais elevados de maturidade (níveis A e B do MR MPS e níveis 4 e 5 do CMMI), realizar a melhoria de processos de software baseando-se apenas em medição tradicional não é suficiente. Nesses níveis é necessário realizar o controle estatístico dos processos (CEP) para conhecer o seu comportamento, determinar seu desempenho em execuções anteriores e, a partir daí, prever seu desempenho em projetos

correntes e futuros, verificando se são capazes de alcançar os objetivos estabelecidos e identificando ações corretivas e de melhoria quando apropriado.

Para utilizar o CEP nesse contexto, é necessário que as medidas definidas e os dados coletados sejam adequados à análise de desempenho de processos. No entanto, as experiências com a realização do CEP nas organizações têm revelado aos pesquisadores e profissionais programas de medição precários, que não produzem medidas que permitam a análise do desempenho e capacidade dos processos [Card, 2004; Kitchenham *et al.*, 2006; Sargut e Demirors, 2006; Raczinski e Curtis, 2008].

Segundo Card (2004) uma das razões que leva as organizações a não realizarem medição de forma adequada ao CEP é a inexistência de orientações sobre como deve ser realizada a medição em um programa de melhoria de processos de acordo com o nível de maturidade organizacional, incluindo-se aspectos relacionados à aplicação futura dos dados. Apesar de existirem normas e modelos que abordam a alta maturidade e determinam quais práticas são necessárias para caracterizá-la, eles não orientam – até mesmo porque não é esse seu objetivo – sobre o que deve ser realizado para implementar essas práticas, entre elas o CEP.

Sendo assim, considerando-se que: (i) é difícil para as organizações de software implantarem os níveis mais altos de maturidade em seus processos, principalmente devido às dificuldades para implantação e realização do CEP; (ii) os estudos e experiências relatados na literatura mostram que grande parte dessas dificuldades está relacionada à realização inadequada da medição; e (iii) os modelos e normas existentes dizem o que esperam que seja feito, porém não fornecem um caminho sobre como fazer; foi proposta em uma tese de doutorado uma estratégia para apoiar as organizações de software que buscam alta maturidade em seus processos na obtenção e manutenção de bases de medidas adequadas ao CEP, bem como na realização de medições apropriadas a esse contexto. A estratégia é composta por uma Ontologia de Medição de Software, um Instrumento para Avaliação de Bases de Medidas considerando Adequação ao CEP e um Conjunto de Recomendações para Medição de Software Adequada ao CEP.

Para apresentar a estratégia proposta na tese de doutorado, este artigo encontra-se assim organizado: na Seção 2 é apresentada uma pequena fundamentação teórica sobre medição de software e controle estatístico de processos; na Seção 3 é apresentada uma visão geral da estratégia proposta; na Seção 4 é apresentada a Ontologia de Medição de Software; na Seção 5 é apresentado o Instrumento para Avaliação de Bases de Medidas; na Seção 6 é apresentado o Conjunto de Recomendações para Medição de Software; na Seção 7 são discutidos alguns trabalhos correlatos; e na Seção 8 são realizadas as considerações finais do artigo.

2. Medição de Software e Controle Estatístico de Processos

A medição de software é considerada uma das atividades mais importantes para a gerência e melhoria de processos e produtos de software. As medidas, ao serem coletadas e armazenadas, podem ser analisadas para fornecerem informações importantes para a tomada de decisão, envolvendo a identificação e realização de ações que orientem os projetos e processos a alcançarem os objetivos para eles estabelecidos [ISO/IEC, 2002]. À medida que uma organização realiza um programa de melhoria de processos, a qualidade de seus processos tende a aumentar, elevando seu nível de maturidade e revelando a necessidade de identificar novos objetivos a serem alcançados.

Dessa forma, assim como a melhoria de processos evolui, a medição também deve evoluir para que seja capaz de fornecer as informações necessárias à tomada de decisão de acordo com os objetivos da organização. Na alta maturidade a evolução da medição é caracterizada pela aplicação do CEP.

O CEP utiliza um conjunto de ferramentas estatísticas para determinar se um processo está sob controle, sob o ponto de vista estatístico. Um processo está sob controle se seu comportamento, descrito por dados coletados para medidas de desempenho definidas para o processo, é estável, ou seja, se suas variações encontram-se dentro dos limites esperados, determinados com base em dados históricos. Um processo estável tem comportamento repetível, sendo possível prever seu desempenho em execuções futuras e, com isso, preparar planos que sejam alcançáveis e melhorá-lo continuamente. Por outro lado, um processo que apresenta variações que ultrapassam os limites esperados é instável e as causas dessas variações, chamadas causas especiais, precisam ser investigadas e tratadas através de ações de melhoria que visem à estabilização [Florac e Carleton, 1999]. Ter estabilizado seus processos é uma característica das organizações que buscam ou encontram-se na alta maturidade.

3. A Estratégia Proposta

Conforme citado na Introdução, a estratégia proposta possui três componentes: uma *Ontologia de Medição de Software* (OMS), um *Instrumento para Avaliação de Bases de Medidas considerando Adequação ao CEP* (IABM) e um *Conjunto de Recomendações para Medição de Software Adequada ao CEP* (CRMS).

A OMS representa o conhecimento útil e necessário referente ao domínio medição de software, considerando tanto aspectos da medição tradicional quanto em alta maturidade, que é a base para a definição e a utilização dos demais componentes da estratégia, uma vez que ela estabelece um vocabulário comum sobre o domínio em questão. O IABM apoia a avaliação de bases de medidas existentes e, quando apropriado, sugere ações que podem ser realizadas para adequá-las para o CEP. O CRMS, por sua vez, auxilia a realização do processo de medição de maneira adequada ao CEP, fornecendo orientações sobre aspectos relevantes à medição nesse contexto. Suas recomendações baseiam-se principalmente nos requisitos presentes no IABM e na conceituação provida pela OMS. Na Figura 1 é apresentada a visão geral da estratégia, incluindo as principais aplicações dos componentes e as relações entre eles (identificadas pelas linhas mais espessas na figura).

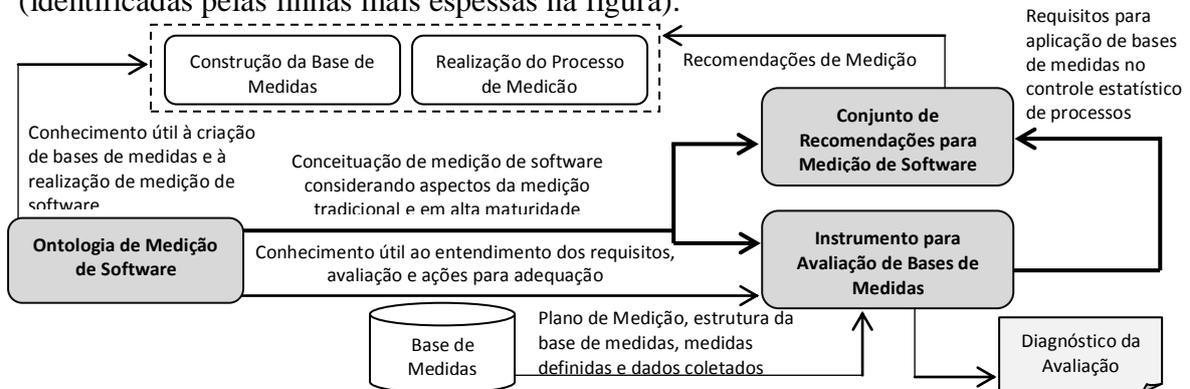


Figura 1 – Visão geral da estratégia proposta.

A aplicação da estratégia proposta em ambientes organizacionais considera dois cenários: (i) organizações que já alcançaram os níveis iniciais de maturidade e, com

isso, possuem bases de medidas com dados coletados ao longo de seus projetos; e (ii) organizações que estão iniciando um programa de melhoria de processos e desejam desde o início definir bases de medidas e realizar medições adequadas aos níveis mais elevados de maturidade. Organizações que se encontram no cenário (i) podem utilizar o IABM para avaliar e adequar, quando possível, suas bases de medidas. A Figura 2 mostra essa aplicação da estratégia.

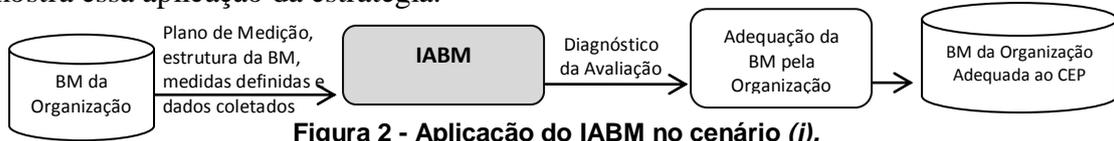


Figura 2 - Aplicação do IABM no cenário (i).

Fundamentalmente, a avaliação de bases de medidas é uma abordagem reativa que busca verificar a adequação de medidas já coletadas e armazenadas, identificar problemas e propor soluções. Essa é a abordagem possível de preparação para o CEP quando a organização dispõe de uma base de medidas resultante de medições realizadas em níveis anteriores de maturidade. Entretanto, não dispor de uma base de medidas adequada pode retardar em muito o início do CEP, sendo, então, necessária uma abordagem pró-ativa, que busque desde o início de um Programa de Medição a definição, coleta e armazenamento de medidas de forma adequada para que, no futuro, a organização possa realizar o CEP. Nesse sentido, organizações que se encontram no cenário (ii) podem utilizar o CRMS e a OMS para definirem sua base de medidas e seu Plano de Medição adequados ao CEP. Nesse caso, como mostra a Figura 3, a OMS e o CRMS fornecem o conhecimento necessário à definição e realização de medições, bem como à criação de uma base de medidas adequada ao CEP.

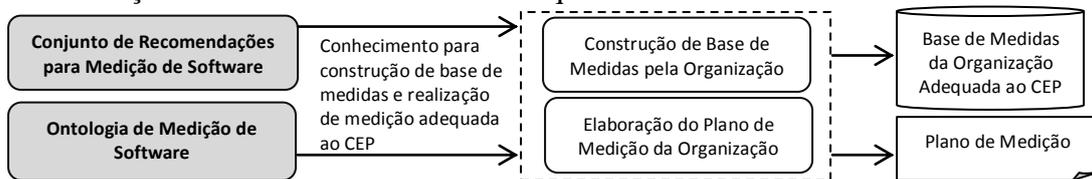


Figura 3 - Aplicação da OMS e do CRMS no cenário (ii).

A realização do CEP em uma organização requer a coleta frequente de dados para as medidas e, algumas vezes, a definição de novas medidas. Sendo assim, as abordagens de aplicação da estratégia descritas devem ser complementadas com a manutenção da adequação da base de medidas e da medição de software. Para isso, quando dados forem coletados para as medidas definidas ou quando novas medidas forem identificadas, considerando necessidades de informação organizacional ou de projetos, a adequação da base de medidas deve ser revista, utilizando-se o IABM. Além disso, as medidas identificadas devem ser definidas e coletadas adequadamente, segundo o conhecimento presente no CRMS e na OMS. A Figura 4 ilustra essa situação.

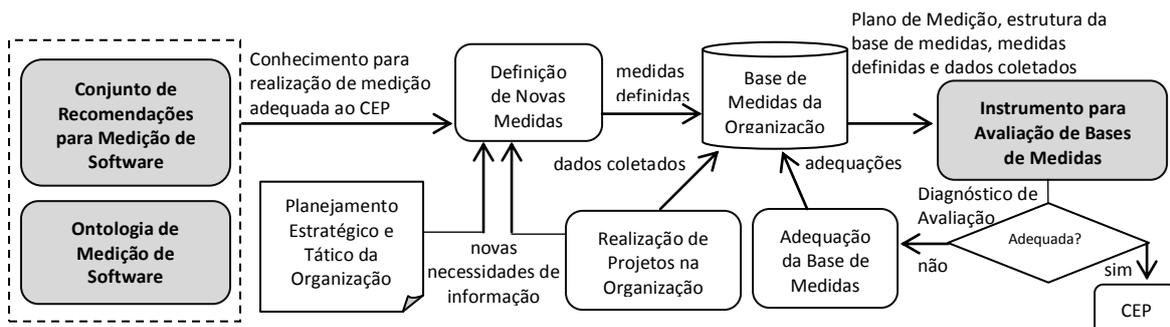


Figura 4 - Manutenção da adequação da base de medidas e da medição de software.

4. A Ontologia de Medição de Software

Ao se decidir pela definição de uma estratégia de apoio à medição de software, percebeu-se a necessidade de selecionar um vocabulário comum a ser utilizado. No entanto, sendo considerada uma disciplina relativamente recente, ainda não foram estabelecidos padrões consensuais para a medição de software. Terminologias, conceitos, procedimentos e métodos de medição de software vêm sendo definidos na última década, porém, em particular, não há consenso para conceitos e terminologias, havendo duplicações e inconsistências nas propostas encontradas na literatura, inclusive nos termos mais comuns da área. Além disso, nenhuma das propostas encontradas oferece uma visão completa da medição de software [García *et al.*, 2006].

Considerando essas limitações, foi definida uma Ontologia de Medição de Software (OMS) [Barcellos *et al.*, 2009; Barcellos *et al.*, 2010] para representar o conhecimento relevante a esse domínio e prover um vocabulário comum que abrange aspectos relevantes a todo o processo de medição de software, abordando tanto a medição tradicional quanto em alta maturidade. Para isso, a ontologia foi definida tendo como base as diversas propostas de conceitos e terminologias da literatura e os requisitos específicos da medição de software em alta maturidade. Além disso, sendo a OMS uma ontologia de domínio, buscando dotá-la de fidedignidade ao mundo real e de clareza conceitual, sua construção seguiu a abordagem proposta por Guarino (1998), tendo sido tomados como base conceitos genéricos de uma ontologia de fundamentação, a UFO (*Unified Foundational Ontology*) [Guizzardi, 2005].

A OMS foi construída seguindo o processo proposto por SABiO - *Systematic Approach for Building Ontologies* [Falbo, 2004], cujas principais atividades são: Identificação do Propósito e Especificação de Requisitos, Captura da Ontologia, Formalização da Ontologia, Integração com Ontologias Existentes, Avaliação da Ontologia e Documentação.

Para abranger o escopo necessário, OMS foi dividida em sete subontologias: *Subontologia de Entidades Mensuráveis*, que trata das entidades que podem ser submetidas à medição e de suas propriedades que podem ser medidas; *Subontologia de Medidas de Software*, que trata da definição de medidas de software; *Subontologia de Objetivos de Medição*, que trata do alinhamento da medição de software com os objetivos estratégicos; *Subontologia de Definição Operacional de Medidas*, que trata do detalhamento de aspectos relacionados à coleta e análise de medidas, estabelecido por uma organização de acordo com seus objetivos de medição; *Subontologia de Medição de Software*, que trata da medição propriamente dita, ou seja, a coleta e armazenamento dos dados para as medidas; *Subontologia de Resultados da Medição*, que trata da análise dos dados coletados para as medidas para obtenção das informações de apoio às decisões; e *Subontologia de Comportamento de Processos*, que trata da aplicação dos resultados da medição na análise do comportamento de processos.

Durante a definição da OMS, alguns conceitos relacionados aos processos de software e às organizações de software foram necessários. Esses conceitos foram reutilizados da Ontologia de Processos de Software definida em [Bertollo, 2006] e da Ontologia de Organização de Software definida em [Villela, 2004]. Para utilizar os conceitos dessas ontologias, assim como a OMS, elas deveriam ser baseadas em UFO. Em relação à Ontologia de Processos de Software, foi utilizada sua evolução adequada à UFO apresentada em [Guizzardi *et al.*, 2008]. Já em relação à Ontologia de Organização de Software, foi realizada uma reengenharia do substrato da ontologia considerado

relevante à OMS, adequando-o à UFO [Barcellos e Falbo, 2009]. Na Figura 5 são apresentadas as subontologias que compõem OMS e as ontologias a ela integradas.

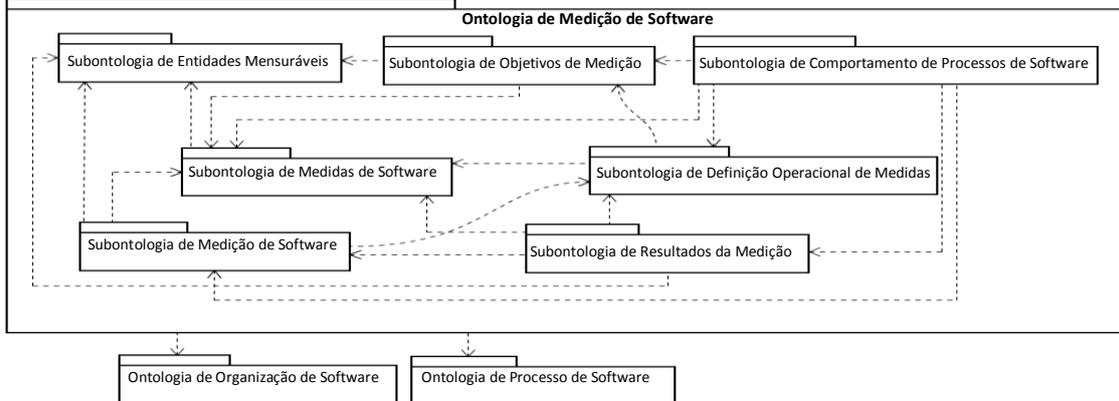


Figura 5 - Ontologia de Medição de Software: subontologias e ontologias integradas.

Seguindo o processo definido em SABiO [Falbo, 2004], a identificação do propósito e especificação dos requisitos de cada subontologia foi realizada através da definição de questões de competência, às quais cada subontologia deve ser capaz de responder (por exemplo, para a Subontologia de Objetivos de Medição: Que medidas podem ser utilizadas como indicadores para analisar o alcance a um objetivo?). A captura e a formalização das subontologias foram realizadas utilizando-se modelos UML, descrições textuais e axiomas. A avaliação foi realizada através da identificação dos conceitos, relações e axiomas necessários para responder a cada questão de competência estabelecida, visando a um compromisso ontológico mínimo, isto é, obter subontologias compostas somente por conceitos, relações e axiomas realmente necessários ao atendimento do propósito delimitado pelas questões de competência. Por fim, para avaliar a aplicação das subontologias no mundo real, foram realizadas instanciações de seus conceitos utilizando-se indivíduos obtidos a partir de dados da literatura e de dados presentes em bases de medidas de organizações. Devido à limitação de espaço, as subontologias da OMS não são descritas aqui.

5. O Instrumento para Avaliação de Bases de Medidas

Apesar de a literatura apontar que um dos maiores obstáculos para se alcançar os níveis mais altos de maturidade se deve a dificuldades relacionadas à medição, não foram encontrados registros de propostas satisfatórias de avaliação de medidas para aplicação no CEP. Assim, decidiu-se pela definição de um Instrumento para Avaliação de Bases de Medidas existentes nas organizações e adequação destas para a realização do CEP. Na Figura 6 é apresentado o processo seguido para desenvolver o IABM.

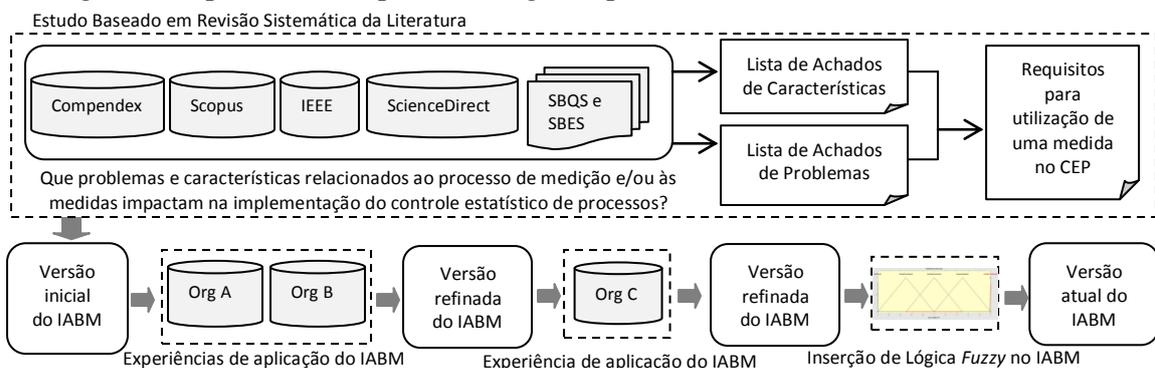


Figura 6. Desenvolvimento do IABM.

Para desenvolver o IABM, inicialmente foi realizado um estudo baseado em revisão sistemática da literatura a partir do qual foram obtidas duas listas de achados: uma de fatores relacionados às medidas e/ou à medição que influenciam positivamente na implementação do CEP (chamados de *características* no estudo) e uma de fatores que influenciam negativamente (chamados de *problemas* no estudo). A partir dessas listas de achados foi identificado um conjunto de requisitos considerados necessários para a aplicação de uma medida no CEP.

O conjunto de requisitos identificado foi, então, utilizado para a criação da primeira versão do IABM, formada por um *checklist* para avaliar cada medida da base de medidas e os dados para ela coletados. Essa versão inicial foi avaliada através de experiências de aplicação em bases de medidas de duas organizações [Barcellos e Rocha, 2008b, a]. O principal objetivo da avaliação da versão inicial foi verificar se os requisitos identificados eram adequados. Para isso, os principais questionamentos realizados foram: (a) Uma medida que atende aos requisitos do IABM pode, realmente, ser aplicada no CEP de forma satisfatória? e (b) Uma medida que não atende aos requisitos do IABM é, realmente, não adequada ao CEP?

Para responder a essas questões, as medidas foram submetidas à avaliação utilizando-se o *checklist* do IABM. Em seguida, os dados das medidas avaliadas foram aplicados em gráficos de controle. Os resultados obtidos com a aplicação dos dados das medidas em gráficos de controle foram ao encontro das avaliações realizadas pelo IABM. Ou seja, as medidas consideradas adequadas ao CEP segundo a avaliação pelo IABM puderam corretamente ser utilizadas nos gráficos de controle e forneceram informações sobre o desempenho dos processos, úteis aos objetivos da organização. Em contrapartida, as medidas que não foram consideradas adequadas ao CEP segundo a avaliação do IABM não puderam ser utilizadas nos gráficos de controle, ou, quando puderam, não foram capazes de descrever o desempenho dos processos e fornecer informações relevantes aos objetivos da organização.

Apesar dos resultados das experiências iniciais de aplicação do IABM terem mostrado, considerando as bases avaliadas, que os requisitos identificados eram adequados, durante a aplicação do IABM percebeu-se que seria necessário fazer uma reestruturação em sua forma de aplicação, uma vez que, para avaliar as medidas propriamente ditas e os dados para elas coletados, foi necessário avaliar também o Plano de Medição e a estrutura da base de medidas. Assim, o IABM foi evoluído, passando a ser composto por quatro *checklists*: um para avaliação do Plano de Medição, um para avaliação da estrutura da base de medidas, um para avaliação das medidas definidas e um para avaliação dos dados coletados para as medidas. Além dessa alteração, na segunda versão do IABM foram descritos detalhadamente os procedimentos de avaliação de cada requisito, bem como ações corretivas possíveis quando um requisito não é atendido.

A segunda versão do IABM foi, então, aplicada para avaliar a base de medidas de uma terceira organização. Essa experiência revelou, ainda, a necessidade de alguns pequenos ajustes no IABM. Os ajustes identificados foram considerados simples, uma vez que estavam relacionados ao texto do IABM, buscando tornar seu entendimento mais claro. Finalmente, após essa experiência de aplicação do IABM, considerando que a avaliação de uma base de medidas tem caráter subjetivo, foram incluídos no IABM alguns princípios da Lógica *Fuzzy* para determinar quão adequada ao CEP é uma base

de medidas. Essa alteração resultou na versão atual do IABM, cuja visão geral é apresentada na Figura 7.

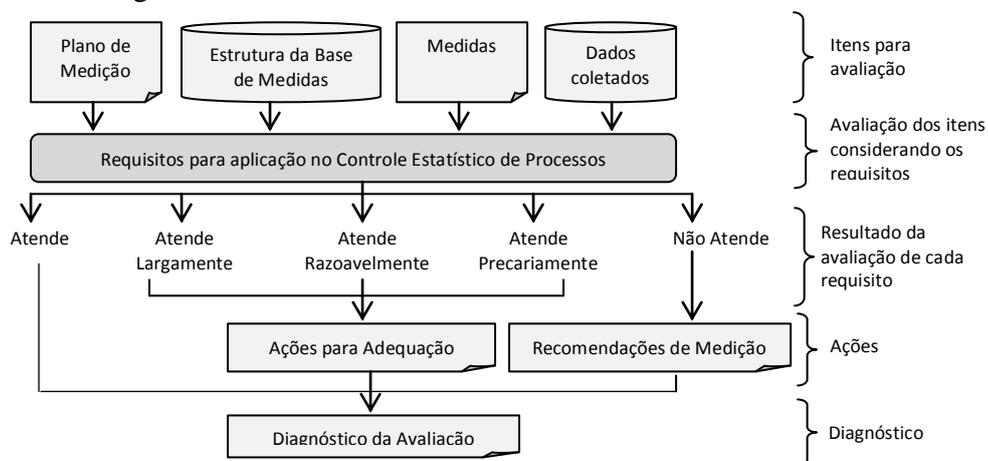


Figura 7. Visão geral do IABM.

A avaliação de uma base de medidas utilizando-se o instrumento definido é composta pela avaliação de quatro itens: o Plano de Medição, a estrutura da base de medidas, as medidas propriamente ditas e os dados coletados para essas medidas. Além disso, uma vez que, de acordo com a abordagem de melhoria de processos de software para a alta maturidade, somente os processos considerados críticos para a organização devem ser submetidos ao CEP, é desejável que a organização identifique esses processos antes da avaliação da base de medidas, a fim de evitar a avaliação desnecessária de medidas não relacionadas a esses processos ou a tendência à escolha de processos que tenham medidas aplicáveis, porém que não sejam críticos.

Cada item considerado pelo instrumento é submetido à avaliação utilizando-se um *checklist* que inclui um conjunto de requisitos. A avaliação de cada item segundo cada requisito pode produzir um dos seguintes resultados: (i) *Atende*: o item satisfaz totalmente o requisito e nenhuma ação de alteração do item avaliado é necessária em relação ao requisito considerado; (ii) *Atende Largamente*, *Atende Razoavelmente* ou *Atende Precariamente*: o item não satisfaz o requisito, mas é possível realizar ações que irão adequá-lo a fim de satisfazer o requisito em questão e, conseqüentemente, permitir sua utilização no CEP. O grau de atendimento do item ao requisito (*Largamente*, *Razoavelmente* ou *Precariamente*) está diretamente relacionado com o esforço necessário para realizar as ações que levarão o item a atender o requisito em questão. Quanto mais esforço, menor o grau de atendimento; e (iii) *Não Atende*: o item não satisfaz o requisito e não há ações possíveis para adequar o item avaliado ao CEP, sendo necessário descartá-lo e redefini-lo, se pertinente.

Quando o resultado da avaliação de um requisito é *Atende Largamente*, *Atende Razoavelmente* ou *Atende Precariamente*, são sugeridas *Ações para Adequação*. Essas ações são orientações providas à organização que visam à realização de correções que permitam a utilização do item avaliado no CEP. Quando o resultado da avaliação de um requisito é *Não Atende*, não há ações de adequação possíveis e o item deve ser descartado da utilização no CEP. Nesse caso, a organização pode ser orientada sobre como é possível atender ao referido requisito através de *Recomendações de Medição* contidas no *Conjunto de Recomendações para Medição de Software*, outro componente da estratégia, descrito na próxima seção deste artigo.

Os resultados da avaliação de uma base de medidas são registrados em um documento denominado *Diagnóstico de Avaliação*, que inclui, além da avaliação detalhada de cada item, sugestões das ações de adequação possíveis e o grau de adequação da base de medidas como um todo ao CEP, dado em percentual. Para determinar o grau de adequação de uma base de medidas são utilizados os princípios da Lógica *Fuzzy* e dos Conjuntos *Fuzzy*.

Na Figura 8 é apresentado o *checklist* utilizado para avaliar as medidas. Ele é aplicado para cada medida avaliada. Os *checklists* para avaliação do Plano de Medição, da estrutura da base de medidas e dos dados coletados para as medidas não são apresentados neste artigo.

Item: Medida _____ Medida avaliada: _____

Legenda: A = Atende; AL = Atende Largamente; AR = Atende Razoavelmente; AP = Atende Precariamente; NA = Não Atende; NFPA = Não foi possível avaliar

Requisitos	Avaliação					
1. A definição operacional da medida é correta e satisfatória.	<input type="checkbox"/>					
A definição operacional da medida inclui corretamente:						
1.1 Definição da medida.	<input type="checkbox"/>					
1.2 Entidade medida.	<input type="checkbox"/>					
1.3 Propriedade medida.	<input type="checkbox"/>					
1.4 Unidade de medida.	<input type="checkbox"/>					
1.5 Tipo de escala.	<input type="checkbox"/>					
1.6 Valores da escala.	<input type="checkbox"/>					
1.7 Intervalo esperado dos dados.	<input type="checkbox"/>					
1.8 Fórmula(s) (se aplicável).	<input type="checkbox"/>					
1.9 Descrição precisa do procedimento de medição.	<input type="checkbox"/>					
1.10 Responsável pela medição.	<input type="checkbox"/>					
1.11 Momento da medição.	<input type="checkbox"/>					
1.12 Periodicidade de Medição.	<input type="checkbox"/>					
1.13 Descrição precisa do procedimento de análise (se indispensável).	<input type="checkbox"/>					
1.14 Periodicidade da análise (se aplicável).	<input type="checkbox"/>					
2. A medida está alinhada a objetivos dos projetos ou da organização.	<input type="checkbox"/>					
A medida está associada a:						
2.1 Objetivo da organização.	<input type="checkbox"/>					
2.2 Objetivo dos projetos.	<input type="checkbox"/>					
3. Os resultados da análise da medida são relevantes às tomadas de decisão.	<input type="checkbox"/>				<input type="checkbox"/>	<input type="checkbox"/>
4. Os resultados da análise da medida são úteis à melhoria de processo.	<input type="checkbox"/>				<input type="checkbox"/>	<input type="checkbox"/>
5. A medida está relacionada ao desempenho de um processo.	<input type="checkbox"/>				<input type="checkbox"/>	<input type="checkbox"/>
6. A medida está relacionada a um processo crítico.	<input type="checkbox"/>				<input type="checkbox"/>	<input type="checkbox"/>
7. A medida está associada a uma atividade ou processo que produz item mensurável.	<input type="checkbox"/>				<input type="checkbox"/>	<input type="checkbox"/>
8. As medidas correlatas à medida estão definidas.	<input type="checkbox"/>					
9. As medidas correlatas à medida são válidas.	<input type="checkbox"/>					
10. A medida possui baixa granularidade.	<input type="checkbox"/>					
11. A medida é passível de normalização (se aplicável).	<input type="checkbox"/>					
12. A medida está normalizada corretamente (se aplicável).	<input type="checkbox"/>					
13. Os critérios de agrupamento de dados para análise da medida estão definidos.	<input type="checkbox"/>					
14. A medida não considera dados agregados.	<input type="checkbox"/>					

Figura 8. *Checklist* para avaliação das medidas de software.

6. O Conjunto de Recomendações para Medição de Software

No contexto de modelos de maturidade, à medida que o nível de maturidade dos processos aumenta, novas necessidades são identificadas e novas medidas são definidas. Então, pode-se dizer que o processo de medição e as medidas utilizadas em uma organização orientam a evolução de um nível de maturidade para outro [Dumke *et al.*, 2004]. No entanto, realizar medição de software adequada ao CEP somente quando as práticas relacionadas à alta maturidade são iniciadas pode retardar sua implantação,

pois, para obter o volume de dados requerido para aplicar uma medida no CEP, medidas adequadas a esse contexto (as quais, geralmente, não são definidas nos níveis iniciais de maturidade) devem ser definidas e os processos devem ser executados diversas vezes em projetos da organização.

Uma maneira de tratar essa questão é realizar medição de software adequada ao CEP desde os níveis anteriores à alta maturidade. Entretanto, ainda não há um conjunto bem estabelecido de orientações com esse propósito. Sendo assim, decidiu-se por definir um Conjunto de Recomendações para Medição de Software Adequada ao CEP, que visa auxiliar as organizações na implementação de um processo de medição adequado ao CEP, fornecendo orientações relacionadas a aspectos considerados relevantes à medição realizada nesse contexto.

As recomendações que compõem o CRMS foram definidas, principalmente, com base nos requisitos do IABM, na conceituação provida pela OMS e no conhecimento obtido através das experiências de utilização desses componentes. Foram consideradas, ainda, orientações, práticas e lições aprendidas registradas em normas, padrões e relatos da literatura. A base que fundamentou a definição do CRMS é mostrada na Figura 9. Os aspectos tratados pelas recomendações foram identificados a partir dos requisitos presentes no IABM. O conhecimento provido pela OMS, pelas experiências de utilização do instrumento e da ontologia, por registros da literatura e por normas e padrões de medição foi utilizado para compor as recomendações.

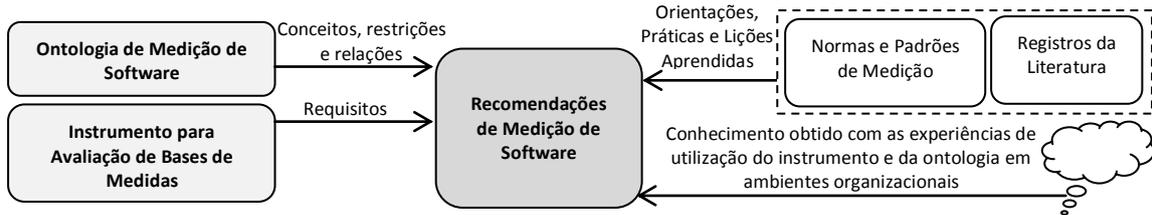


Figura 9. Definição do Conjunto de Recomendações para Medição de Software.

O CRMS é composto por vinte recomendações organizadas em cinco grupos. Na Figura 10 é apresentada a visão geral do CRMS, onde são identificados os aspectos tratados pelas recomendações que compõem cada um de seus grupos. Em seguida, cada grupo de recomendações é descrito.

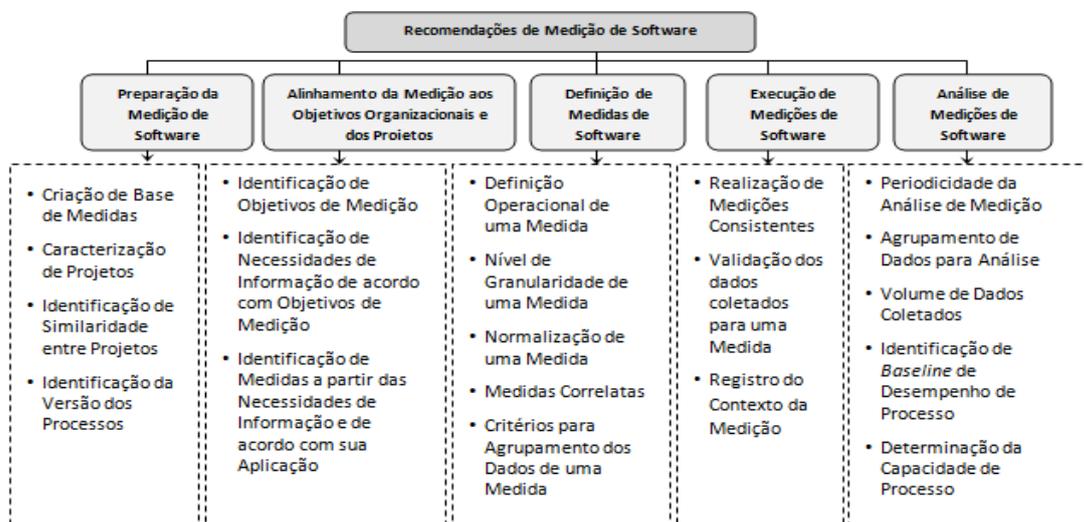


Figura 10. Visão Geral do Conjunto de Recomendações para Medição de Software.

- *Preparação da Medição de Software*: contém recomendações relacionadas a aspectos que devem ser considerados antes da implantação de medição em uma organização e sem os quais não é possível realizar a medição de forma adequada.
- *Alinhamento da Medição de Software aos Objetivos Organizacionais e dos Projetos*: contém recomendações que visam à realização de medições alinhadas aos objetivos de negócio da organização e aos objetivos específicos dos projetos. Abordam a elaboração do Plano de Medição da organização e dos projetos considerando o Planejamento Estratégico da organização. Para isso, guiam a identificação de medidas úteis e alinhadas aos objetivos estabelecidos.
- *Definição de Medidas de Software*: contém recomendações para definir medidas adequadamente. Inclui o estabelecimento de definições operacionais para as medidas e trata outros aspectos relevantes, a saber: nível de granularidade, normalização, medidas correlatas e critérios para agrupamento de dados das medidas.
- *Realização de Medições de Software*: contém recomendações para a execução de medições de software, que consiste da coleta e armazenamento de dados para as medidas.
- *Análise de Medições de Software*: contém recomendações para que a análise dos dados coletados para as medidas forneça as informações necessárias identificadas no Plano de Medição, apoiando, assim, a tomada de decisões e a identificação de ações corretivas e de melhoria.

A avaliação do CRMS foi realizada por especialistas através da técnica de Revisão por Pares, tendo sido selecionados para a revisão os avaliadores MR MPS [Softex, 2009] habilitados a realizarem avaliações dos níveis A e B do modelo, uma vez que possuem conhecimento teórico e prático da utilização do CEP em organizações de alta maturidade. Considerando os resultados da avaliação, o CRMS foi revisto. A seguir na Tabela 1, como exemplo, são apresentadas as recomendações que tratam o aspecto *Definição Operacional de uma Medida*, presente no grupo *Definição de Medidas de Software*.

Tabela 1. Recomendações definidas para *Definição Operacional de uma Medida*.

<p>Propósito: Orientar a elaboração da definição operacional de uma medida. A definição operacional de uma medida inclui informações detalhadas sobre a medida, principalmente no que diz respeito à sua coleta e análise.</p>
<p>Fundamentação Teórica: A repetitividade da medição de uma medida está diretamente relacionada com a completeza e precisão de sua definição operacional. Uma definição operacional incompleta, ambígua ou fracamente documentada possibilita que diferentes pessoas entendam a medida de maneiras diferentes e, conseqüentemente, colem dados inválidos, realizem medições incomparáveis ou análises incorretas, o que torna a medição inconsistente e ineficiente [Kitchenham <i>et al.</i>, 2001].</p> <p>A definição operacional de uma medida deve ser estabelecida de acordo com sua aplicação. Por exemplo, medidas aplicadas na análise de desempenho de processos, diferentemente das medidas com aplicação no monitoramento e controle tradicionais, devem incluir em seu procedimento de análise a utilização de técnicas do controle estatístico de processos.</p> <p>Em uma organização que deseja definir e coletar medidas adequadas ao controle estatístico de processos desde os níveis iniciais de maturidade, as definições operacionais das medidas, inicialmente, são orientadas ao monitoramento e controle tradicionais. Porém, para que os dados coletados para essas medidas sejam futuramente úteis ao controle estatístico de processos, as definições operacionais das medidas devem garantir que os dados coletados e armazenados sejam úteis ao controle estatístico de processos [Barcellos <i>et al.</i>, 2009].</p>

Tabela 1. Recomendações definidas para *Definição Operacional de uma Medida* (cont.).

Recomendações:

R1. Estabelecer uma definição operacional para as medidas, a qual inclua as seguintes informações:

- i. *Nome*: nome da medida.
- ii. *Definição*: descrição sucinta da medida.
- iii. *Mnemônico*: sigla utilizada para identificar a medida.
- iv. *Tipo de Medida*: classificação da medida quanto à sua dependência funcional, podendo uma medida ser uma medida base ou uma medida derivada.
 - v. *Entidade Medida*: entidade que a medida mede. Exemplos: organização, projeto, processo, atividade, recurso humano, recurso de hardware, recurso de software e artefato, dentre outros.
 - vi. *Propriedade Medida*: propriedade da entidade medida quantificada pela medida. Exemplos: tamanho, custos, defeitos, esforço etc.
 - vii. *Unidade de Medida*: unidade de medida em relação à qual a medida é medida. Exemplos: pessoa/mês, pontos de função, reais etc.
- viii. *Tipo de Escala*: natureza dos valores que podem ser atribuídos à medida. Exemplos: escala nominal, escala intervalar, escala ordinal, escala absoluta e escala taxa.
- ix. *Valores da Escala*: valores que podem ser atribuídos à medida. Exemplos: números reais positivos etc. Para medidas com escala do tipo absoluta ou taxa, ao determinar os valores da escala, é preciso identificar a precisão a ser considerada (0, 1 ou 2 casas decimais).
- x. *Intervalo esperado dos dados*: limites de valores da escala definida de acordo com dados históricos ou com metas estabelecidas. Exemplo: [0, 10].
- xi. *Procedimento de Medição*: descrição do procedimento que deve ser realizado para coletar uma medida. A descrição do procedimento de medição deve ser clara, objetiva e não ambígua.
- xii. *Fórmula de Cálculo de Medida*: fórmula utilizada no procedimento de medição de medidas derivadas, para calcular o valor atribuído à medida considerando-se sua relação com outras medidas ou com outros valores. Exemplo: aderência ao cronograma = tempo real / tempo estimado.
- xiii. *Responsável pela Medição*: papel desempenhado pelo recurso humano responsável pela coleta da medida. É importante que o responsável pela medição seja fonte direta das informações a serem fornecidas na medição. Exemplos: analista de sistemas, programador, gerente do projeto etc.
- xiv. *Momento da Medição*: momento em que deve ser realizada a coleta e registro de dados para a medida. O momento da coleta deve ser uma atividade do processo definido para o projeto ou de um processo organizacional. Exemplos: na atividade Homologar Especificação de Requisitos, na atividade Realizar Testes de Unidade etc.
- xv. *Periodicidade de Medição*: frequência de coleta da medida. Exemplos: diária, mensal, uma vez por fase, uma vez por projeto, uma vez em cada ocorrência da atividade designada como momento da medição etc. É indispensável que haja coerência entre a periodicidade de medição e o momento de medição.
- xvi. *Procedimento de Análise*: descrição do procedimento que deve ser realizado para representar e analisar os dados coletados para uma medida, incluindo, além do procedimento propriamente dito, as ferramentas analíticas que devem ser utilizadas (por exemplo: histograma, gráfico de controle XmR etc.). A descrição do procedimento de análise deve ser clara, objetiva e não ambígua. Um procedimento de análise de medição pode ser baseado em critérios de decisão (por exemplo, utilizando-se uma meta como referência) e, nesse caso, os critérios de decisão considerados (incluindo suas premissas e conclusões) devem ser claramente estabelecidos. Medidas que não são analisadas isoladamente não precisam ter procedimento de análise definido. Por exemplo: se a medida número de requisitos alterados só for submetida à análise quando utilizada na composição da medida taxa de alteração de requisitos, não há necessidade de definir seu procedimento de análise.
- xvii. *Momento da Análise de Medição*: momento em que deve ser realizada a análise de dados coletados para a medida. O momento da análise deve ser uma atividade do processo definido para o projeto ou de um processo organizacional como, por exemplo, em atividades de monitoramento de projeto.
- xviii. *Periodicidade da Análise*: frequência de análise de dados da medida. Exemplos: diária, mensal, uma vez por fase, uma vez por projeto, uma vez em cada ocorrência da atividade designada como momento da análise etc. É indispensável que haja coerência entre a periodicidade de análise de medição e o momento da análise de medição¹.

¹ A *periodicidade de análise* é um aspecto tratado em detalhes em outras recomendações do CRMS.

Tabela 1. Recomendações definidas para *Definição Operacional de uma Medida* (cont).

<p><i>xiv. Responsável pela Análise:</i> papel desempenhado pelo recurso humano responsável pela análise da medida. É importante que o responsável pela análise de medição seja apto a aplicar o procedimento de análise e tenha conhecimento organizacional que propicie a correta interpretação dos dados e fornecimento de informações que apoiem as tomadas de decisão. Exemplos: gerente do projeto, gerente de qualidade etc.</p>
<p>R2. Estabelecer a definição operacional da medida de acordo com sua aplicação.</p>
<p>R3. Estabelecer, para as medidas identificadas nos níveis iniciais de maturidade, mas que poderão ser futuramente utilizadas no controle estatístico dos processos, definições operacionais que permitam desde os níveis iniciais, coleta e armazenamento frequente e adequado dos dados necessários à realização do controle estatístico de processos.</p>
<p>R4. Para utilizar no controle estatístico de processos medidas identificadas nos níveis iniciais de maturidade, estabelecer novas definições operacionais voltadas para aplicação na análise de desempenho dos processos, incluindo, por exemplo, um procedimento de análise adequado, que utilize as técnicas do controle estatístico de processos.</p>

7. Trabalhos Correlatos

Considerando trabalhos correlatos, em relação à OMS tem-se as ontologias de medição de software propostas em [García *et al.*, 2006] e em [Dalmoro, 2006]. No entanto, em ambas as propostas os autores não consideram aspectos relacionados à alta maturidade nem utilizam uma ontologia de fundamentação como base. Em relação à avaliação da base de medidas, embora tenham sido encontrados vários trabalhos que destacam essa necessidade, apenas um [Tarhan e Demirors, 2006] inclui especificamente uma proposta para avaliação de medidas considerando CEP. Nele os autores propõem uma abordagem para seleção dos processos que serão submetidos ao CEP que considera a existência de medidas úteis um dos critérios para a seleção de um processo. Apesar de incluir avaliação de medidas, até onde teve-se acesso, a abordagem proposta pelos autores é bastante limitada nesse contexto e os próprios autores afirmam que ela não é suficiente para selecionar medidas adequadas ao CEP. Por fim, no contexto do CRMS, foram encontrados poucos trabalhos que tratam da medição em alta maturidade, destacando-se as iniciativas de Dumke [Dumke *et al.*, 2004]. No entanto, seus trabalhos tratam mais especificamente da aplicação dos dados das medidas no CEP do que da definição das medidas e coleta dos dados.

8. Considerações Finais

Neste artigo foi apresentada uma estratégia proposta em uma tese de doutorado com o objetivo de auxiliar as organizações que buscam a alta maturidade em seus processos na avaliação de suas bases de medidas e realização de medição adequada a ao CEP. A estratégia é composta por três componentes: uma Ontologia de Medição de Software, um Instrumento para Avaliação de Bases de Medidas e um Conjunto de Recomendações para Medição.

Dada a limitação de tempo, no contexto da tese de doutorado não foi possível utilizar toda a estratégia proposta em uma mesma organização desde os níveis iniciais de maturidade até a realização do CEP nos níveis mais elevados. Considerando esse fato e, além disso, a existência, no Brasil, de poucas organizações que estão implementando as práticas da alta maturidade, para avaliar a estratégia, conforme descrito no artigo, foram realizadas avaliações intermediárias, à medida que seus componentes foram desenvolvidos. Os resultados obtidos foram ao encontro dos

resultados esperados, ou seja, considerando as avaliações realizadas, os componentes se mostraram úteis às organizações e adequados ao contexto considerado.

Como trabalho futuro, planeja-se a utilização de toda a estratégia em ambientes organizacionais. Considerando cada componente individualmente, em relação ao IABM, planeja-se a implementação de uma ferramenta de apoio (atualmente o IABM é composto, basicamente, por planilhas eletrônicas) e a realização de novas experiências de aplicação. No contexto do CRMS, planeja-se sua utilização em ambientes organizacionais, o que permitirá avaliá-lo sob a ótica das organizações, levando em consideração aspectos técnicos (adequação das medidas e dados coletados ao se utilizar as recomendações) e organizacionais (impacto da utilização do CRMS na implementação dos níveis mais elevados de maturidade, por exemplo, em relação a tempo e custos). Por fim, em relação à OMS, encontra-se em definição em uma dissertação de mestrado um repositório de medidas adequado à alta maturidade, que utiliza como principal base a conceituação provida pela OMS.

Referências

- BARCELLOS, M. P., FALBO, R. A. (2009) *"Using a Foundational Ontology for Reengineering a Software Enterprise Ontology"*, Lecture Notes in Computer Science, v. 5833, p. 179-188.
- BARCELLOS, M. P., FALBO, R. A., DALMORO, R. (2010) *"A Well-Founded Software Measurement Ontology"*, In Proceedings of the 6th International Conference on Formal Ontology in Information Systems (FOIS 2010), Toronto - Canadá.
- BARCELLOS, M. P., ROCHA, A. R., FALBO, R. A. (2009) *"An Ontology-based Approach for Software Measurement and Suitability Measures Basis Evaluation to Apply Statistical Software Process Control in High Maturity Organizations"*, In Proceedings of the ER2009 PhD Colloquium, Gramado - RS.
- BARCELLOS, M. P., ROCHA, A. R. (2008a) *"Avaliação de Bases de Medidas considerando sua Aplicabilidade ao Controle Estatístico de Processos de Software"*, VII Simpósio Brasileiro de Qualidade de Software (SBQS'08), Florianópolis – SC.
- BARCELLOS, M. P., ROCHA, A. R. (2008b) *"Uma Abordagem de Apoio à Realização de Controle Estatístico de Processos de Software em Organizações de Alta Maturidade"*, XXXIV Conferência Latinoamericana de Informática (CLEI'08), Santa Fé - Argentina.
- BERTOLLO, G. (2006) *"Definição de Processos em um Ambiente de Desenvolvimento de Software"*, Tese de Mestrado, Departamento de Informática, Centro Tecnológico, Universidade Federal do Espírito Santo (UFES), Vitória - ES, Brasil.
- CARD, D. N. (2004) *"Statistical Techniques for Software Engineering Practice"*, In Proceedings of the 26th International Conference on Software Engineering - ICSE'2004, Scotland, UK, p. 722-723.
- CHRISISS, M. B., KONRAD, M., SHRUM, S. (2006) *"CMMI (Second Edition): Guidelines for Process Integration and Product Improvement"*, Addison-Wesley.
- DALMORO, R., 2008, *"Avaliação e Melhoria de Processos de Software: Conceituação e Definição de um Processo para Apoiar a sua Automatização"*, Dissertação de Mestrado, Departamento de Informática, Centro Tecnológico, Universidade Federal do Espírito Santo, Vitória – Brasil.
- DUMKE, R., CÔTÉ, I., ANDRUSCHAK, O. T. (2004) *"Statistical Process Control (SPC) - A Metric-based Point of View of Software Processes Achieving the CMMI Level Four"*, Technical Report, Dept. of Computer Science, University of Magdeburg, Germany.

- FALBO, R. A. (2004) "*Experiences in Using a Method for Building Domain Ontologies*", In Proceedings of the Fourth International Conference on Quality Software (QSIC'2004), IEEE Computer Society, Braunschweig, Germany, p. 162-169.
- FLORAC, W. A., CARLETON, A. D. (1999) "*Measuring the Software Process: Statistical Process Control for Software Process Improvement*", Addison Wesley.
- GARCÍA, F., BERTOIA, M. F., CALERO, C., VALLECILLO, A., RUIZ, F., PIATTINI, M., GENERO, M. (2006) "*Towards a Consistent Terminology for Software Measurement Information and Software Technology*", Information and Software Technology, v. 48, n. 8, p. 631-644.
- GUARINO, N. (1998) "*Formal Ontology and Information Systems*", In Proceedings of International Conference in Formal Ontology and Information Systems, pp 3-15, 1998.
- GUIZZARDI, G. (2005) "*Ontological Foundations for Structural Conceptual Models*", Universal Press, The Netherlands, ISBN 90-75176-81-3.
- GUIZZARDI, G., FALBO, R. A., GUIZZARDI, R. S. S. (2008) "*Grounding Software Domain Ontologies in the Unified Foundational Ontology (UFO): The case of the ODE Software Process Ontology*", In Proceedings of the XI Iberoamerican Workshop on Requirements Engineering and Software Environments, Recife - Brasil.
- ISO/IEC (2002) "*ISO/IEC 15939 – 2002 (E) Software Engineering – Software Measurement Process*", International Organization for Standardization and the International Electrotechnical Commission, Geneva, Switzerland.
- ISO/IEC (2003) "*ISO/IEC 15504-2 - Information Technology – Software Process Assessment*", International Organization for Standardization and the International Electrotechnical Commission, Geneva, Switzerland.
- ISO/IEC (2008) "*ISO/IEC 12207:2008 - Systems and Software Engineering - Software Life Cycle Process*", International Organization for Standardization and the International Electrotechnical Commission, Geneva, Switzerland.
- KITCHENHAM, B., HUGHES, R. T., LINKMAN, S. G. (2001) "*Modeling Software Measurement Data*", IEEE Transactions on Software Engineering, v. 27, n. 9, p. 788-804.
- KITCHENHAM, B., KUTAY, C., JEFFERY, R., CONNAUGHTON, C. (2006) "*Lessons Learnt from the Analysis of Large-scale Corporate Databases*", In Proceedings of the 28th International Conference on Software Engineering (ICSE'06), Shanghai, China, p. 439-444.
- RACZINSKI, B., CURTIS, B. (2008) "*Software Data Violate SPC's Underlying Assumptions*", IEEE Software, v. 25, n. 3, p. 49-50.
- SARGUT, K. U., DEMIRORS, O. (2006) "*Utilization of Statistical Process Control (SPC) in Emergent Software Organizations: Pitfalls and Suggestions*", Software Quality Journal, v. 14, n. 5, p. 135-157.
- SOFTTEX (2009) "*MPS.BR: Melhoria de Processo do Software Brasileiro - Guia Geral : 2009*", Disponível em: <http://www.softex.br/mpsbr>.
- TARHAN, A., DEMIRORS, O. (2006) "*Investigating Suitability of Software Process and Metrics for Statistical Process Control*", Lecture Notes in Computer Science, v. 4257, p. 88-99.
- VILLELA, K. (2004) "*Definição e Construção de Ambientes de Desenvolvimento de Software Orientados à Organização*", Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro - Brasil.