



ABORDAGEM PARA ESPECIFICAÇÃO DE EXPERIMENTOS *IN SILICO* EM
ENGENHARIA DE *SOFTWARE* E ÁREAS AFIM

Wallace Martinho Pereira

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Ciências em Engenharia de Sistemas e Computação.

Orientador: Guilherme Horta Travassos

Rio de Janeiro
Fevereiro de 2011

ABORDAGEM PARA ESPECIFICAÇÃO DE EXPERIMENTOS *IN SILICO* EM
ENGENHARIA DE *SOFTWARE* E ÁREAS AFIM

Wallace Martinho Pereira

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM
ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Guilherme Horta Travassos, D.Sc.

Prof^a. Marta Lima de Queiros Mattoso, D.Sc.

Prof. Marcio Oliveira Barros, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
FEVEREIRO DE 2011

Martinho Pereira, Wallace

Abordagem para especificação de experimentos in silico em engenharia de software e áreas afim/ Wallace Martinho Pereira – Rio de Janeiro: UFRJ/COPPE, 2011.

XIV, 236 p.: il.; 29,7 cm.

Orientador: Guilherme Horta Travassos.

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2011.

Referências Bibliográficas: p. 141-151.

1. *Workflow* científico. 2. Concepção. 3. Experimentos *in silico*. 4. Experimentos *in virtuo*. 5. Engenharia de Software Experimental. I. Travassos, Guilherme Horta II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Aos meus Pais, exemplos e incentivo de vida.

À minha irmã e irmão, Valeska e William.

À minha namorada, Ana.

Agradecimentos

À minha Mãe, Deise, pelo amor, compreensão e apoio desde sempre.

Ao meu Pai, Adir, em memória, pelo apoio e incentivo aos estudos.

À minha Irmã e Irmão, Valeska e William, pela amizade e conselhos dados a mim durante minha vida.

À minha Namorada, pelo amor, apoio, compreensão e incentivo, ao longo de cinco anos e por toda esta pós-graduação.

Ao meu Orientador, Guilherme Travassos, pela grande dedicação, conselhos e motivação ao longo de três anos, contribuindo para minha formação profissional e como pessoa. Agradeço pela orientação, incentivo e por me conduzir durante este curso de pós-graduação e, por acreditar em mim e no meu trabalho.

Aos professores, Márcio Barros e Marta Mattoso, por participarem de minha banca de defesa de mestrado.

Especialmente aos amigos, Rodrigo Spindola, Jobson Massolar e Paulo Sergio pelas dicas, a ajuda e os conselhos durante todo o período.

Aos Companheiros da COPPE, Fortuna, Marco Antônio, Marcos, Taísa, Vitor, Rafael, Sílvia e Rafael Espírito, Karen, Breno, Eduardo, Daniel, Anderson pela amizade, sugestões e ajuda nos momentos que precisei.

À COPPE por prover a infraestrutura. Ao CNPQ pelo apoio financeiro. Ao projeto Galileu pela oportunidade de desenvolver e aplicar este trabalho. Ao CEDERJ pela oportunidade de desenvolver minha formação profissional.

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

ABORDAGEM PARA ESPECIFICAÇÃO DE EXPERIMENTOS *IN SILICO* EM
ENGENHARIA DE *SOFTWARE* E ÁREAS AFIM

Wallace Martinho Pereira

Fevereiro/2011

Orientador: Guilherme Horta Travassos

Programa: Engenharia de Sistemas e Computação

Este trabalho propõe uma abordagem desenvolvida a partir de uma pesquisa científica baseada na condução de estudos secundários e primários que possui o objetivo de apoiar a concepção de *workflow* científico, em nível abstrato, para experimentos científicos baseado em simulação. São considerados dois aspectos do contexto atual: (1) necessidade de documentação e representação em níveis mais altos de abstração e (2) métodos que auxiliem na concepção do *workflow*. Resultados de um estudo em campo indicam que esta abordagem contribui para identificação e formalização do experimento científico como um *workflow* e na identificação de defeitos nesta especificação inicial.

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

APPROACH TO SPECIFY *IN SILICO* EXPERIMENTS ON SOFTWARE
ENGINEERING AND RELATED SCIENCE AREAS

Wallace Martinho Pereira

February/2011

Advisor: Guilherme Horta Travassos

Department: Computer Science and Systems Engineering

This work proposes an approach developed by following a scientific research methodology based on the conduction of primary and secondary studies with the objective of supporting the conception of scientific workflows, on high abstraction level, for scientific experiments based on simulation. It considers two aspects of the current context: (1) the need for workflow documentation and representation on higher levels of abstraction and (2) methods supporting the conception of scientific workflows. Results of an experimental study executed in the field has indicated this approach leads to the identification and formalization of the scientific experiment as a workflow and allows the identification of defects in this initial workflow specification.

ÍNDICE

Índice de Figuras	xi
Índice de Tabelas	xiii
Capítulo 1 - INTRODUÇÃO	1
1.1 Contexto e Motivação: Descrição do Problema	1
1.2 Questões de Pesquisa.....	4
1.3 Objetivos	6
1.4 Trabalhos relacionados	7
1.5 Método de Pesquisa	9
1.6 Organização do Trabalho	10
Capítulo 2 - Conceitos sobre Experimentação baseada em simulação, <i>workflow</i> e uml	13
2.1 Fundamentação teórica	13
2.1.1 Experimentação	13
2.1.2 <i>Workflow</i> científico	16
2.1.3 Ciclo de vida do <i>workflow</i> científico e do experimento	19
2.1.4 Sistemas Gerenciadores de <i>Workflow</i> científico - SGWfC.....	22
2.2 Definição dos conceitos sobre elementos do domínio de <i>workflow</i> científico	23
2.3 Linguagens e notações para representação de <i>workflow</i> científico.....	25
2.4 Revisão da notação UML 2.....	27
2.4.1 Notação do diagrama de classes	27
2.4.2 Notação do diagrama de atividades	29
2.5 Considerações finais do capítulo	33
Capítulo 3 - <i>Quasi</i> revisão sistemática sobre composição de experimentos baseados em <i>workflow</i> científico	35
3.1 Introdução	35
3.1.1 Protocolo da <i>quasi</i> revisão Sistemática	36
3.2 Execução da <i>quasi</i> Revisão Sistemática	45
3.2.1 Primeira rodada de execução – Maio 2009	45
3.2.2 Segunda rodada de execução – Outubro 2010	50
3.2.3 Considerações sobre a execução da revisão	52
3.2.4 Avaliação da qualidade das referências incluídas:	54
3.3 Resumo das abordagens.....	56
3.4 Análise dos resultados.....	64
3.5 Considerações finais do capítulo	74

Capítulo 4 - Abordagem para concepção de <i>workflow</i> científico abstrato.....	77
4.1 Introdução	77
4.2 Descrição do procedimento de concepção de <i>workflow</i> abstrato.....	79
4.2.1 Detalhamento da fase “Especificar <i>workflow</i> científico”	80
4.2.2 Detalhamento da fase “Validar <i>workflow</i> científico”	82
4.3 Especificação de <i>workflow</i> científico em nível abstrato.....	87
4.3.1 Descrição textual através de Formulários.....	88
4.3.2 Representação gráfica do <i>workflow</i> científico.....	94
4.4 Prova de Conceito – Domínio de Evolução de <i>Software</i>	96
4.4.1 Objetivo da Avaliação.....	96
4.4.2 Descrição do Domínio	97
4.4.3 Aplicação do procedimento para composição	98
4.5 Considerações finais do capítulo	104
Capítulo 5 - Evolução da abordagem para concepção de <i>workflow</i> científico abstrato	
.....	107
5.1 Introdução	107
5.2 Modificações na abordagem para concepção.....	108
5.2.1 Modificações no procedimento de concepção	108
5.2.2 Modificações na representação textual	109
5.2.3 Modificações na representação gráfica do <i>workflow</i> científico.....	111
5.3 Heurísticas para garantia da qualidade da especificação de <i>workflow</i>	
científico	114
5.3.1 Heurísticas para organização das tarefas de detecção de defeitos	114
5.3.2 Heurísticas para verificação do conteúdo da especificação de <i>workflow</i>	117
5.4 Aplicação em Campo – Projeto GALILEU.....	121
5.4.1 Domínio de Exploração de Petróleo em Águas Profundas	121
5.4.2 Aplicação do procedimento	122
5.4.3 Análise da aplicação em campo	130
5.5 Considerações Finais do Capítulo	131
Capítulo 6 - Conclusões.....	133
6.1 Considerações finais	133
6.2 Resultados obtidos	135
6.3 Contribuições da pesquisa	137
6.4 Limitações	138
6.5 Futuras linhas de pesquisa.....	140
Referências Bibliográficas	141
Anexo A – Strings de busca da <i>quasi</i> revisão sistemática	152

Anexo B – Lista de todas as referências	163
Anexo C – Dados extraídos para análise na <i>quasi</i> -Revisão.....	205
Anexo D – Tamplate para especificação de workflow abstrato	216
Anexo E – Metamodelo para experimentação utilizando workflow científico	222

ÍNDICE DE FIGURAS

Figura 1.1. Passos do método de pesquisa aplicado no desenvolvimento da abordagem.....	11
Figura 2.1. Processo experimental, definido por WOHLIN ET. AL. (2000), estendido por AMARAL (2003).....	16
Figura 2.2. Representação da relação entre nível de abstração e quantidade de informação em um <i>workflow</i>	19
Figura 2.3. Ciclo de vida do experimento científico <i>in silico</i> , adaptado de MATTOSO ET AL. (2010).	21
Figura 2.4. Elemento Classe da UML 2.....	28
Figura 2.5. Associação simples.	28
Figura 2.6. Relacionamento de agregação e composição, respectivamente.....	29
Figura 2.7. Relacionamento de especialização e autorelacionamento.	29
Figura 2.8. Extrato do Metamodelo da UML 2.2 contendo elementos da notação do diagrama de atividades. Também estão representados os relacionamentos entre esses elementos.	30
Figura 2.9. Notação gráfica de Atividade (a), Ação (b) e elemento CallBehaviorAction (c).	30
Figura 2.10. Notação de Arestas: Fluxo de controle (a) e Fluxo de dados (b).....	31
Figura 2.11. Notação Pin: Input pin e Output pin (a). Notação de Parameter node (b).	32
Figura 2.12. Elementos de controle e decisão.	32
Figura 2.13. Elementos de início e fim de fluxo.....	33
Figura 3.1. Distribuição anual de referências sobre composição e quantidade de abordagens novas.....	65
Figura 4.1. Abordagem para concepção de <i>workflow</i> científico em nível abstrato na versão inicial.	78
Figura 4.2. Método para concepção de experimentos baseados em <i>workflow</i> científico, em nível abstrato. Representado na notação de diagrama de atividade, UML 2.	80
Figura 4.3. Detalhamento da fase “Especificar <i>workflow</i> científico”, composta por duas tarefas: “Definir modelo inicial de <i>workflow</i> científico” e “Identificar e modelar requisitos do <i>workflow</i> científico”.....	81
Figura 4.4. Detalhamento da fase “Validar <i>workflow</i> científico”, contendo tarefas de inspeção e validação da Especificação de <i>workflow científico abstrato</i>	83
Figura 4.5. Perfil UML - Estereótipos e elementos do diagrama de atividades UML 2.2, versão um.	96

Figura 4.6. Processo para Observação de Evolução de <i>Software</i> (Araújo, 2009).	97
Figura 4.7. Modelo de contexto para o estudo “ <i>Simulação da Evolução de Software</i> ”. 99	
Figura 4.8. <i>Workflow</i> abstrato inicial para o estudo de <i>Simular a Evolução de Software</i>	100
Figura 4.9. Sub- <i>workflow</i> para atividade composta “ <i>Gerar equações para simulação</i> ”.	100
Figura 4.10. Exemplo da planilha para relato de discrepâncias encontradas na inspeção.	103
Figura 4.11. <i>Workflow</i> abstrato corrigido para o estudo de “ <i>Simular a Evolução de Software</i> ”.	105
Figura 5.1. Estrutura da abordagem para composição de <i>workflow</i> científico	108
Figura 5.2. Estereótipos e elementos principais do diagrama de atividades UML 2.2.114	
Figura 5.3. Fase de “Validar <i>workflow</i> científico” do procedimento proposto nesta dissertação.....	115
Figura 5.4. Representação dos procedimentos de “Analisar fadiga de risers” (CORREA, 2003).....	122
Figura 5.5. Modelo de contexto para experimento in silico Análise de fadiga de <i>risers</i>	123
Figura 5.6. Modelos adaptados da especificação. (a) <i>workflow</i> abstrato de “Analisar de fadiga de risers”. (b) modelo da atividade composta “Analisar dos movimentos da plataforma”.....	125
Figura 5.7. (a) Extrato do diagrama de Ferramentas. (b) Extrato do diagrama de Artefatos.	126
Figura 5.8. Distribuição dos defeitos pelas categorias.	127
Figura 5.9. Exemplo de <i>workflows</i> concretos definidos a partir da especificação de <i>workflow</i> abstrato.	129

ÍNDICE DE TABELAS

Tabela 2.1. Etapas do ciclo de vida de MATTOSO ET AL. (2010) mapeadas para o processo experimental de WOHLIN ET AL. (2000) / AMARAL (2003).....	21
Tabela 2.2. Elementos da experimentação baseada em <i>workflow</i> científico e suas descrições.....	25
Tabela 3.1. Objetivo específico relacionado à questão primária.	36
Tabela 3.2. Objetivo específico relacionado à questão secundária do estudo.....	36
Tabela 3.3. Questões para avaliação da qualidade dos trabalhos.	40
Tabela 3.4. Tabela com valores a serem utilizados na Q.4.....	40
Tabela 3.5. Artigos de controle para a revisão <i>quasi</i> sistemática.....	42
Tabela 3.6. Quantidade de referências mantidas sem duplicatas para questão primária.	47
Tabela 3.7. Quantidade de referências base de dados após a primeira fase de avaliação para a questão primária.	47
Tabela 3.8. Quantidade de referências nos anais após a primeira fase de avaliação para a questão primária.	47
Tabela 3.9. Quantidade total de referências incluídas na segunda fase de avaliação para questão primária.	48
Tabela 3.10. Referências recuperadas para a questão secundária.....	49
Tabela 3.11. Quantidade de referências após a primeira fase de avaliação da questão secundária.	50
Tabela 3.12. Quantidade final de referências incluídas e aptas à extração da questão secundária.	50
Tabela 3.13. Quantidade de referências recuperadas e mantidas sem duplicatas para String primária, Outubro 2010.	51
Tabela 3.14. Quantidade de referências incluídas para extração de dados questão primária, Outubro 2010.	51
Tabela 3.15. Quantidade de referências recuperadas e mantidas sem duplicatas para String secundária, Outubro 2010.....	52
Tabela 3.16. Quantidade de referências incluídas para extração de dados questão secundária, Outubro 2010.....	52
Tabela 3.17. Percentual de representatividade das bases digitais no conjunto de referências da questão primária, por etapa de inclusão.	53
Tabela 3.18. Percentual de representatividade das bases digitais no conjunto de referências da questão secundária, por etapa de inclusão.....	53

Tabela 3.19. Informações de referência dos artigos incluídos na questão primária e secundária.	54
Tabela 3.20. Resultado da avaliação da qualidade das referências incluídas.	55
Tabela 3.21. Questões oriundas dos objetivos de pesquisa.	64
Tabela 3.22. Descrição e Categorias dos problemas identificados.	66
Tabela 3.23. Problemas tratados pelas abordagens.	67
Tabela 3.24. Tecnologias de construção utilizadas por cada abordagem.	70
Tabela 3.25. Características das abordagens identificadas.	72
Tabela 3.26. Categoria de procedimentos de verificação listados por cada referência.	74
Tabela 4.1. Classificação das discrepâncias, adaptado de [SHULL, 1998].	84
Tabela 4.2. Formulário de Atividade e seus campos.	91
Tabela 4.3. Formulário de Artefato e seus campos.	92
Tabela 4.4. Formulário de Ferramenta e seus campos.	93
Tabela 4.5. Mapeamento entre elementos do domínio de <i>workflow</i> científico para elementos do diagrama de atividades da UML 2.2.	95
Tabela 4.6. Formulário atividade <i>Criar equações da simulação</i> retirado da especificação.	102
Tabela 4.7. Formulário artefato <i>Dados da simulação da evolução</i> retirado da especificação.	102
Tabela 4.8. Formulário ferramenta <i>Tabela_Excel</i> retirado da especificação.	103
Tabela 5.1. Novos campos e suas descrições do formulário de Artefato.	110
Tabela 5.2. Mapeamento final entre elementos do domínio de <i>workflow</i> científico para elementos do diagrama de atividades da UML 2.2.	112
Tabela 5.3. Heurísticas para inspeção e suas categorias.	119
Tabela 5.4. Dados de cada inspetor relacionados ao número de defeitos e tempo gasto na inspeção.	127

CAPÍTULO 1 - INTRODUÇÃO

Neste capítulo é apresentado o contexto do trabalho relacionado ao uso de simulação na realização da ciência e as vantagens e as suas conseqüências. São descritas a motivação desta pesquisa, abordando os desafios da modelagem computacional e experimentação baseada em simulação, e a questão de investigação que norteia a pesquisa. São também apresentados os seus objetivos, o método de pesquisa adotado, o histórico deste trabalho e a organização deste texto.

1.1 Contexto e Motivação: Descrição do Problema

A Engenharia de *Software* (ES) vem utilizando a Experimentação como instrumento para a criação de um corpo de conhecimento e, para que apresente validade científica, todos seus itens precisam ser verificados perante a realidade através de estudos experimentais (JURISTO & MORENO, 2001). Estes estudos permitem ao pesquisador um maior controle da situação e a manipulação do comportamento do ambiente de forma direta, precisa e sistemática (WOHLIN ET AL., 2000). Na Engenharia de *Software*, os estudos experimentais visam observar a validade dos itens de conhecimento quando relacionados a seus possíveis comportamentos em processos de *software* e como podem afetar o produto gerado. Atualmente, diferentes estudos experimentais podem ser aplicados para realizar esta verificação, podendo contar com a participação de profissionais em ambientes variando entre indústria e academia.

Todavia, existem limitações em situações nas quais o tempo para examinar o comportamento sob investigação é demasiadamente longo, o que pode acabar por inviabilizar a participação de profissionais e tornar a observação mais difícil, resultando riscos de continuidade da pesquisa. Condições como essa têm motivado o uso cada vez mais freqüente de estudos baseados em simulação na Engenharia de *Software* Experimental (ZHANG ET AL., 2008). O uso da simulação como aparato de realização de suas pesquisas, embora seja prática recente em Engenharia de *Software*, é corrente em outras áreas da ciência, e.g., Biologia, Engenharia, Física, dentre outras (MATTOS ET AL., 2008). Engenheiros de *software* têm executado seus estudos baseados em simulação ao longo dos anos para observar comportamentos relacionados à gerência de risco (LIU ET AL., 2009), processos de *software* (ZHANG

ET AL., 2008) e decaimento de *software* (RUS ET AL., 2003; ARAÚJO, 2009), dentre outros.

Os estudos que fazem uso de ambientes simulados são denominados experimentos *in virtuo* e *in silico*. Esses estudos permitem observar o mundo real através de simulação em ambientes virtuais, compostos por modelos computacionais que estão representados, em geral, por programas de computador (e.g. sistema ou *software*). A simulação ocorre através da execução desses programas em infraestruturas computacionais. Em especial, os experimentos *in silico* são altamente dependentes de recursos computacionais (DEELMAN ET AL., 2009). Os estudos baseados em simulação nas outras áreas de ciência são denominados também de experimentos científicos (TROGER & FERNANDES, 2004).

A principal diferença entre os estudos baseados em simulação é relacionada à presença de participantes reais ou simulados. Nos experimentos *in silico* tanto os participantes quanto o ambiente são simulados, ao contrário nos experimentos *in virtuo*, onde o ambiente simulado sofre interação de participantes reais. Um estudo baseado em simulação, apesar das questões envolvendo a qualidade dos modelos computacionais e possíveis problemas éticos (BASILI ET AL., 2008), pode trazer vantagens quando comparado a estudos executados em laboratório ou indústria, como permitir maior controle do ambiente, menor custo de execução e a possibilidade de antecipação a futuros riscos à pesquisa. Também, existe a possibilidade de observar, de forma restrita, a viabilidade das tecnologias de *software* sob investigação (TRAVASSOS & BARROS, 2003).

Uma opção para apoiar os estudos *in virtuo* e *in silico* seria através do uso de abordagens para definição de *workflows* de negócio (HOLLINGSWORTH, 1994), visando automatizar esses experimentos. O *workflow* de negócio surgiu em meados da década de 1970, com o intuito de automatizar e coordenar fluxos de trabalho em escritórios (GIL, 2007). Atualmente, existem diversas abordagens para apoiar a definição de *workflow* de negócio, tais como *Object Thinking by Proforma*, citada em (IENDRIKE, 2008) e a modelagem através de casos de uso (JACOBSON ET. AL., 1994). Contudo, deve-se considerar que as características dos ambientes de pesquisa diferem dos escritórios de negócio (BASILI ET. AL., 2008), pois existem:

- Limitações quanto à disponibilidade de recursos humanos para construção dos *workflows*. Em geral, as equipes são compostas por poucos indivíduos e com carga de trabalho alta e pouco tempo disponível;
- Limitações quanto ao nível de conhecimento em tecnologias para identificação de requisitos e desenvolvimento de sistemas e *workflows*. Em

geral, as equipes são compostas por pesquisadores não especialistas em ciência da computação, e;

- Variações no conjunto de atividades que pode conter diversas alternativas de métodos, ferramentas e sistemas. Essas variações geram muitas combinações de fluxos de execução que possuem o mesmo objetivo.

Assim, a utilização de abordagens para definição de *workflows* de negócio na área de experimentação baseada em simulação pode não ser trivial, pois os contextos são muito diferentes em termos de participantes e recursos disponíveis. Desta forma, para apoiar os estudos baseados em simulação, foram criadas tecnologias como *workflow* científico e Sistemas Gerenciadores de *Workflow* Científico (SGWfC). O *workflow científico* é um modelo que representa as atividades do estudo organizadas em uma seqüência a fim de realizar um determinado objetivo e, em geral, estas são realizadas automaticamente por programas (DEELMAN ET AL., 2009). Os *workflows* são interpretados e executados pelos SGWfCs que, em geral, permitem sua especificação, modelagem e execução. Os benefícios do uso destas tecnologias para experimentação são: registro da proveniência dos dados; automação da execução do fluxo de atividades; controle e invocação das ferramentas; manipulação dos dados consumidos e gerados (MATTOS ET AL., 2008).

Entretanto mesmo com essas novas tecnologias, os experimentos *in virtuo* e *in silico*, naturalmente, adicionam complexidade a realização de pesquisas, pois esses requerem maior apoio computacional para a execução dos seus diferentes programas e uma infraestrutura complexa disponível e configurada *a priori*. Também existe necessidade de maior conhecimento do domínio pelos pesquisadores onde a pesquisa será executada para que seja realizada a modelagem computacional do ambiente, do participante do estudo e do objeto sob investigação (TRAVASSOS & BARROS, 2003). Isso tudo torna a *concepção* e *execução* desses estudos não trivial para o pesquisador, o que pode representar uma dificuldade para a modelagem computacional de estudos baseados em simulação.

De fato, a modelagem computacional já foi identificada como um dos desafios para computação por diversos grupos de pesquisa: UKCRC (UKCRC, 2005; KAVANAGH & HALL, 2008); CRA (CRA, 2005); SBC (SBC, 2006). Em especial, a SBC (2006) detalha os principais desafios para 2006 - 2016 relacionados à modelagem computacional como:

- Desempenho: desafio relacionado ao processamento dos dados, ao uso de infraestruturas e a ambientes para alto desempenho.

- Identificação de requisitos: desafio relacionado ao desenvolvimento de novas técnicas de identificações de requisitos.
- Armazenamento: desafio relacionado ao armazenamento do alto volume de dados gerados e dos fatores associados à sua manipulação.
- Colaboração: desafio relacionado ao impacto da colaboração entre engenheiros de *software* e outros cientistas e ao uso de ferramentas que auxiliem a modelagem colaborativa.
- Visualização: desafio relacionado ao desenvolvimento de algoritmos e técnicas de visualização dos resultados gerados.

Como se percebe, a modelagem computacional é uma área de pesquisa em aberto com muitos desafios a serem vencidos. Assim, à medida que experimentos baseados em simulação, em especial da categoria *in silico*, são cada vez mais utilizados na Engenharia de *Software* e outras áreas da ciência, torna-se cada vez mais necessário que sejam desenvolvidas tecnologias para auxiliar em sua concepção e execução.

1.2 Questões de Pesquisa

Uma preocupação para a realização de experimentos *in silico* se refere à utilização dos SGWfCs. Esses sistemas geralmente usam linguagens específicas, obrigando os pesquisadores a descrever o *workflow* científico em baixo nível de abstração exclusivamente para este sistema. Por conseguinte, os pesquisadores se concentram mais nas questões de implementação e menos na definição de requisitos essenciais do experimento *in silico*, o que torna toda a tarefa de concepção mais complexa. Assim, a concepção do estudo se torna semelhante à atividade de "codificação", porque a maioria das questões diz respeito às definições de recursos computacionais, como por exemplo, escrita de linhas de comando e escolha de endereços *IP*.

Mesmo o emprego de esforços no provimento de mecanismos para modelagem visual, *e.g.*, Kepler (ALTINTAS ET AL., 2004) ou Vistrails (CALLAHAN ET AL., 2006), não foram suficientes para resolver o problema descrito anteriormente. Pois, ainda existe a codificação dos programas e definição de detalhes arquiteturais nesses SGWfC juntamente com a concepção do estudo. Adicionalmente, observa-se a falta de métodos para auxiliar os cientistas na concepção e identificação de requisitos para experimentos *in silico* (SBC, 2006; BASILI ET AL. 2008). Portanto, a concepção é normalmente realizada *ad hoc* e diretamente no SGWfC (baixo nível de abstração).

Assim, não há garantia de que as informações relevantes para apoiar o estudo experimental foram retratadas, exceto as decisões arquiteturais de baixo nível oriundas das necessidades computacionais do experimento científico.

O ideal seria inicialmente conceber o *workflow* científico em alto nível de abstração (independente de Sistema Gerenciador de *Workflow* Científico) e posteriormente definir detalhes de *implementação* relacionados a aspectos arquiteturais e restritivos do seu funcionamento. Na verdade, a concepção através de diferentes níveis de abstração é citada por diversos autores (LUDASCHER ET AL., 2003; VERDI ET AL., 2007; OGASAWARA ET AL., 2009) como boa prática a ser empregada na concepção, diminuindo a complexidade de tal tarefa. Entretanto, considerando o conhecimento do autor desta dissertação sobre o tema, não existe orientação disponível na literatura técnica sobre o tipo de informação necessária (requisitos) e tão pouco sobre quais são as tarefas e perfis para a concepção desses experimentos *in silico* (e *in virtuo*) em alto nível de abstração: o que apenas se encontra é algum auxílio na definição de restrições e condições relativas aos diferentes SGWfC.

Nós acreditamos que os experimentos *in silico* (e *in virtuo*) representam uma estratégia promissora e necessária para acelerar a experimentação e ciclos de aprendizagem em Engenharia de *Software*. No entanto, na revisão da literatura técnica não foi possível encontrar tecnologias maduras para lidar com esses experimentos, desde a concepção de *workflow* em níveis mais abstratos, até sua implementação em um SGWfC.

Apesar de relacionados, os *workflows* de negócio e científico diferem em uma característica: quem desenha o *workflow*. No contexto do negócio, existem especialistas em especificação disponíveis, e notações e ferramentas são criadas para utilização por Engenheiros de *Software* ou cientistas da computação. Entretanto, no contexto científico não há tanta disponibilidade de recursos com alto conhecimento em desenvolvimento e especificação de requisitos, pois, em geral, as equipes são pequenas e formadas por especialistas do domínio (BASILI ET AL., 2008). Todo este contexto nos motivou a investigação no sentido de apoiar as abordagens relativas à concepção de *workflow* científico em experimentos *in virtuo* e *in silico* na Engenharia de *Software* e áreas afim.

Assim, apesar da concepção de *workflow* como um todo ser um campo de pesquisa em aberto, o problema a ser tratado pela abordagem descrita nesta dissertação está relacionado à concepção de *workflow* científico em alto nível de abstração e sua representação em um formato estruturado e definido. Além disso, esta abordagem foi concebida para ser utilizada por pesquisadores e especialistas do

domínio, que pretendem utilizar *workflow científico* nas suas simulações e que não são, necessariamente, Engenheiros de *Software* ou cientistas da computação. Desta forma, a abordagem é composta por poucos instrumentos, artefatos e tarefas, considerando a restrição de recursos humanos e de nível de conhecimento sobre computação. Além disso, esta abordagem não contempla a concepção de *workflow científico* para um SGWfC em particular, isto é, não se preocupa com a implementação do *workflow* nestes sistemas e na definição de aspectos arquiteturais e recursos computacionais.

Considerando este cenário, esta pesquisa propõe uma abordagem de apoio à identificação de *workflows científicos* para experimentos *in silico* (e *in virtuo*) em alto nível de abstração, aplicando também técnicas de garantia da qualidade nos modelos de *workflow* definidos. Deve-se esclarecer que o escopo da pesquisa não diz respeito ao desenvolvimento de sistemas ou *softwares* que representam os modelos computacionais a serem executados, mas sim apoiar na representação da ordem de execução das atividades do experimento *in silico* como modelos de *workflow científico*.

A hipótese desta pesquisa considera que *é possível minimizar os riscos envolvidos com a concepção de experimentos in silico (e in virtuo) que utilizam tecnologia de workflow científico, reduzindo problemas de omissão, fato incorreto, ambigüidade, informação estranha e inconsistência nos seus requisitos identificados.*

1.3 Objetivos

O principal objetivo desta pesquisa consiste na definição de abordagem que auxilie o engenheiro de *software* (ou pesquisadores) nas tarefas de concepção e verificação dos requisitos de experimentos *in silico* (e *in virtuo*) que utilizam a tecnologia de *workflow científico*. Esse objetivo pode ser decomposto em:

1. **Identificar o cenário sobre composição de *workflow científico*:** identificar através do uso de *quasi-revisões* sistemáticas o cenário sobre composição de *workflow científico* para experimentos *in silico*;
2. **Identificar características que definem elementos do *workflow científico*:** identificar, através da revisão da literatura, as características que definem os elementos do *workflow científico*;
3. **Definir o procedimento para concepção de *workflow científico*:** estabelecer as atividades de apoio à identificação e verificação de requisitos do *workflow científico*, assim como seus artefatos consumidos e produzidos;

4. **Definir modelos para representação gráfica do *workflow* científico:** definir a notação a ser utilizada na representação do modelo de *workflow* científico, estendendo-a para representar os seus elementos;
5. **Definir representação textual para especificar os requisitos do *workflow* científico:** estruturar, a partir das características definidas, uma forma de capturar textualmente as informações sobre os requisitos do *workflow* científico.

A fim de limitar o escopo, esta pesquisa não tem como objetivo em nenhum momento definir os seguintes itens:

- Linguagem universal de representação e/ou execução de *workflow* científico;
- Novo sistema gerenciador de *workflow* científico e/ou máquinas de execução de *workflow* científico;
- Novos paradigmas de simulação de modelos computacionais e experimentos *in silico*;
- Criação de técnicas ou métodos para desenvolvimento de *software* que representem modelos computacionais.

1.4 Trabalhos relacionados

Algumas boas práticas relacionadas ao uso de *workflow* científico (LUDASCHER ET AL., 2003; YU & BUYYA, 2005; DEELMAN ET AL., 2009; VERDI ET AL., 2007) sugerem que se deve explorar diferentes níveis de abstração quando construindo seus modelos, utilizando inicialmente abstrações de conceitos em mais alto nível e postergando a definição de detalhes de implementação. Apesar das recomendações, poucos trabalhos são encontrados propondo métodos ou processos ou procedimento que auxiliem na concepção dos *workflows* para experimentos *in silico*.

Um destes trabalhos é VERDI ET AL. (2007), que definem um processo para capturar *workflow* conceitual, inspirado nos conceitos industriais para fluxo automatizado de negócios. O processo explora três diferentes modelos: modelo representando o fluxo de controle; modelo hierárquico dos *workflows*, que representa a ordem e os níveis de atividades; e um modelo que representa o fluxo de dados para cada atividade. No entanto, representar os fluxos do experimento através desses modelos pode exigir um maior esforço na manutenção, pois são três tipos de modelos

diferentes, podendo levar a inconsistências entre eles. Além disso, não há nenhuma indicação de que o processo descrito por VERDI ET AL. oferece um mecanismo de verificação, exceto por propor a validação do *workflow* científico através de *walkthroughs*, realizado pelos mesmos pesquisadores que criaram o documento, o que pode influenciar os resultados.

Em outro trabalho, GIL ET AL. (2007b) propõem um processo para compor os *workflows* em diferentes níveis de abstração. Os autores definem as etapas do processo e os papéis a serem desempenhados em cada etapa. O processo é mais abrangente do que a abordagem proposta nesta dissertação, porque lida com a concepção dos *workflows* em alto nível de abstração e na sua posterior implementação. No entanto, a abordagem de GIL ET AL. não traz maiores detalhes, pois não descreve as informações necessárias para nova concepção e nem apresenta como devem ser os instrumentos utilizados. Além disso, a abordagem representa o experimento como um *workflow* em alto nível, todavia toda a abstração é relacionada apenas a único fluxo de execução, sem apresentar possíveis variações em atividades ou caminhos de execução.

O trabalho descrito por PLLANA ET AL. (2005) se relaciona à abordagem proposta nesta dissertação devido ao uso de modelos UML para representar *workflows* científicos. PLLANA ET AL. propõem um sistema gerenciador no qual os seus modelos são expressos como diagrama de atividades. Para tal, eles realizam uma extensão da notação para representar as funcionalidades do sistema, como por exemplo, paralelismo. Apesar de PLLANA ET AL. utilizarem o diagrama de atividades, eles diferem no nível de abstração que realizam a concepção, pois na abordagem deles, o *workflow* já define detalhes de implementação na interface do sistema gerenciador.

Considerando a perspectiva de reutilização das informações do experimento científico, existem iniciativas como myExperiment (DE ROURE ET AL., 2009) e Vistrails (CALLAHAN ET AL., 2006) que permitem o armazenamento e posterior uso de *workflows* científicos por outros pesquisadores ou grupos de pesquisa. Contudo, os *workflows* científicos armazenados se encontram representados em nível concreto, já definidos para um SGWfC em particular e ligados a uma infra-estrutura computacional específica, podendo necessitar de adaptações e acarretar, assim, problemas no seu uso. Adicionalmente, além dos modelos de *workflow* científico, não existe garantia de existência de uma documentação auxiliar contendo informações sobre o *workflow* escolhido para ser reutilizado.

Considerando a concepção de *workflow* científico, OGASAWARA ET AL. (2009) propõe uma abordagem denominada GexpLine, que permite, a partir de uma

representação estruturada, a derivação semi-automática de *workflows* científicos concretos para SGWfCs. Entretanto, a representação feita na GexpLine não se encontra no mesmo nível de abstração que a abordagem descrita nesta dissertação, principalmente pelo fato dos modelos da GexpLine conterem informações de infraestrutura computacional previamente definidas e necessárias para a derivação dos *workflows* concretos. De fato, as abordagens GexpLine e a descrita nesta dissertação são complementares, pois uma lida diretamente com a concepção de *workflow* em nível abstrato, enquanto a outra pode utilizar essas informações como insumo para a concepção dos *workflows* concretos.

1.5 Método de Pesquisa

O método de pesquisa utilizado neste trabalho segue os passos descritos na Figura 1.1, e ao seu final, espera-se como resultado a definição da abordagem para concepção proposta. A seguir, tem-se uma descrição de cada um dos passos executados:

- **Realizar Revisão Inicial da Literatura:** executado no segundo semestre de 2008, nesta atividade realizou-se uma revisão *ad hoc* da literatura técnica sobre composição de *workflow* científico para experimentos *in silico* (e *in virtuo*). O foco desta revisão inicial foi entender os conceitos básicos desta área de pesquisa, os principais termos e identificar os elementos do domínio de *workflow* científico e suas características;
- **Executar Quasi-Revisão Sistemática da literatura:** consistiu na realização de revisão controlada da literatura. Os fundamentos da revisão sistemática (estudo secundário) auxiliam na obtenção de resultados reaplicáveis, verificáveis e possivelmente mais precisos sobre o assunto que se pretende pesquisar (BIOLCHINI ET AL., 2005). Iniciada no segundo semestre de 2008 e executada no primeiro semestre de 2009, o foco desta revisão foi investigar o estado da arte e da prática da composição de *workflow* científico caracterizando as abordagens;
- **Definir a abordagem para composição:** iniciado no segundo semestre de 2008 e continuado no primeiro semestre de 2009, foi desenvolvido de forma incremental e utilizou como base as informações identificadas na revisão *ad hoc* e na *quasi* revisão sistemática. Ao final, chegou-se à abordagem para apoiar a definição e verificação de requisitos de experimentos *in silico* representados por *workflow* científico;

- **Realizar prova de conceito:** realizado no segundo semestre de 2009. O objetivo era identificar problemas na abordagem proposta, através da sua aplicação em um experimento *in virtuo*, e assim obter indícios da sua viabilidade.
- **Evoluir abordagem para composição:** realizado no primeiro semestre de 2010. Neste os problemas identificados na abordagem, após sua aplicação na prova de conceito, foram resolvidos e uma nova versão da abordagem foi criada.
- **Realizar estudo de campo:** iniciado no primeiro semestre de 2010 e concluído no segundo semestre de 2010. O objetivo foi avaliar a aplicabilidade da abordagem proposta em um ambiente real de pesquisa, desenvolvendo uma documentação para um experimento *in silico* e, assim, obter indícios da sua viabilidade.
- **Atualizar os resultados da Quasi-Revisão Sistemática:** realizado no segundo semestre de 2010, consistindo na reaplicação do protocolo de pesquisa da *quasi* revisão sistemática com o intuito de atualizar os dados e obter caracterização do cenário mais atual sobre composição de *workflow* científico.

1.6 Organização do Trabalho

Este capítulo apresentou as idéias motivadoras para o desenvolvimento dessa dissertação, o método de pesquisa utilizado para definição da proposta e a hipótese direcionadora da pesquisa. Estes tópicos serão refinados ao longo dos próximos capítulos. O texto dessa dissertação é estruturado da seguinte forma, além desse

Capítulo 1 de introdução:

- **Capítulo 2** – Os conceitos sobre experimentação são apresentados. Também é descrita a tecnologia de *workflow* científico utilizada na realização de simulação. São revisadas algumas notações de modelos e linguagens de representação de *workflows* existentes na literatura técnica.
- **Capítulo 3** – A *Quasi* revisão sistemática realizada é apresentada, descrevendo a questão de pesquisa e o seu protocolo. Nesse capítulo também é apresentada a análise dos resultados obtidos e as conclusões geradas a partir desses dados.
- **Capítulo 4** – A estrutura da abordagem para composição de *workflow*, em sua versão inicial, é apresentada. Neste capítulo descreve-se o

procedimento para a concepção de *workflow* científico abstrato. Também é descrita a prova de conceito aplicada no domínio de evolução de *software*, com intuito de identificar problemas e avaliar a viabilidade da abordagem;

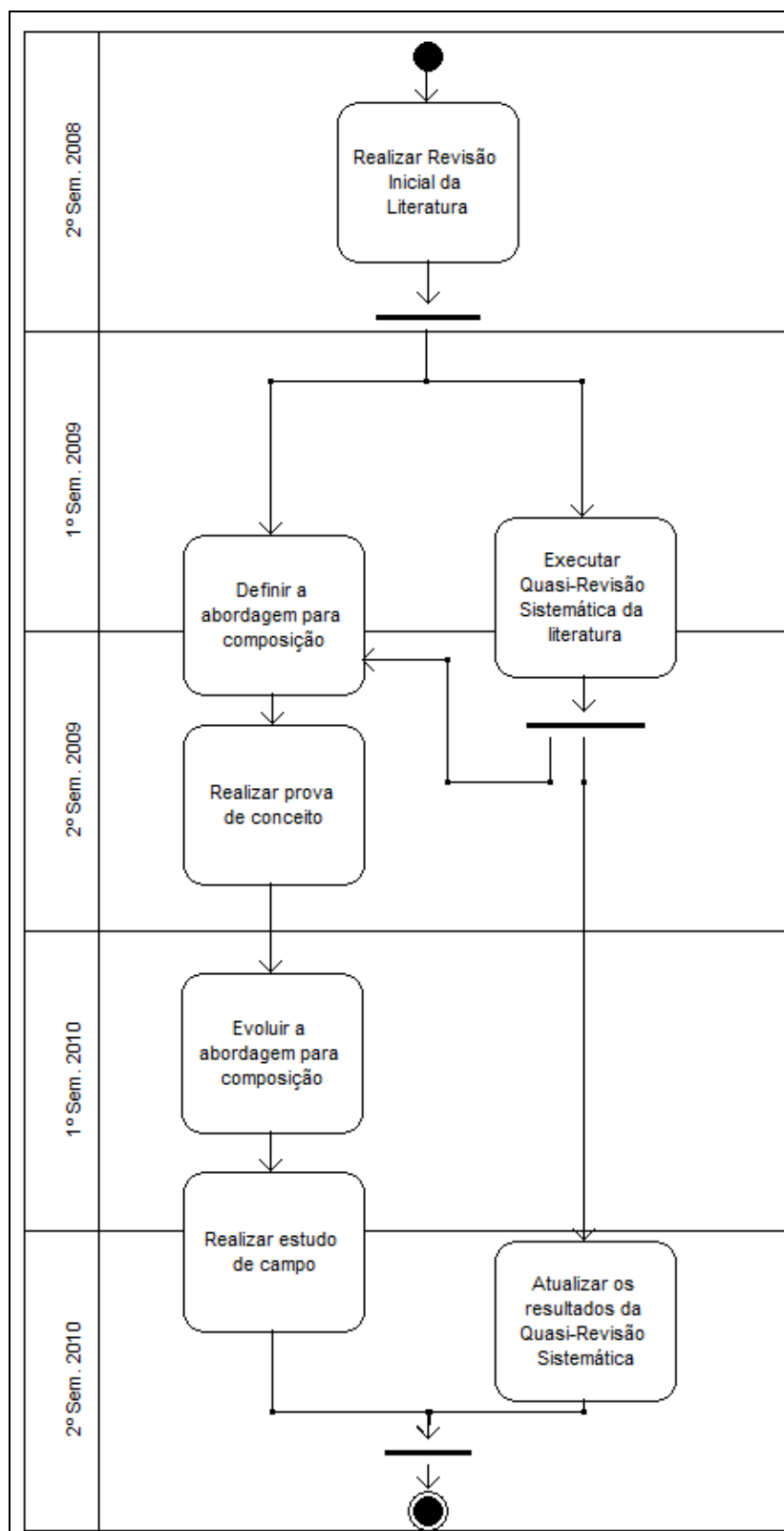


Figura 1.1. Passos do método de pesquisa aplicado no desenvolvimento da abordagem.

- **Capítulo 5** – A evolução da estrutura da abordagem para composição de *workflow* é descrita neste capítulo. São apresentadas as modificações realizadas no procedimento, nos instrumentos e a definição de heurísticas para o uso na abordagem. Adicionalmente, é descrita a avaliação experimental da abordagem, que consiste na sua aplicação de estudo de campo em um projeto real em larga escala.
- **Capítulo 6** – As conclusões da pesquisa, a descrição dos resultados obtidos, as principais contribuições e as limitações desta pesquisa são descritos neste capítulo.

Além destes, esta dissertação contém cinco anexos:

- **Anexo A** – Anexo contendo as strings de buscas derivadas para cada base digital, que foram utilizadas no protocolo de revisão *quasi* sistemática.
- **Anexo B** – Anexo contendo a lista de todas as referências retornadas após a execução da string de busca e sua classificação quanto à inclusão e exclusão.
- **Anexo C** – Anexo contendo a tabela completa da síntese dos dados extraídos das referências incluídas, incluindo o trecho extraído da própria referência.
- **Anexo D** – Anexo contendo um template do documento que descreve a especificação do *workflow* científico abstrato.
- **Anexo E** – Anexo contendo o metamodelo auxiliar dos elementos do domínio de *workflow* científico.

CAPÍTULO 2 - CONCEITOS SOBRE EXPERIMENTAÇÃO BASEADA EM SIMULAÇÃO, WORKFLOW E UML

Neste capítulo é apresentada a revisão dos conceitos básicos sobre a experimentação e a tecnologia de workflow científico necessários para apoiar o entendimento do tema dessa dissertação. Esta revisão ad hoc da literatura técnica aborda o conceito de experimentação, de workflow científico e seus ciclos de vida, comparando-os brevemente. Também são apresentados os requisitos dos Sistemas Gerenciadores, que visam apoiar o ciclo de vida do workflow científico. Por último é apresentado um conjunto de linguagens e notações para representação gráfica de workflow científico, em especial, a notação de diagrama de atividades da UML 2.0.

2.1 Fundamentação teórica

2.1.1 Experimentação

O uso de experimentação é motivado quando o pesquisador deseja ter o controle sobre o ambiente e manipular o comportamento do objeto sob avaliação de maneira precisa, direta e sistemática, sendo seu principal objetivo avaliar uma hipótese previamente definida. Contudo, realizar experimentação não é uma tarefa trivial, necessitando de preparação adequada do protocolo do estudo experimental, o gerenciamento da sua execução e sua posterior análise (WOHLIN ET. AL. 2000).

Estudos experimentais se diferenciam de acordo com o nível de controle que o pesquisador possui sobre o ambiente onde são executados, permitindo dividi-los em categorias distintas de estudos. A categorização ajuda o pesquisador a explicitar características do experimento a ser executado, como o nível de controle, influência de fatores externos e a possibilidade de generalizar o seu resultado (TRAVASSOS & BARROS, 2003).

Travassos & Barros (2003) apresentam uma taxonomia para estudos experimentais em Engenharia de *Software* composta por quatro categorias, dividindo assim os experimentos em:

- *In vivo*: experimentos aplicados por participantes em ambientes e em circunstâncias reais (e.g., um projeto de desenvolvimento de *software* em uma empresa).
- *In vitro*: experimentos executados em ambientes controlados, onde o pesquisador possui maior controle sobre os fatores que podem afetar os resultados. Em geral, os estudos são feitos em laboratórios ou universidades.
- *In virtuo*: experimentos onde o ambiente é representado através de modelos computacionais, que descrevem a realidade na qual os participantes interagem diretamente.
- *In silico*: experimentos onde tanto os participantes quanto o ambiente e objeto de estudo são descritos como modelos computacionais, não havendo (ou reduzindo ao máximo) qualquer tipo de intervenção humana.

Uma característica importante de experimentos *in virtuo* e *in silico* é a possibilidade de simulação do mundo real, conduzida em um ambiente virtual composto por modelos numéricos (TRAVASSOS & BARROS, 2003). Estes modelos são representações computacionais de um fenômeno ou elemento do mundo real. Em geral, os modelos computacionais estão representados na forma de *software*, sistemas ou fontes de informações. Em diversas outras áreas a simulação é utilizada para auxiliar a tomada de decisão ou avaliação de um determinado fenômeno. Experimentos *in silico* são encontrados, dentre outras áreas, na bioinformática (STEVENS ET AL., 2004), meteorologia (GANNON ET AL. 2007), ecologia (PENNINGTON ET AL. 2007) e exploração de petróleo *offshore* (OLIVEIRA ET AL. 2009). Nestes domínios, os experimentos *in silico* são mais conhecidos como experimentos científicos.

A exploração da simulação confere aos experimentos *in virtuo* e *in silico* vantagens quanto às ameaças ao controle e custos quando comparados aos experimentos *in vivo* ou *in vitro* (TRAVASSOS & BARROS, 2003). Por exemplo, utilizando um experimento *in silico* pode-se representar, através de modelos computacionais, um grande número de participantes interagindo com o ambiente e o objeto sob avaliação. Desta forma, a simulação dessa interação permite extrair indícios de que o comportamento sob estudo é válido. O teste e a avaliação do estudo antes de sua aplicação no mundo real (*in vivo*) ou ambiente controlado (*in vitro*) podem detectar possíveis problemas, evitando custos de reaplicação, principalmente em larga

escala; ou a perda do estudo que devido a pouca disponibilidade de participantes tornaria proibitivo sua reexecução com a mesma população no caso de falha (BARROS, 2001).

Outro exemplo da exploração de experimentos *in silico* é a possível utilização de modelos computacionais para simular o decaimento da qualidade de sistemas de software ao longo do tempo. Notoriamente, conforme sofre evoluções e manutenções ao longo de sua vida útil, o software pode apresentar perda de desempenho, anomalias na execução, dentre outros comportamentos não previstos, ou seja, o decaimento de sua qualidade. Desta maneira, podem-se observar esses possíveis comportamentos, utilizando-se simulação para realizar a passagem de tempo sobre os ciclos de manutenção e com isso obter indícios de como o software se comportará futuramente, permitindo a ação preventiva para evitá-los (ARAÚJO, 2009).

Apesar das vantagens, o uso de experimentos baseados em simulação introduz novas questões como, por exemplo, exigir mais conhecimento do pesquisador sobre o domínio do estudo, pois ele precisa representar o ambiente, participante e/ou objeto sobre avaliação como modelo computacional. Outro ponto em questão é relacionado à qualidade dos modelos computacionais criados para a simulação, que influenciam diretamente na qualidade dos dados produzidos, o que, portanto, podem comprometer o resultado do experimento, direcionando-o para resultados inválidos ou até mesmo conclusões errôneas. Adicionalmente a todas essas questões, deve-se definir uma infraestrutura computacional mais complexa para o estudo, afinal os modelos são, em geral, representados por diferentes softwares que precisam ser executados em recursos computacionais específicos (e.g. servidores, cluster, dentre outros) definidos *a priori* (TRAVASSOS & BARROS, 2003).

Contudo, antes de definir a infraestrutura computacional na qual a simulação será realizada, assim como nos estudos *in vivo* e *in vitro*, o pesquisador precisa tratar outros pontos a fim de criar seu experimento, como a definição do objetivo, das hipóteses, dos instrumentos, dentre outras etapas. Também é foco de preocupação e preparação o empacotamento de todo material utilizado no experimento, pois assim permite-se que este possa ser repetido e avaliado por outros pesquisadores interessados. O empacotamento auxilia na proveniência (FREIRE ET AL. 2008) do experimento, isto é, o registro das informações sobre determinado objeto ou processo utilizado, permitindo-se avaliar posteriormente sua qualidade e autenticidade.

Considerando todas essas necessidades, WOHLIN ET AL. (2000) definem um processo para experimentação, a princípio genérico para qualquer tipo de estudo experimental. O processo definido por WOHLIN ET AL. contém cinco atividades, sendo elas: definição; planejamento; execução; análise e interpretação; e,

apresentação e empacotamento. Apesar deste processo auxiliar no gerenciamento dos estudos experimentais, AMARAL (2003) estendeu o processo modificando a etapa de empacotamento, tornando-a paralela às demais etapas. Assim, a coleta e armazenamento da proveniência abrangem todas as outras atividades, o que segundo AMARAL, diminuiria a possibilidade de omissões de informação importantes na repetição do estudo. A Figura 2.1 representa o processo experimental proposto.

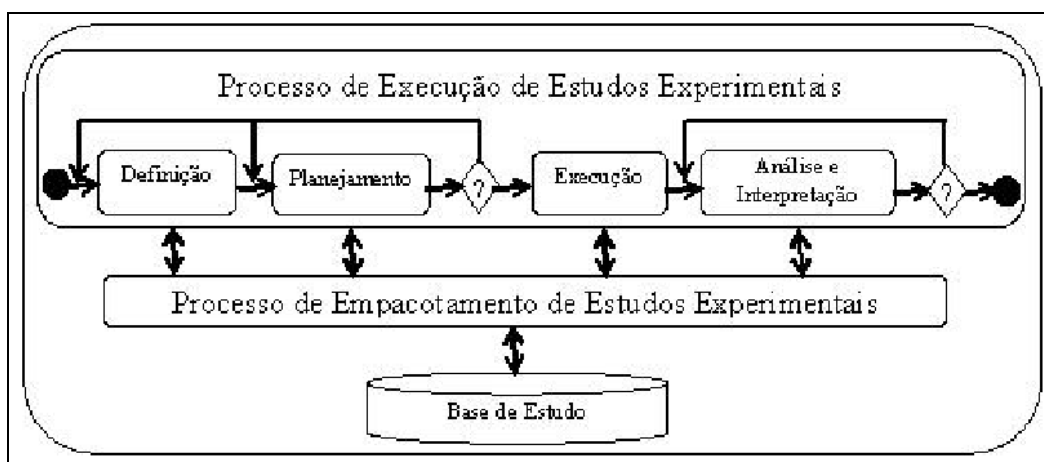


Figura 2.1. Processo experimental, definido por WOHLIN ET. AL. (2000), estendido por AMARAL (2003).

Basicamente, na etapa de definição o pesquisador expressa em termos de objetivos o experimento a ser realizado. Após isto, o pesquisador realiza o seu planejamento, definindo seu projeto do experimento, criando e estabelecendo os instrumentos, analisando os aspectos de validade dos resultados. Todas estas definições são documentadas no Plano de Estudo Experimental. Importante destacar que este plano deve tratar questões que busquem diminuir os riscos na execução do experimento, para que este não venha a ser invalidado. Por isso, ao final do planejamento deve ocorrer uma avaliação do plano, podendo decidir replanejar o experimento ou dar continuidade a ele. Na etapa de execução, o pesquisador coloca em prática o que foi planejado para o estudo, coletando os dados gerados para a posterior interpretação e análise. Na etapa seguinte, o pesquisador analisa os resultados, podendo fazer uso de métodos estatísticos, sínteses, assim interpretando os dados e apresentando suas conclusões.

2.1.2 Workflow científico

Na ciência, o uso da computação surge como um elemento importante em conjunto com a teoria e experimentação (DEELMAN ET AL. 2009). Então, surge o conceito de e-Science, cuja definição é: “Apoio ao cientista para o desenvolvimento de

ciência em larga escala utilizando infraestrutura computacional correspondente” (MATTOSO & DA CRUZ, 2008). Este apoio vem através do uso de tecnologias como o *workflow*, ou neste contexto, chamado **workflow científico**. Muitas áreas já utilizam esta tecnologia, tais como Biologia, Química, Engenharia e Física (MATTOS ET. AL., 2008). O seu uso tem por meta principal facilitar o trabalho do pesquisador, de forma que este possa se concentrar no resultado do experimento científico e não em como será executada a parte computacional da qual ele depende.

Uma característica do *workflow científico* é o conceito de orquestração, que determina a ordem e a seqüência das atividades necessárias para alcançar o objetivo do experimento científico. O *workflow* é o *template* da orquestração e, também, é a representação do experimento científico. No momento em que são inseridos os dados e esse é designado para um determinado problema, torna-se uma instância do *workflow científico* (DEELMAN ET AL. 2009).

Uma característica importante está relacionada à possibilidade de representar o *workflow científico* em níveis de abstração (LUDASCHER ET AL., 2003; YU & BUYYA, 2005; DEELMAN ET AL., 2009; OGASAWARA ET AL., 2009). O nível abstrato permite ao pesquisador um maior grau de liberdade, pois este não precisa se preocupar com questões relacionadas à *implementação*, isto é, os recursos que serão utilizados, por exemplo, ambientes em grade (grid) (FOSTER E KESSELMAN, 1999) ou em *cluster* (BAKER, BUYYA E HYDE, 1999), mas, sim, com o comportamento do experimento e como será expresso através do *workflow científico*. Enquanto isso, o nível concreto está ligado fortemente à tecnologia, porque aqui são associados quais são os recursos computacionais necessários para suportar a execução deste *workflow*.

Contudo, na literatura técnica, são encontradas diversas divisões e classificações sobre o conceito de abstração em *workflows científicos*. LUDASCHER ET AL. (2003) define dois níveis de abstração, que ele nomeia em abstrato e executável. Para eles, o *workflow* abstrato seria uma rede de atividades abstratas, contendo as definições conceituais de seus insumos e produtos, enquanto o *workflow* executável seria uma rede de atividades concretas, isto é, serviços disponíveis do ambiente. O relacionamento entre eles se daria através de um mapeamento semântico entre os termos utilizados no *workflow* abstrato e os recursos do *workflow* concreto. Em YU & BUYYA (2005), o modelo de *workflow* pode ser classificado em dois níveis: abstrato, que não referencia nenhum recurso do ambiente em grade, e concreto, que liga as suas atividades aos recursos. Outra classificação, proposta por GIL (2007), considera que são três os níveis de abstração, sendo eles:

- i.* **Estrutura do workflow (*Workflow template*):** identifica os tipos de componentes a serem invocados e o fluxo de dados entre eles, sendo bem próximo ao *workflow* em nível abstrato;
- ii.* **Instância do workflow (*workflow instance*):** especifica os dados de entrada necessários para uma análise, não se prendendo a detalhes, como por exemplo, localidade do recurso computacional para execução;
- iii.* **Workflow executável (*executable workflow*):** é criado quando uma instância do *workflow* é ligada aos recursos computacionais necessários para sua execução.

Importante ressaltar que mesmo não havendo uma única classificação sobre os níveis de abstração, a representação de um *workflow* pode apresentar vários níveis intermediários entre o abstrato e concreto. Isto dependerá da quantidade e do tipo de informação que o *workflow* contém e a maneira como o pesquisador deseja lidar com a abstração. A Figura 2.2 apresenta uma representação da relação entre o nível de abstração e a quantidade de informação representada em um modelo de *workflow* científico. Ao se relacionar a Figura 2.2 e a classificação proposta por GIL (2007), tem-se que o nível “Estrutura do workflow” está mais próximo do nível de representação, o “Workflow abstrato”; já o nível “Instância do workflow” estaria compreendido nos níveis intermediários de abstração, pois nele já se encontram detalhes de infraestrutura computacional, localização dos dados, definição de ferramentas e programas de computador, dentre outras informações; por último, o nível “Workflow executável” está alocado na parte mais baixa da Figura 2.2, no que seria o “Workflow concreto”, já pronto para execução em uma “engine” com os dados reais a serem consumidos na execução. Destaca-se que a Figura 2.2 não é uma classificação definitiva, mas sim uma genérica que será adotada por este trabalho.

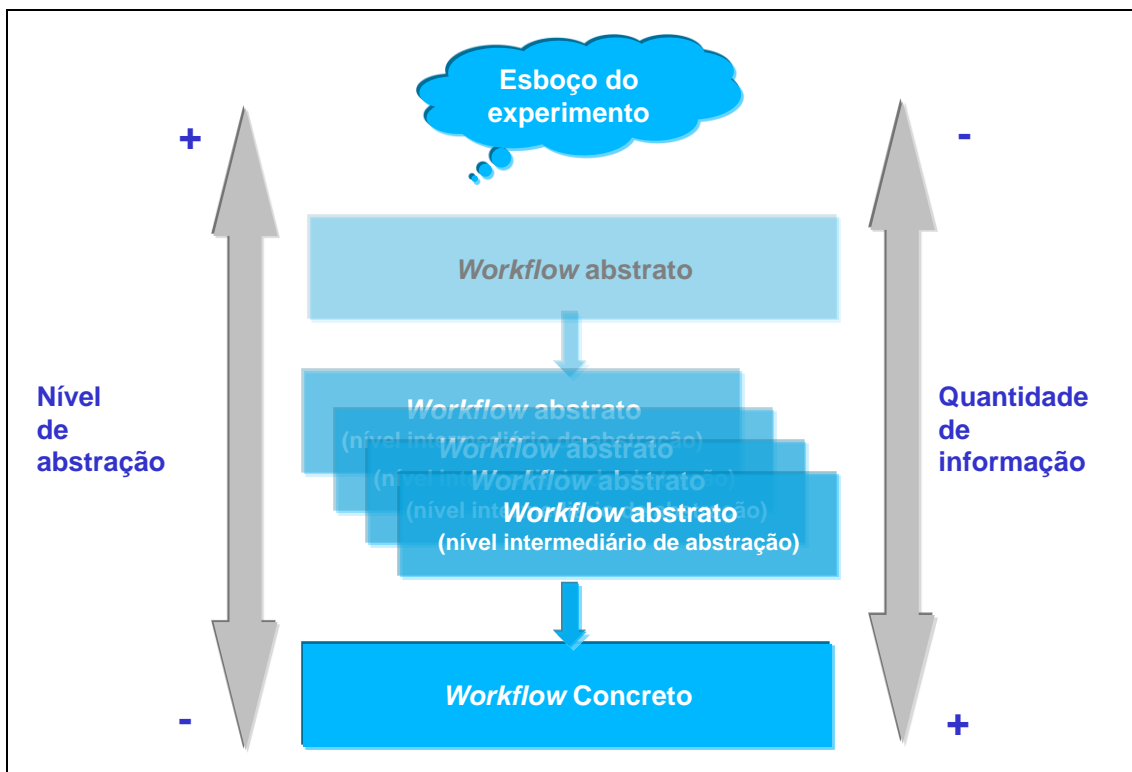


Figura 2.2. Representação da relação entre nível de abstração e quantidade de informação em um *workflow*.

Outra característica do *workflow* científico é relacionada à forma como a seqüência (fluxo) de atividades pode ser estruturada. Existem duas categorias de estruturas, a *DAG* (*Directed Acyclic Graph*) e a *non-DAG* (*Not Directed Acyclic Graph*) (YU & BUYYA, 2005). Numa estrutura baseada em *DAG* cada atividade representa um nó de um grafo, podendo haver vários filhos e a sua ordem de execução está ligada à seqüência dos nós. Além disso, a representação *DAG* contém estruturas paralelas que representam a capacidade de execução de atividades de forma concorrente. Já a representação *non-DAG*, entretanto, permite a repetição de atividades, possuindo estruturas de iteração na seqüência, como *loops* ou ciclos.

2.1.3 Ciclo de vida do *workflow* científico e do experimento

O *workflow* científico possui um ciclo de vida associado a sua existência (DEELMAN ET AL., 2009) e, em geral, é muito parecido com o processo de experimentação ou ciclo de vida experimental. Na literatura técnica, são encontradas diversas propostas para descrever e representar o ciclo de vida de um *workflow* científico. Como exemplo, pode-se citar LIVNY ET AL. (1994) que propôs um ciclo de vida baseado na definição genérica de estudos experimentais e que poderia ser aplicado nas mais diversas categorias de experimentos. Entretanto, um dos seus

problemas é estar centrado somente na etapa de execução, não detalhando outras etapas, como e.g. proveniência. Outro exemplo é BOSE ET AL. (2005), que estende o trabalho de LIVNY ET AL (1994), abordando a importância de se registrar a proveniência das informações geradas.

Contudo, MATTOSO ET AL. (2010) argumentam em seu trabalho que os ciclos de vida propostos possuem uma perspectiva similar e não atenderiam todas as suas etapas (composição; mapeamento; execução; proveniência). Além disso, esses ciclos de vida não distinguiriam o conceito de experimento da instância de *workflow*, não contemplando a possibilidade de lidar com abstrações. Para MATTOSO ET AL. (2010), o experimento científico é a abstração de nível mais alto, contendo diferentes alternativas de atividades, enquanto os *workflows* executáveis seriam instâncias de um caminho possível dentro do experimento, mapeados para recursos computacionais específicos e utilizando dados determinados para aquela rodada de execução.

A Figura 2.3 apresenta os relacionamentos entre as fases do ciclo de vida proposto em MATTOSO ET AL. (2010). A Composição é a primeira etapa dentro do ciclo de vida e é o momento onde são especificadas a seqüência, a ordem e as dependências em um *workflow* concreto ou abstrato. Contém a etapa de concepção do *workflow* e, possivelmente, a etapa de busca e recuperação. A importante subfase denominada concepção é o momento no qual o pesquisador concebe o *workflow* científico que representará seu experimento. A busca e recuperação têm por objetivo reutilizar *workflows* já concebidos para re-executá-los ou adaptá-los para novos objetivos.

A etapa de Execução é o momento no qual o *workflow* concreto é recuperado, juntamente com os insumos (dados) que serão consumidos, e este é enviado para uma “*engine*” (máquina de execução) que mantém o monitoramento da execução. A “*engine*” é responsável, além do monitoramento, pelo registro de *log* de eventos, invocação das ferramentas e transmissão dos dados. Nesta etapa, o *workflow* concreto é efetivamente ligado aos recursos computacionais, e.g. servidores, ambientes em grade, que irão executá-lo e por consequência gerar os dados da simulação. A etapa de Análise é o momento no qual o pesquisador realiza a avaliação dos dados gerados, visualizando-os, avaliando as hipóteses, tirando conclusões sobre os resultados, e, por fim, produzindo novos conhecimentos científicos.

A etapa central é a Proveniência, relacionada ao registro dos dados e metadados do *workflow* científico. A proveniência refere-se a todo histórico sobre a criação dos dados, sejam eles insumos, intermediários ou produtos, e do próprio protocolo do experimento, como hipóteses, modelos de *workflow* em diferentes níveis de abstração. Com esses dados há possibilidade de reaplicar os experimentos em outros contextos

para gerar novos dados ou reproduzi-lo a fim de verificar se os resultados anteriores são válidos ou verdadeiros. Diferente das demais, a etapa de proveniência acompanha todo o ciclo de vida do *workflow* científico, provendo e registrando as informações que determinam a qualidade e sua autenticidade.

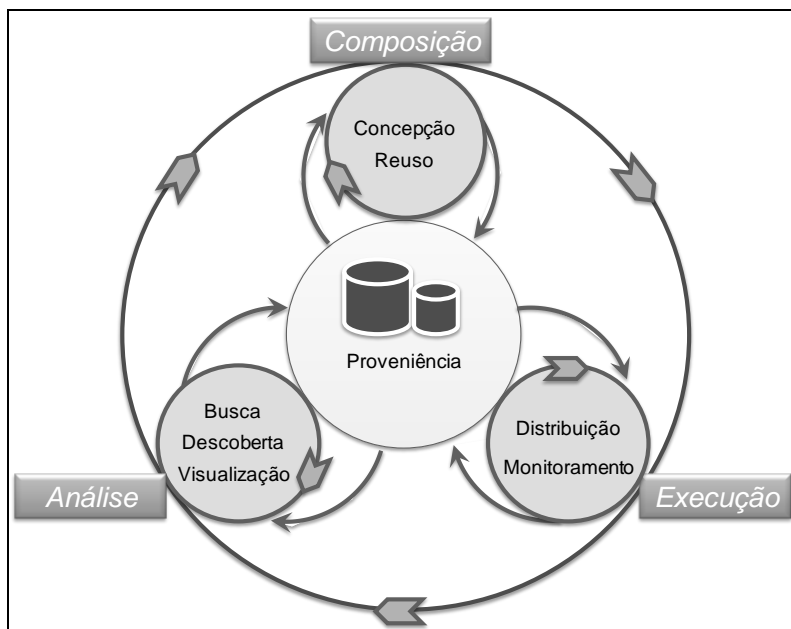


Figura 2.3. Ciclo de vida do experimento científico *in silico*, adaptado de MATTOSO ET AL. (2010).

Realizando uma comparação com o processo geral proposto por WOHLIN ET AL. (2000) e estendido por AMARAL (2003), a proveniência do *workflow* científico seria equivalente à etapa de acompanhamento do processo experimental. De fato, as etapas do ciclo de vida de MATTOSO ET AL. (2010) podem ser mapeadas para o processo experimental de WOHLIN ET AL. (2000) / AMARAL (2003), conforme a Tabela 2.1.

Tabela 2.1. Etapas do ciclo de vida de MATTOSO ET AL. (2010) mapeadas para o processo experimental de WOHLIN ET AL. (2000) / AMARAL (2003).

Etapa do ciclo de vida do <i>workflow</i>	Atividade do processo experimental	Características semelhantes
Composição	Definição e Planejamento	Definição do protocolo e preparação do experimento. No ciclo de vida, há a criação do <i>workflow</i> científico concreto a ser executado (instrumento).
Execução	Execução	Execução seguindo o protocolo, gerando-se os dados e informações.
Análise	Análise e Interpretação	Análise dos dados e informações geradas.
Proveniência	Empacotamento	Armazenamento das informações que caracterizam os dados criados, mantendo-se a qualidade e autenticidade. Facilitar a replicação do experimento.

2.1.4 Sistemas Gerenciadores de *Workflow* científico - SGWfC

Os Sistemas Gerenciadores de *Workflow* Científico (SGWfC) são responsáveis por invocar aplicações utilizadas nas atividades do *workflow*, controlando os dados passados como entradas e saídas. Em geral, permitem que *workflows* sejam modelados e especificados através de interfaces gráficas, assim como realizam o monitoramento da sua execução. É importante ressaltar que as funcionalidades providas pelos SGWfCs variam de um para outro. Porém, de uma maneira geral espera-se atender os seguintes requisitos (MATTOS ET AL., 2008).

- Incluir e estender dados, processos e recursos utilizados dentro do *workflow*, além de mecanismos de controle de execução;
- Auxiliar na definição e redefinição das propriedades dos processos, dados e recursos determinados para um *workflow* científico específico;
- Verificar entradas e saídas definidas para cada atividade do *workflow*, além de disponibilizar meios para conversão dos dados ou verificar se os resultados gerados eram esperados;
- Otimizar o *workflow* científico definido, durante sua execução. Também deve permitir o acompanhamento da execução e interrupções temporárias do fluxo;
- Possibilitar o agendamento da execução para um determinado momento desejado, ou até mesmo de intervalos de tempo, e;
- Armazenar tanto os dados quanto os meta-dados produzidos pelo *workflow*. Também deve permitir a um usuário a capacidade de consultar e atualizar os meta-dados.

Existem diversos SGWfCs descritos na literatura técnica, um exemplo é o Kepler (ALTINTAS ET AL., 2004), sistema utilizado para construção e execução de *workflows* científicos, largamente usado na área de Bioinformática. O Kepler permite uma modelagem visual, através do conceito de módulos, no qual o usuário do sistema precisa configurar os recursos computacionais dentro do SGWfC. Desta maneira, trabalha com o conceito de *workflow* concreto, pois não existe a modelagem somente do comportamento do experimento. O Kepler é capaz de verificar o modelo, contudo, a verificação só é executada quando solicitada pelo usuário. Entretanto, não permite a

reutilização de modelos criados dentro de outros modelos, o que acaba por dificultar a reutilização de *workflows*, pois toda a modelagem deve ser repetida em outro momento.

Outro exemplo de SGWfC é o sistema Wings (KIM, GIL & RATNAKAR, 2006; GIL ET AL., 2007a), sistema de modelagem de *workflow* que usa representação semântica e técnicas de planejamento para apoiar a criação de *templates* de *workflow* e suas respectivas instâncias. Instâncias do *workflow* devem ser submetidas ao sistema Pegasus (KIM, GIL & RATNAKAR, 2006), para alocação dos recursos em ambientes em grade e a execução do *workflow*. O *workflow* é representado em uma linguagem chamada *DAX (Directed Acyclic Graph in XML)*. Os elementos que compõem o *template* do *workflow* são: *Files; Componentes; Restrições; Input; Output*. Estes elementos são especificados através de uma ontologia do domínio da área científica ou engenharia do experimento. A ontologia do domínio e os *templates* de experimentos precisam ser criados por um especialista, cabendo a um pesquisador menos experiente a tarefa de instanciar o *workflow*, utilizando a interface gráfica para inserir os dados. A última etapa é a alocação dos recursos computacionais (e.g.: ambientes em grade) para o *workflow* executável instanciado, o que é feito pelo sistema Pegasus que também é responsável por movimentar os dados entre os recursos computacionais no decorrer do fluxo. A representação deste *workflow* é *DAG*, portanto não suporta *laços*.

Em geral, nos diferentes SGWfC, os *workflows* criados não são exportáveis para outros sistemas, e isto ocorre porque eles são dependentes da plataforma na qual foram concebidos e possuem portabilidade reduzida. Mesmo para a sua especificação e modelagem não é possível estar independente da plataforma, com isso, o reuso desses *workflows* previamente definidos pode ser extremamente limitado em contextos e ambientes diferentes do experimento original.

2.2 Definição dos conceitos sobre elementos do domínio de *workflow* científico

A tecnologia de *workflow* científico é um instrumento disponível para realizar experimentação *in silico* (e *in virtuo*) e como tal contém conceitos próprios. Neste contexto, o *workflow* se caracteriza por um modelo representando o experimento científico que contém um ou mais fluxos possíveis de execução. Cada fluxo de execução é composto por um conjunto de atividades próprias, encadeadas a fim de representar a ordem de inicialização e execução (TAYLOR ET AL., 2007; DEELMAN ET AL., 2009).

Em alguns casos, um *workflow* pode ser composto por outros *workflows* e assim, quando existe esta relação, define-se que estes últimos representam *Sub-workflows* (YU & BUYYA, 2006). Os *Sub-Workflows* possuem as mesmas características de um *workflow*, ou seja, fluxo de execução interno, representação da ordem de suas atividades, porém fazem parte do fluxo de execução do *workflow* principal.

Uma atividade é definida como um passo do experimento científico, um processo ou lógica utilizado na simulação (INGALLS, 2002). Uma atividade transforma, manipula ou gera algum artefato durante sua execução. Em geral, é considerado um elemento atômico, isto é, não contém fluxo de execução ou outras atividades internas. Para este caso, também existe o conceito de sub-atividade, que seria uma atividade pertencente à outra (BARBARA, MEHROTRA & RUSINKIEWICZ, 1996), muito próximo ao conceito expresso em *sub-workflow*. Além disso, o conceito de sub-atividade também pode ser utilizado para expressar a relação entre o *workflow* e a atividade (FAHRINGER ET AL., 2004).

Um artefato pode ser definido como uma informação criada, modificada ou manipulada durante a execução de uma atividade, ou representações de um dado ou informação utilizado, ou ainda um conhecimento científico gerado através da análise do pesquisador sobre uma informação (FLOWER, 2003).

Durante sua execução, as atividades são apoiadas em maior ou menor intensidade por ferramentas, que se caracterizam por serem recursos computacionais, como sistemas, *softwares* ou programas (FLOWER, 2003). Quando uma ferramenta apoia mais intensamente uma atividade, a própria execução da ferramenta ou conjunto de ferramentas utilizadas representa a realização da atividade e, neste caso, a participação do pesquisador durante a execução não é essencial. Entretanto, quando uma ferramenta apoia com menos intensidade uma atividade, a execução da ferramenta não é equivalente à realização da atividade, sendo mais importante a participação do pesquisador ou cientista.

Todos os conceitos descritos se apresentam geralmente nos modelos de *workflow* científico no nível concreto, pois são os elementos entendidos pelos Sistemas Gerenciadores de *Workflow* científicos. Na Tabela 2.2, tem-se os elementos apresentados e suas definições dentro do contexto de experimentação suportada por *workflow* científico.

Tabela 2.2. Elementos da experimentação baseada em *workflow* científico e suas descrições.

Elemento	Descrição
Workflow	<i>Workflow</i> se caracteriza por conter um encadeamento das atividades principais do experimento científico, formando um fluxo de execução.
Sub-workflow	Sub- <i>Workflow</i> se caracteriza como um <i>workflow</i> que compõe o fluxo de execução de outro <i>workflow</i> , contém um sub-conjunto de atividades.
Atividade	Atividade se caracteriza como a representação de um passo dentro do fluxo de execução do <i>workflow</i> do experimento científico.
Artefato	Artefato se caracteriza como informação criada, transformada, modificada ou manipulada de alguma forma por atividades.
Ferramentas	Ferramenta se caracteriza como um recurso computacional utilizado para apoiar parcial ou totalmente uma atividade. Em geral, são sistemas, <i>software</i> e programas que auxiliam na execução.
Sub-atividade	Atividade que pertence a um sub- <i>workflow</i> ou a outra atividade.

2.3 Linguagens e notações para representação de *workflow* científico

Atualmente existe uma gama de linguagens que permitem a representação de *workflow* descrita na literatura técnica. Algumas dessas linguagens, além de permitir descrever o *workflow* textualmente, contêm notações gráficas associadas que são utilizadas na composição do *workflow* de maneira visual. Nem todas as linguagens inicialmente eram utilizadas para conceber *workflow* científico. Na verdade, muitas surgiram para descrever o *workflow* de negócio, que é a tecnologia utilizada para automatizar uma sequência de atividades de negócio (HOLLINGSWORTH, 1994). Entretanto, com o interesse em automatizar as atividades científicas, algumas linguagens foram aplicadas por grupos de pesquisa no desenvolvimento de *workflow* científico para seus experimentos, como em SUE & YUE (2010) com uso de BPEL (OASIS, 2007) e OGASAWARA ET AL. (2009) com o uso de XPD (WfMC, 2008).

Na última década, outras linguagens surgiram diretamente para a área científica, buscando atender a especificidades do *workflow* científico, e.g. manipulação de grandes volumes de dados, a utilização de recursos computacionais em grade, dentre outras. Essas linguagens acabaram por ser incorporadas aos SGWfCs como parte das soluções propostas (FAHRINGER ET AL., 2004). O objetivo desta seção não é se estender e analisar detalhadamente todas as linguagens disponíveis, mas apresentar um conjunto significativo para exemplificar os esforços na área de automação de atividades.

A notação BPMN (*Business Process Modeling Notation*) tem como principal objetivo prover uma notação gráfica a ser entendida por qualquer indivíduo envolvido na criação de um *workflow* de negócio, além de servir como uma ponte de ligação

entre o projeto do *workflow* e sua *implementação* (OMG, 2010b). Outro objetivo é servir como notação gráfica para diferentes linguagens de execução baseadas em XML, isto é, a notação BPMN não é uma linguagem para execução, mas um padrão de notação para modelagem. Com o BPMN pode-se modelar tanto fluxo de controle (sequência de execução de atividades) quanto fluxo de dados. Com o BPMN se é capaz de representar pontos de decisão sobre os fluxos (controle e dados) e oferece a possibilidade de representar estruturas de iteração.

A linguagem XPD (XML Process Definition Language) (WfMC, 2008) tem por objetivo descrever de forma declarativa um *workflow*, em formato XML, para que possa ser utilizado em diferentes ferramentas. A XPD define um esquema XML para especificar de forma declarativa parte do *workflow* e, além disso, é capaz de representar o *workflow* tanto semanticamente (descrição XML) quanto graficamente (modelos). A XPD é relacionada à notação BPMN, pois foi definida para armazenar todos os elementos da notação e assim podem-se modelar os fluxos de controle e dados, além de representar estruturas de decisão e iteração.

A linguagem BPEL (*Business Process Execution Language*) tem como objetivo especificar um conjunto de serviços *web* que representam um *workflow* de negócio. BPEL é uma linguagem de execução e não possui uma notação gráfica para representar visualmente os *workflows*. A BPEL foi desenvolvida para prover a definição da sequência de interações entre os serviços *web* e o fluxo de dados transmitidos de um para outro durante a execução (OASIS, 2007).

A linguagem SCUFL (*Simple Conceptual Unified Flow Language*) tem como objetivo representar um *workflow* científico de forma simplificada no SGWfC Taverna (OINN ET AL., 2004). A SCUFL é uma linguagem baseada em XML que contém uma notação gráfica associada, permitindo a representação de *workflow* em alto nível de abstração (através de modelos). A SCUFL contém elementos que permitem a representação de fluxo de dados e de controle, havendo também a definição de estruturas para representação de iterações e *sub-workflow*, contudo a representação de estruturas de decisão não é apoiada.

A linguagem AGWL (*AGWL: Abstract Grid Workflow Language*) surgiu da necessidade de conceber *workflows* científicos para serem utilizados em ambientes computacionais em grade. O objetivo da AGWL é descrever um *workflow* para aplicações em grade através de um modelo em alto nível de abstração, protegendo o usuário de detalhes de *implementação*. A linguagem é baseada em XML e não possui uma notação gráfica associada, servindo como descrição para outras notações. A AGWL contém estruturas de decisão e de iteração e permite a representação de fluxo de controle e de dados entre as atividades (FAHRINGER ET AL., 2004).

A UML - *Unified Modeling Language* surgiu para especificar e modelar artefatos de sistemas de *software* orientado a objetos, sendo composta por um conjunto de notações que permitem a modelagem de diversas perspectivas desses sistemas (OMG, 2009). Contudo, não só softwares são representados, mas modelos de *workflow*, incluindo *workflow* científico, também podem através do uso da UML, mais especificamente através da representação pela notação de Diagrama de Atividades. Um exemplo é o trabalho descrito por PLLANA ET AL. (2005), que propõem um sistema gerenciador no qual os seus modelos são expressos como diagrama de atividades. Para tal, eles realizam uma extensão da notação para representar as funcionalidades do sistema, como por exemplo, paralelismo. Como o diagrama de atividades e classes da UML são utilizados na abordagem proposta nesta dissertação, estes serão apresentados com mais detalhes na próxima seção.

2.4 Revisão da notação UML 2

A UML 2 é a versão mais atual e, como descrita, é uma linguagem de modelagem padronizada e composta por um conjunto de diagramas capazes de representar graficamente diferentes visões do sistema ou *software* em desenvolvimento (OMG, 2009). Cada diagrama contém uma notação padrão associada, capaz de representar um aspecto do sistema. A versão atual da UML conta com um conjunto de 14 diagramas, que podem ser agrupados quanto à visão que representam: Estruturais e Comportamentais. Os diagramas estruturais enfatizam a estrutura do *software* e o que o compõe, em geral, representando-o de forma estática, enquanto os diagramas comportamentais enfatizam o que deve ocorrer no *software*, sendo usados para descrever as suas funcionalidades. Nesta seção são revisadas as notações de dois diagramas utilizados na abordagem proposta nesta dissertação: o diagrama estrutural denominado Diagrama de Classes e outro comportamental denominado Diagrama de Atividades.

2.4.1 Notação do diagrama de classes

A notação de diagrama de classes da UML 2 (OMG, 2009) representa a estrutura e as relações entre as classes que compõem o domínio do sistema ou *software*. Esse diagrama pertence à categoria de modelos que representam a estrutura estática do sistema. No diagrama estão modeladas de maneira estática as classes do *software* e as relações existentes entre elas. Esta notação é revisada nesta seção, pois é utilizada na proposta descrita nesta dissertação.

Uma classe é definida como a descrição de um conjunto de objetos que partilham a mesma especificação de funcionalidades, restrições e semântica. No

diagrama de classes, estas são representadas graficamente através de retângulos, divididos em três compartimentos. Na Figura 2.4, no sentido de cima para baixo, tem-se o primeiro compartimento que representa o nome da classe, no segundo compartimento são descritos os atributos das classes, e no terceiro compartimento temos os métodos das classes.

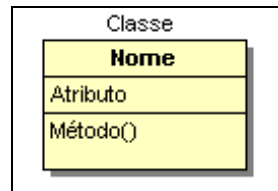


Figura 2.4. Elemento Classe da UML 2.

No diagrama de classes são definidos tipos diferentes de relacionamento entre as classes. A Associação é um relacionamento básico entre duas classes, significando que elas colaboram entre si. A Figura 2.5 representa a associação simples. Numa relação podem ser expressas as cardinalidades, que representam a quantidade de objetos de outra classe que um objeto pode se relacionar. Há também a descrição dos papéis envolvidos em uma relação e o nome da associação.

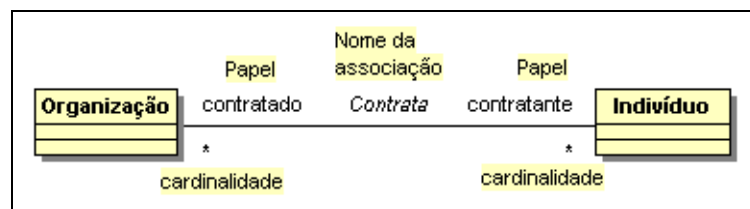


Figura 2.5. Associação simples.

Existem ainda outros relacionamentos que expressam uma semântica específica como todo-parte, ou seja, nesta relação existe uma classe “parte” que “pertence” a uma classe “todo”. Um destes relacionamentos é a composição no qual uma parte está necessariamente relacionada a um único todo e ainda há o relacionamento de agregação na qual uma parte pode ser compartilhada com outros todos. Esses dois relacionamentos estão representados na Figura 2.6.

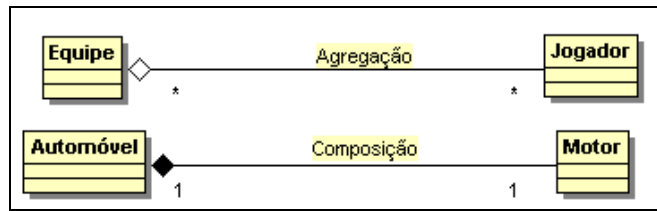


Figura 2.6. Relacionamento de agregação e composição, respectivamente.

A Figura 2.7 apresenta a notação gráfica de outros dois relacionamentos existentes na notação, Especialização e Auto relacionamento. A especialização é a representação de refinamentos nas abstrações, podendo identificar classes mais genéricas ou mais específicas em um diagrama de classes. Já o auto relacionamento é uma associação entre objetos da mesma classe, sendo que cada objeto tem um papel distinto na associação.

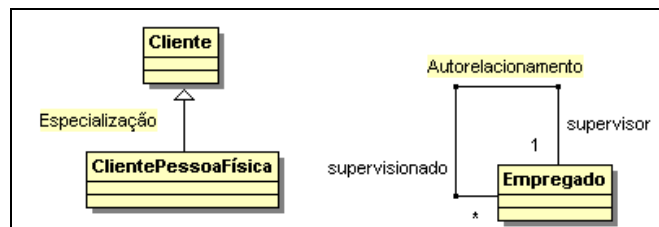


Figura 2.7. Relacionamento de especialização e autorelacionamento.

2.4.2 Notação do diagrama de atividades

O diagrama de atividades da UML 2 é um modelo que descreve os comportamentos de um sistema. Modelos comportamentais, em geral, determinam quando um comportamento específico do sistema deve iniciar e quais são suas entradas (BOCK, 2003a). O diagrama de atividades segue as abordagens de modelagem tradicionais representando os fluxos de dados e controle, definindo que outras ações devem ser iniciadas de acordo com o término de ações predecessoras e a disponibilidade de dados de entrada necessários.

A Figura 2.8 apresenta o extrato do metamodelo da UML 2, contendo os elementos pertencentes ao diagrama de atividades e também representa s relacionamentos entre eles. A UML 2 define diversos elementos a serem utilizados na modelagem de diagramas de atividade, os quais contém conceitos e significados próprios. Serão apresentados, entretanto, resumidamente nesta seção.

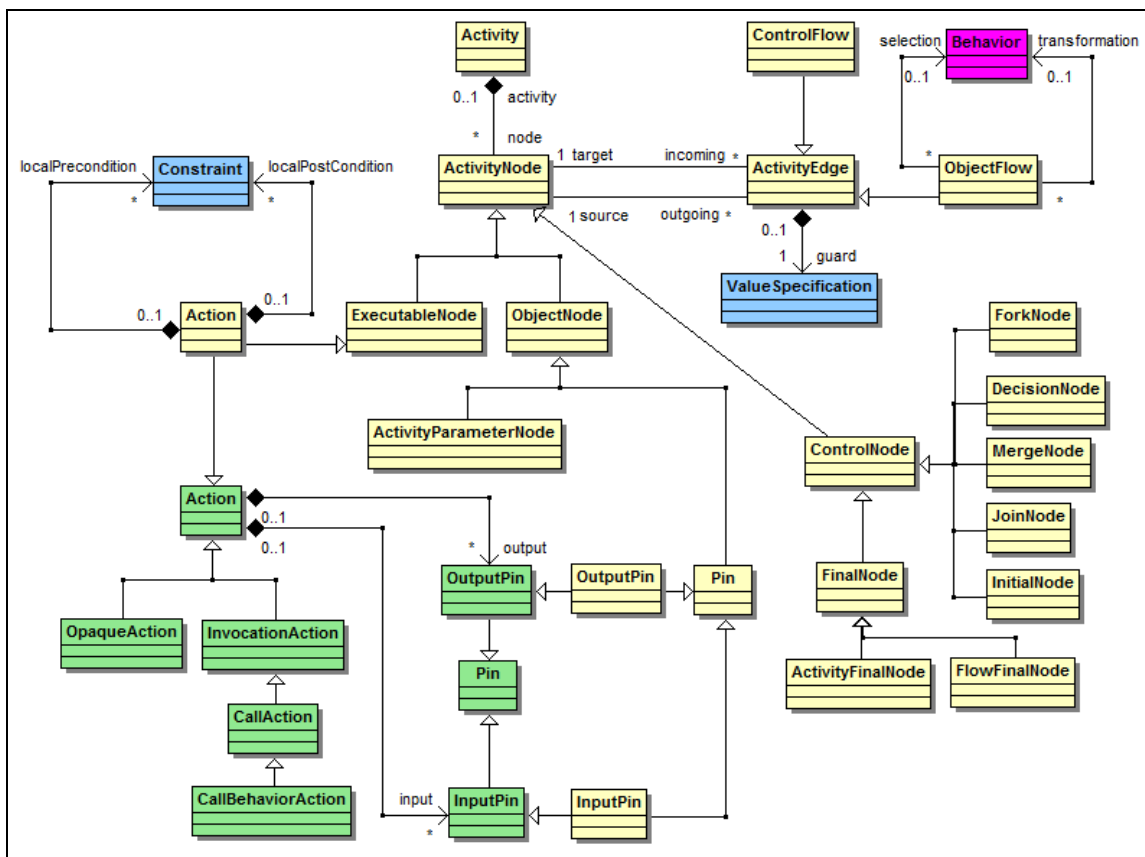


Figura 2.8. Extrato do Metamodelo da UML 2.2 contendo elementos da notação do diagrama de atividades. Também estão representados os relacionamentos entre esses elementos.

Elementos Atividades e Ação:

Na especificação UML 2, uma atividade representa um comportamento definido para o sistema. Na notação ela é representada pelo elemento **Atividade** (*Activity*) (Figura 2.9 - a). A atividade pode ser composta por sub-elementos (atividades, ações, nós de decisão, arestas) encadeados a formar um fluxo de execução.

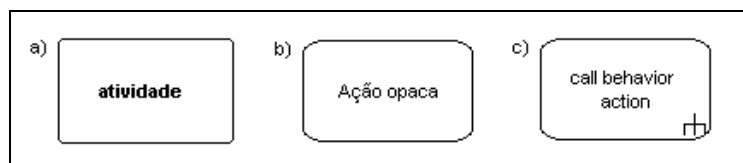


Figura 2.9. Notação gráfica de Atividade (a), Ação (b) e elemento CallBehaviorAction (c).

Na UML, **Ações** são elementos contidos diretamente em atividades, um dos tipos denominado como **Ação Opaca** (*Opaque Action*) e sua notação gráfica representada na Figura 2.9 - b. As ações são um tipo especial de elemento nas atividades, que se conectam a outros elementos através arestas e representam um passo atômico dentro do fluxo de execução, não podendo conter sub-elementos

(outras Atividades ou tipos de Ação). Na UML 2, a atomicidade difere a Ação e a Atividade (BOCK, 2003a).

Na notação de diagrama de atividades, para indicar que uma determinada atividade (comportamento) é invocada durante a execução de um fluxo, pode-se utilizar um tipo de Ação especializada, já definida na própria especificação da UML 2. Essa Ação é chamado de *CallBehaviourAction* (Figura 2.9 - c).

Elementos arestas de ligação do modelo:

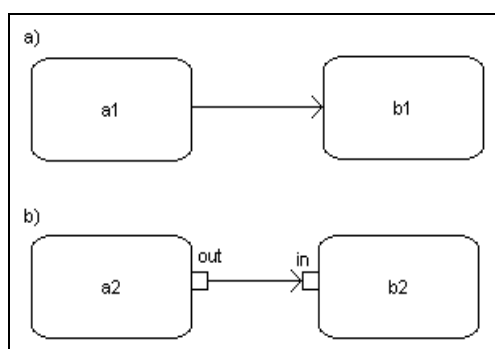


Figura 2.10. Notação de Arestas: Fluxo de controle (a) e Fluxo de dados (b).

As ligações e o encadeamento entre as Ações e Atividades no diagrama de atividades são representados através das **Arestas** (*Activity Edge*). A Figura 2.10 contém a representação gráfica das arestas no diagrama de atividades da UML. Na notação são definidos dois tipos de arestas: **Fluxo de controle** (*Control Flow Edge*) e **Fluxo de dados** (*Object Flow Edge*). A aresta Fluxo de controle (Figura 2.10 – a) representa o encadeamento da seqüência de execução de Atividades e Ações. A aresta Fluxo de controle não transmite informações ou dados entre os dois elementos conectados, somente indica ordem de execução. Já a aresta de Fluxo de dados (Figura 2.10 – b), é caracterizada por indicar não só a ordem entre atividades, mas também por transmitir dados e/ou informações, representando assim movimentação das informações entre os elementos Atividades ou Ações (BOCK, 2003b).

Elementos portas e parâmetros do modelo:

As arestas de Fluxo de dados são conectadas às **Portas** (*Pin*) ou aos **Parâmetros** (*Parameter Node*), tendo suas notações representadas na Figura 2.11 – a e Figura 2.11 – b, respectivamente (BOCK, 2004). Ambos os elementos Portas e Parâmetro têm o conceito de sentido dos dados associados, isto é, têm sentido de entrada quando esperam um dado como **insumo** (*Input Pin* ou *Parameter node*) ou

sentido de saída quando disponibilizam um dado como **produto** (*Output Pin* ou *Parameter node*).

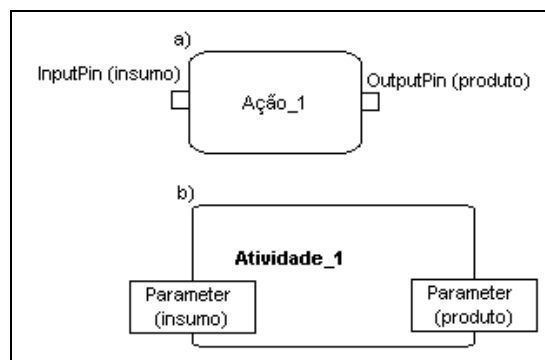


Figura 2.11. Notação Pin: Input pin e Output pin (a). Notação de Parameter node (b).

Sintaticamente na UML, a Atividade só pode possuir elementos do tipo Parâmetros *Parameter node* associado, utilizado tanto para representar insumos quanto produtos, enquanto, a Ação só pode possuir elementos do tipo Porta (*pin*) associada a ela. Uma Porta é especializada dependendo do sentido em que os dados estão chegando para Ação, caso sejam dados de entrada (insumo) é representada por *Input pin* e caso sejam dados de saída é representada por *Output pin*.

Elementos de controle e decisão do modelo:

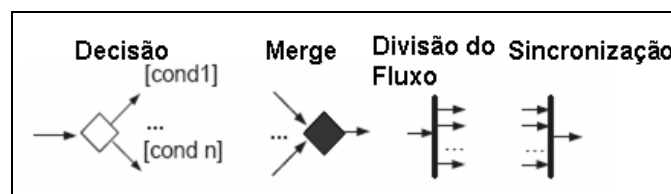


Figura 2.12. Elementos de controle e decisão.

Na notação do diagrama de atividades podemos encontrar os **nós de controle** (*Control Node*) (BOCK, 2003c). Existem sete elementos especializados, representados nas Figura 2.12 e Figura 2.13. O principal objetivo dos nós de controle é realizar operações nos fluxos de dados e controle presentes em uma atividade (comportamento do sistema). Os nós de controle para manipulação de fluxos são:

- Tipo **Decisão** (*Decision Node*) representa o início de um comportamento condicional, no qual cada aresta de saída representa um caminho de execução que pode ser seguido. O nó do tipo Decisão deve apresentar as condições de guarda (entre colchetes []), que representam os possíveis resultados ou condições da decisão.

- Tipo **Merge** (*Merge Node*) é utilizado quando se deseja marcar o fim de um comportamento condicional, imediatamente passando para uma aresta de saída todos os fluxos de dados ou de controle.
- Tipo **Divisão de fluxo** (*Fork Node*) representa a divisão do fluxo de dados ou de controle em diversos fluxos de saída ao mesmo tempo; há uma replicação dos controles e dados oriundos dos fluxos de entrada.
- Tipo **Sincronização** (*Join Node*) representa a junção de todos os fluxos de entrada (dados e controles) em um único fluxo de saída. Esta saída só estará pronta quando todos os fluxos de entrada estiverem disponíveis para o nó do tipo Sincronização.

Na notação ainda existem os nós de controle que representam o início e término de fluxos de execução. Eles estão representados na Figura 2.13. O **nó de início** (*Initial Node*) é utilizado para representar o início de um fluxo de execução, que pode ser um Fluxo de dados ou Fluxo de controle, enquanto o **nó final de atividade** (*Activity Final Node*) é utilizado para representar o término do fluxo de execução da atividade, isto é, se um desses elementos é alcançado durante a execução, então o fluxo de execução se encerra. Ainda há o **nó final de fluxo** (*Flow Final Node*), que também determina o término, mas somente indica o fim de um possível caminho de execução e não de toda a atividade.

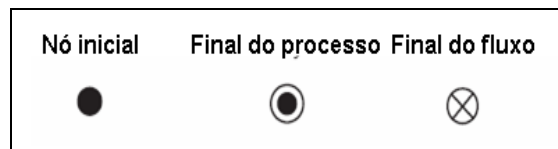


Figura 2.13. Elementos de início e fim de fluxo.

2.5 Considerações finais do capítulo

Nesse capítulo foram abordados os conceitos relacionados à experimentação, descrevendo o processo de experimentação e as diferentes categorias existentes. Dentre essas, em especial, se encontram os experimentos *in virtuo* e *in silico* cuja característica é a simulação do ambiente, objetos sobre avaliação e, possivelmente, participantes através do uso de modelos computacionais. Esses modelos estão geralmente na forma de *softwares* que precisam interagir entre si para alcançar o objetivo definido para o experimento *in silico* (e *in virtuo*).

Entretanto, como há grande dependência de recursos computacionais, deve-se definir previamente uma infraestrutura computacional (servidores, grades, dentre

outras) para execução desses experimentos. Isto acarreta em maior complexidade no planejamento e tendo isso em vista, tecnologias surgiram para facilitar essas atividades, sendo uma delas o *workflow* científico. *Workflow* científico tem por objetivo automatizar as atividades do experimento, fazendo a invocação das ferramentas (*softwares*) que são utilizadas e transmitindo os dados gerados entre elas.

Na área de simulação, sistemas foram desenvolvidos com o intuito de apoiar o ciclo de vida do *workflow* científico (composição, execução, análise e proveniência), denominados Sistemas Gerenciadores de *Workflow* Científico (SGWfC) e neste capítulo foram apresentados os principais requisitos para que esse objetivo seja alcançado. Foram também abordados os conceitos que fazem parte dos modelos de *workflow* em geral, como atividade, ferramenta, *sub-workflow* e artefato. Por último, foram descritas algumas linguagens para modelagem de *workflow* em geral e posteriormente foram abordados os diagrama de atividades e diagrama de classe da UML 2, as notações de modelagem utilizadas na abordagem proposta neste trabalho.

CAPÍTULO 3 - QUASI REVISÃO SISTEMÁTICA SOBRE COMPOSIÇÃO DE EXPERIMENTOS BASEADOS EM WORKFLOW CIENTÍFICO

Neste capítulo é apresentada uma quasi revisão sistemática da literatura técnica sobre composição de workflow científico para experimentos in virtuo e in silico, cujo principal objetivo é caracterizar o estado da arte e prática nesta área de estudo. A sua execução indicou a necessidade de se tornar disponível uma abordagem que descreva explicitamente os procedimentos para concepção de workflow científico no nível abstrato.

3.1 Introdução

Um dos principais desafios em experimentos *in virtuo* e *in silico* está relacionado ao processo de gerência de *workflows* científicos (MATTOSO ET AL., 2008), pois um dos problemas encontrados é como conceber um *workflow* científico abstrato para um experimento e, posteriormente, transformá-lo em um *workflow* concreto. Esta questão traz preocupações relativas à documentação do conhecimento utilizado para criar o *workflow* científico e a sua correção. No ciclo de vida de um experimento científico em larga escala, este problema se encontra na fase de composição, mais especificamente na etapa de concepção.

Desta forma, antes de propor qualquer solução, é necessário realizar uma pesquisa na literatura técnica a fim de caracterizar quais são e se existem abordagens ou métodos que sejam utilizados para a composição de *workflow* científico e como essas abordam a garantia de qualidade. Então, para revisar a literatura foi decidido executar uma *quasi*-revisão sistemática. Uma *quasi*-revisão sistemática é um tipo de estudo secundário (TRAVASSOS ET AL., 2008), no qual o objetivo é identificar, avaliar e interpretar os trabalhos existentes na literatura técnica sobre um determinado tópico de interesse, baseada em uma estratégia de pesquisa formal, através de um protocolo de revisão (TRAVASSOS ET AL., 2008). Mas diferentemente da Revisão sistemática tradicional (KITCHENHAM, 2004; KITCHENHAM ET. AL., 2009), o principal objetivo da *quasi* revisão é a caracterização da área ou não existe objeto para comparação.

A *quasi* revisão sistemática tem como objetivo explorar a literatura técnica a fim de caracterizar as abordagens ou métodos utilizados para a composição de *workflow* científico para experimentos *in silico* ou *in virtuo*. A *quasi* revisão sistemática diz respeito a aplicação de abordagens na construção de *workflows* pela perspectiva do Engenheiro de *software* ou cientistas. Também, como objetivo secundário, existe interesse em caracterizar se dentre as abordagens existentes, há alguma que utilize técnica de verificação e ou validação para garantia de qualidade. Abaixo, na Tabela 3.1 e Tabela 3.2, estão descritos os objetivos específicos deste estudo, baseados na abordagem GQM (SOLINGEN & BERGHOUT, 1999). A seguir é descrito o protocolo que será utilizado na *quasi*-revisão sistemática.

Tabela 3.1. Objetivo específico relacionado à questão primária.

Analisar:	As abordagens ou métodos para construção de <i>workflows</i> científicos.
Com o propósito de:	Caracterizar.
Em respeito a:	Sua aplicação na construção de <i>workflows</i> (Fases do Ciclo de vida, Tipo de atividades executadas, instrumentos executados e processos).
Do ponto de vista do:	Engenheiro e cientista.
No contexto da:	Composição de <i>workflows</i> para experimentos científicos <i>in silico</i> e <i>in virtuo</i> .

Tabela 3.2. Objetivo específico relacionado à questão secundária do estudo.

Analisar:	As abordagens ou métodos para composição de <i>workflows</i> científicos.
Com o propósito de:	Caracterizar.
Em respeito a:	Existência de apoio às atividades de verificação e validação (inspeção) utilizadas (Tipo, instrumentos e procedimentos).
Do ponto de vista do:	Engenheiro e cientista.
No contexto da:	Composição de <i>workflows</i> para experimentos científicos <i>in silico</i> e <i>in virtuo</i> .

3.1.1 Protocolo da *quasi* revisão Sistemática

O protocolo contém a questão de pesquisa; os métodos utilizados para executar a revisão; os critérios de inclusão e exclusão para os estudos primários recuperados; documentação da estratégia de busca utilizada para permitir que possa ser avaliado quanto ao rigor e à completude; os procedimentos de avaliação da qualidade dos trabalhos retornados; e a estratégia para extração dos dados. A seguir, os requisitos do protocolo de pesquisa são apresentados tomando por base a proposta de BIOLCHINI ET AL. (2005).

Escopo de pesquisa:

O escopo da revisão é sobre composição de *workflows* científicos e não são contempladas as outras fases do ciclo de vida de um *workflow* científico, isto é, sua execução, sua análise e sua proveniência. Esta pesquisa também não tem o objetivo

de caracterizar Sistemas Gerenciadores de *Workflows* Científicos (SGWfC), mas sim, focar nas abordagens que possam ser utilizadas por estes.

Questão de pesquisa:

A questão de pesquisa é baseada no escopo e no objetivo do estudo (Tabela 3.1 e Tabela 3.2), definida como:

- **Questão de pesquisa:** Que abordagens (ou métodos) são utilizadas para a composição de *workflows* científicos?
- **Questão de pesquisa adicional:** Quais destas abordagens (ou métodos) oferecem algum procedimento de garantia de qualidade, particularmente, para verificar (revisão) e validar dos *workflows* científicos?
- **População:** Pesquisas ou trabalhos baseados em *workflow* científico, descritos na literatura técnica.
- **Intervenção:**
 - Para a questão principal: abordagens (ou métodos) para a composição de *workflow* científico.
 - Para questão secundária: abordagens (ou métodos) de composição *workflow* científico que oferecem algum procedimento de garantia de qualidade dos *workflows* científicos.
- **Resultados:** abordagens (ou métodos) de composição de *workflow* científico. Identificação dos procedimentos para garantia de qualidade de *workflow* científico.
- **Design Experimental:** espera-se a apresentação de pelo menos um exemplo de aplicação ou uma prova de conceito.

Estratégia para pesquisa

A estratégia utilizada para pesquisa contempla a busca em bases de dados eletrônicas, incluindo journals e anais de conferências. São também utilizados os Anais do SBBB e *e-Science* (base bdbcomp¹). As editoras selecionadas entre os periódicos disponíveis nas bases de dados: SCOPUS, IEEE Xplore, Science Direct e Compendex EI (internacionais); Anais SBBB e *e-Science* (nacionais).

¹ <http://www.lbd.dcc.ufmg.br/bdbcomp/>

Os termos utilizados na pesquisa são em inglês, por essa ser a língua utilizada nas bases de dados eletrônicas consultadas. Os principais termos utilizados na pesquisa são:

- *Workflow: experiment, process.*
- *Scientific: e-science, science*
- *Composition: development, modeling, modelling, specification, design, definition, orchestration, conception, description, analysis, representation.*
- *Approach: method, technique, model, process.*
- *Inspection: review, verification, validation, reading, revise.*

O idioma dos trabalhos retornados pode ser inglês, português ou espanhol, pois são os idiomas compreensíveis para os pesquisadores executando a revisão. Os tipos de trabalho podem ser tanto teórico, experimental ou descritivo, pois o tópico de interesse não está restrito a nenhum tipo de *design* experimental.

Critério de seleção de trabalhos e procedimentos

Nem todos os trabalhos e artigos retornados após a execução da busca são relevantes à questão de pesquisa, pois podem conter os termos definidos para a questão, contudo não abordarem o tópico de interesse. Assim, devem ser definidos critérios de inclusão/exclusão para aplicação no conjunto retornado. Se o artigo atende a todos os critérios ele é incluído para a próxima etapa de avaliação; se não, ele é excluído. Para essa pesquisa os critérios são:

- Os trabalhos podem estar em inglês, português ou espanhol.
- Os trabalhos devem tratar da fase de composição do ciclo de vida do *workflow* científico.
- Questão primária: Trabalhos (estudo, relato, proposta ou descrição) que apresentem uma abordagem (ou método) para composição de *workflow* científico.
- Questão secundária: Trabalhos (estudo, relato, proposta ou descrição) que apresentem uma abordagem (ou método) para composição de *workflow* científico, que utilize verificação e validação (inspeção) para controle de qualidade.

O procedimento de seleção contém duas etapas: preliminar e final. No procedimento preliminar, a estratégia de pesquisa é aplicada por um pesquisador, assim identificando trabalhos potenciais, retornados pela base de dados. Todos os trabalhos são documentados e suas citações são importadas para a ferramenta

JabRef², no formato RIS, com todos os campos selecionados. Depois, baseado nos títulos, no resumo (abstract) e palavras chaves (keywords), o pesquisador realiza marcações, seguindo a classificação: I – incluído; E – excluído; N – não definido. Um segundo pesquisador revisa a lista e evolui a classificação. Ao final os trabalhos são excluídos caso não apresentem uma abordagem para a composição de *workflow* científico. Caso não fique claro através dos títulos, no resumo (abstract) e palavras chaves (keywords) se o trabalho está em conformidade com o critério anterior, este é incluído para uma posterior análise qualitativa. Os resultados são revisados por todos os pesquisadores envolvidos e quaisquer desacordos são discutidos e resolvidos e, caso não haja consenso sobre determinado estudo, o mesmo é incluído.

O procedimento de seleção final consiste na cópia de todos os trabalhos incluídos como resultados da pesquisa inicial. Esses são revisados por no mínimo um dos pesquisadores, sendo avaliados quanto a sua qualidade através de uma lista. Essa revisão conclui a seleção de artigos aptos ao processo de extração de dados. Entretanto, caso seja encontrado mais de um trabalho relatando uma mesma abordagem, é selecionado aquele considerado mais completo e quaisquer desacordos sobre os trabalhos revisados conjuntamente são discutidos e resolvidos. Em casos que o acordo não seja alcançado, o artigo deve ser incluído.

Listas de verificação e procedimentos para avaliação da qualidade dos trabalhos

A qualidade dos trabalhos é avaliada através de um conjunto de questões conforme apresentado na Tabela 3.3. A pontuação para cada uma das questões foi atribuída de maneira distinta, entretanto o somatório total dos valores é no máximo de 15 pontos. Através de uma avaliação qualitativa da resposta seguindo uma escala, ela é convertida em um valor numérico.

A questão Q.1 de nível superior é calculada através da soma das suas respectivas sub-questões, Q.1.1, Q.1.2 e Q.1.3. As sub-questões são respondidas seguindo o seguinte padrão de valoração: Sim então o valor igual a 1 (um), valor é atribuído quando o trabalho descreve detalhadamente a informação desejada; Parcial então o valor igual a 0,5 (meio), valor é atribuído quando o trabalho descreve, parcialmente ou de maneira insuficiente, a informação desejada; Não então o valor igual a 0 (zero), valor é atribuído quando o trabalho não descreve a informação desejada.

² <http://jabref.sourceforge.net/>

As questões Q.2 e Q.3 não possuem nenhuma sub-questão, sendo respondidas seguindo o padrão de valoração apresentado: Para “Sim” então o valor igual a 1 (um), valor é atribuído quando o trabalho descreve a informação desejada; Para “Não” então o valor igual a 0 (zero), valor é atribuído quando o trabalho não descreve a informação desejada.

Tabela 3.3. Questões para avaliação da qualidade dos trabalhos.

Questão
Q.1. O trabalho descreve detalhadamente a abordagem proposta?
Q.1.1. O trabalho descreve as atividades que compõem a abordagem?
Q.1.2. O trabalho descreve os instrumentos utilizados pela abordagem?
Q.1.3. O trabalho descreve a ordem na qual as atividades são executadas?
Q.2. O trabalho descreve qual é o contexto para qual a abordagem pode ser aplicada?
Q.3. O trabalho descreve quais são o perfil técnico e o tipo de conhecimento necessário ao responsável pela aplicação da abordagem?
Q.4. O trabalho apresenta uma avaliação da abordagem de boa qualidade?
Q.4.1. O trabalho apresenta algum relato de uso ou estudo experimental que retrate a aplicação da abordagem?
Q.4.2. Quão bom é este relato de uso ou estudo experimental?
Q.5. O trabalho foi publicado em qual tipo de fonte?

O valor da questão Q.4 é calculado através da soma dos valores atribuídos para suas sub-questões, conforme Tabela 3.4. A sub-questão Q.4.1 é respondida através da seguinte classificação: “Estudo de caso”; “Experimento controlado”; “Relato de uso e prova de conceito”; e “Explanatório ou descritivo”. A sub-questão Q.4.2 é respondida seguindo a seguinte escala de valores: “Excelente”; “Bom”; “Regular”; “Pobre”; “Muito pobre”. A partir das respostas um valor da tabela é atribuído a questão Q.4.

Tabela 3.4. Tabela com valores a serem utilizados na Q.4.

		Q.4.2				
		Excelente	Bom	Regular	Pobre	Muito Pobre
Q.4.1	Estudo de caso	7	6	5	4	3
	Experimento controlado	6	5	4	3	2
	Relato de uso e prova de conceito	5	4	3	2	1
	Explanatório ou descritivo	4	3	2	1	0

A questão Q.5 é respondida seguindo o seguinte padrão de valoração: “Journal”, “Book” ou “Book chapter” recebe o valor igual a 3 (dois); “Conference” ou “Symposium” recebe o valor igual a 1,5 (um e meio); “Workshops” ou “Not indexed conference” ou “New conference” recebe o valor igual a 0,5 (meio).

Estratégia para extração de dados

Um artigo incluído pode conter mais informações do que a necessária para responder a questão de pesquisa, por isso deve se definir quais são os dados essenciais a serem extraídos. Essa extração também facilita a revisão e análise dessas informações. O primeiro conjunto diz respeito às informações para referência, contendo: Título, Autores, Fonte, Categoria. Outra informação importante é a descrição dos problemas na composição lidados pela abordagem.

Existe um conjunto de dados que serve para responder a questão primária:

- Tipo de fonte literária: teórico, experimental ou descritivo. “Teórico” é um artigo que apresenta somente a teoria, não apresentando um exemplo real de uso. “Experimental” é um artigo que apresenta um estudo experimental, associado a uma abordagem/processo descrita dentro deste artigo. “Descritivo” é um artigo que apresenta a descrição de uma abordagem/processo, apresentando um exemplo real de uso ou aplicação.
- Tecnologia de construção: Podem ser paradigmas de construção, a forma de estruturação da abordagem.
- Descrição do método ou abordagem para composição de *workflow* científico;
 - Fase do ciclo de vida: Utilizando a classificação proposta em (MATTOSO ET AL., 2010).
 - Instrumentos utilizados: Aparelho, objeto ou utensílio que serve para executar o procedimento/processo de concepção.
 - Mecanismos utilizados: Disposição das partes que constituem uma máquina, um aparelho. Também pode ser o manejo de um instrumento ou técnica.
 - Artefatos gerados e/ou utilizados: Definição adotada: Designação dada a qualquer objeto produzido ou consumido durante a execução do procedimento/processo de concepção.
 - Atividades executadas para a aplicação da abordagem ou método;
- Tipo de aplicação ou domínio para os quais o método ou abordagem é apropriado (em caso de discussão teórica) ou utilizado (em casos de artigos experimentais).
- Número de projetos já aplicados.

Existe um conjunto de dados para a questão secundária:

- Especificação da técnica de inspeção utilizada;

- Descrição da técnica de inspeção utilizada;
- Instrumentos utilizados na inspeção;
- Procedimentos da técnica utilizada;
- Tipo de aplicação ou domínio para os quais a inspeção é apropriada (em caso de discussão teórica) ou utilizada (em casos de artigos experimentais).

A estratégia de extração dos dados é aplicada por um pesquisador, assim identificando e extraindo as informações dos trabalhos selecionados. Todas as informações extraídas são documentadas em tabelas. Em uma segunda etapa, outro pesquisador seleciona uma amostra representativa das informações extraídas, revisando todos os dados. O pesquisador classifica a extração das informações como: I – insatisfatória; S – satisfatória. Ao final, o pesquisador pode solicitar a extração de todas as informações novamente, caso avalie como insatisfatória.

Resultados esperados (trabalhos de controle)

A Tabela 3.5 apresenta os artigos de controle que são utilizados nas execuções da busca da questão primária e secundária. Nesta tabela é apresentada a distribuição dos artigos pelas bases de buscas escolhidas e se são retornados na execução da string de busca primária e/ou secundária. Na execução da questão primária, todos os artigos de controle são esperados no retorno, além de atenderem aos critérios de inclusão/exclusão. Já na execução da questão secundária, o artigo de nº 4 não é esperado no retorno, pois não define formas de garantir a qualidade dos produtos gerados pela abordagem ou método proposto.

Tabela 3.5. Artigos de controle para a revisão *quasi* sistemática.

Nº	Artigos de controle	Scopus	IEEE Xplore	Compendex EI	Science Direct	String Primária	String Secundária
1	VERDI ET AL., 2007	Sim	Não se aplica	Não se aplica	Não se aplica	X	X
2	SORDE ET AL., 2007	Sim	Sim	Sim	Não se aplica	X	X
3	LIU ET AL., 2007	Sim	Não se aplica	Sim	Sim	X	X
4	LIU ET AL., 2006	Sim	Sim	Sim	Não se aplica	X	

PICO

A técnica PICO (PAI ET AL., 2004) é utilizada para auxiliar na definição da string de busca da pesquisa, oriunda do domínio da medicina. Esta técnica consiste na definição dos termos a ser utilizada para a população do estudo, a intervenção sendo

pesquisada, a comparação que pode ser utilizada para avaliar os resultados e os resultados esperados. Para esta pesquisa a população foi definida como: pesquisas ou trabalhos baseados em *workflow* científico, descritos na literatura técnica. Sendo os termos utilizados:

- *Scientific workflow, scientific experiment, science process, scientific process, workflow in e-science, e-science workflow.*

A intervenção definida para a questão principal da pesquisa foi: abordagens (ou métodos) para composição de *workflow* científico. Sendo os termos utilizados:

- *Composition, development, modeling, modeling, specification, design, definition, orchestration, conception, description, analysis, representation.*
- *Approach, method, technique, model, process.*

Já para questão secundária, a intervenção foi definida como: abordagens (ou métodos) para composição de *workflow* científico que ofereçam algum procedimento de verificação ou validação para garantia de qualidade de *workflows* científicos. Sendo os termos utilizados:

- *Composition, development, modeling, modeling, specification, design, definition, orchestration, conception, description, analysis, representation.*
- *Approach, method, technique, model, process.*
- *Inspection, review, verification, validation, reading, revise.*

Nesta pesquisa não foi definida nenhuma comparação. Os resultados foram definidos para a questão principal como: abordagens ou métodos para composição de *workflow* científico.

- *Composition, development, modeling, modeling, specification, design, definition, orchestration, conception, description, analysis, representation.*
- *Approach, method, technique, model, process.*

Enquanto para questão secundária, os resultados foram definidos como: Identificação dos procedimentos de verificação (revisão) e validação (inspeção) para garantia de qualidade de *workflow* científico.

- *Composition, development, modeling, modeling, specification, design, definition, orchestration, conception, description, analysis, representation.*
- *Approach, method, technique, model, process.*
- *Inspection, review, verification, validation, reading, revise.*

Uma string genérica foi gerada, contendo a combinação dos termos definidos para população do estudo, intervenção observada e os resultados esperados como retorno. A partir dessa string, as strings específicas foram definidas para cada base digital, pois cada uma possui mecanismo de busca diferente, interpretando de forma distinta a combinação. As *strings* derivadas para cada base não estão listadas nesta seção, pois são demasiadamente grandes, contudo estão disponíveis no anexo A deste documento.

String para questão primária: (scientific workflow OR scientific workflows OR scientific experiment OR scientific experiments OR science process OR science processes OR scientific process OR scientific processes OR workflow in e-science OR e-science workflow OR e-science workflows) **AND** ((development approach) OR (development method) OR (development technique) OR (development model) OR (development process) OR (composition approach) OR (composition method) OR (composition technique) OR (composition model) OR (composition process) OR (modeling approach) OR (modeling method) OR (modeling technique) OR (modeling model) OR (modeling process) OR (modelling approach) OR (modelling method) OR (modelling technique) OR (modelling model) OR (modelling process) OR (specification approach) OR (specification method) OR (specification technique) OR (specification model) OR (specification process) OR (design approach) OR (design method) OR (design technique) OR (design model) OR (design process) OR (definition approach) OR (definition method) OR (definition technique) OR (definition model) OR (definition process) OR (orchestration approach) OR (orchestration method) OR (orchestration technique) OR (orchestration model) OR (orchestration process) OR (conception approach) OR (conception method) OR (conception technique) OR (conception model) OR (conception process) OR (description approach) OR (description method) OR (description technique) OR (description model) OR (description process) OR (analysis approach) OR (analysis method) OR (analysis technique) OR (analysis model) OR (analysis process) OR (representation approach) OR (representation method) OR (representation technique) OR (representation model) OR (representation process))

String para questão secundária: ((scientific workflow OR scientific workflows OR scientific experiment OR scientific experiments OR science process OR science processes OR scientific process OR scientific processes OR workflow in e-science OR e-science workflow OR e-science workflows) **AND** ((development approach) OR (development method) OR (development technique) OR (development model) OR (development process) OR (composition approach) OR (composition method) OR (composition technique) OR (composition model) OR (composition process) OR

(modeling approach) OR (modeling method) OR (modeling technique) OR (modeling model) OR (modeling process) OR (modelling approach) OR (modelling method) OR (modelling technique) OR (modelling model) OR (modelling process) OR (specification approach) OR (specification method) OR (specification technique) OR (specification model) OR (specification process) OR (design approach) OR (design method) OR (design technique) OR (design model) OR (design process) OR (definition approach) OR (definition method) OR (definition technique) OR (definition model) OR (definition process) OR (orchestration approach) OR (orchestration method) OR (orchestration technique) OR (orchestration model) OR (orchestration process) OR (conception approach) OR (conception method) OR (conception technique) OR (conception model) OR (conception process) OR (description approach) OR (description method) OR (description technique) OR (description model) OR (description process) OR (analysis approach) OR (analysis method) OR (analysis technique) OR (analysis model) OR (analysis process) OR (representation approach) OR (representation method) OR (representation technique) OR (representation model) OR (representation process)))
AND (inspection OR review OR verification OR validation OR reading OR revise)

3.2 Execução da *quasi* Revisão Sistemática

3.2.1 Primeira rodada de execução – Maio 2009

Todas as referências recuperadas nas Bases de dados digitais foram obtidas através da execução das *strings* de busca. A busca foi realizada sem restrição de data inicial, até Maio de 2009 (data da execução das *strings* de busca). As referências recuperadas foram exportadas no formato RIS. Entretanto, as referências contidas nos anais de conferências nacionais (SBBD e e-Science), por estarem hospedados em uma base que não permite a execução de *strings* de busca complexas (acima de 255 caracteres), foram recuperadas integralmente para cada ano. Desta forma, a seleção inicial ficou a cargo do pesquisador, que leu todos os títulos e resumos das referências. As referências dos anais de conferências foram recuperadas na base BDBCOMP, sendo que a conferência SBBD foi pesquisada desde 1999 até a 2008, enquanto a conferência e-Science foi de 2007 e 2008.

A base de dados ACM não foi utilizada, pois a sua ferramenta de busca apresentava limitações no tamanho da *string* de busca (255 caracteres). Isto acarretaria na criação de diversas sub-*strings* de busca, tanto para questão primária quanto questão secundária, aumentando demasiadamente o esforço para o gerenciamento das referências retornadas. Somado a isso, a base digital Scopus indexa as referências presentes na base de dados ACM. Portanto, optou-se pela não

utilização da ACM Digital Library como fonte de dados e execução, pois a sua utilização não acrescentaria muito ao resultado. Contudo, deve-se destacar que sempre há possibilidade da não utilização da base de dados ACM representar a perda de alguma referência.

Todos os trabalhos recuperados das bases de dados digitais (Scopus, IEEE Xplore, Compendex EI e Science Direct) foram importados para a ferramenta de gerenciamento de referências JabRef (versão 2.5). Através dela, foi possível manusear, excluir as referências repetidas e categorizá-las. Para registrar os dados extraídos de cada referência, foram criados campos relacionados aos itens definidos no critério de extração de dados. Já para os trabalhos recuperados na base bdbcomp, os trabalhos foram importados e armazenados em um documento texto.

Durante o processo de exportação, problemas ocorreram com a base digital IEEE Xplore, pois essa não permite que seja exportada mais de uma referência por vez. Isto foi contornado através do uso de uma ferramenta de captura de referências, Zotero³. Assim, as referências foram importadas para sua base de dados e, posteriormente, exportadas para o formato RIS, e ao final importadas para o JabRef. As demais bases digitais permitiram a exportação das referências sem problemas, pois a quantidade de artigos retornados estava dentro dos parâmetros permitidos por elas.

As strings de busca foram aplicadas nas bases digitais, conforme descritas no protocolo de execução da *quasi*-revisão sistemática. A Tabela 3.6 apresenta a quantidade de referências recuperadas por base digital e anais de conferências (online). Devido à existência de duplicatas nas referências recuperadas, uma avaliação quanto à repetição foi necessária. As referências repetidas foram eliminadas, mantendo-se somente um registro. Todos os controles definidos foram recuperados nas bases digitais utilizadas.

A quantidade de referências mantidas após a eliminação das duplicatas está também representada na Tabela 3.6. As referências obtidas na base BDBCOMP não precisaram ser avaliadas quanto à repetição de artigos, pois se tratavam de trabalhos publicados nas conferências a cada ano, portanto não havia repetição. Além disso, se tratavam de artigos em Português não indexados nas bases digitais utilizadas.

³ <http://www.zotero.org/>

Tabela 3.6. Quantidade de referências mantidas sem duplicatas para questão primária.

Base de dados	Nº de recuperados	Nº de mantidos	% do total
Scopus	280	271	75,3 %
IEEE Xplore	104	76	21,1 %
Compendex EI	57	5	1,4 %
Science Direct	40	8	2,2 %
TOTAL	481	360	100 %
Anais de Conferências	Nº de recuperados	Nº de mantidos	% do total
SBBB	240	240	95 %
e-Science	12	12	5 %
TOTAL	252	252	100 %

Após a eliminação das referências repetidas, uma avaliação inicial das referências mantidas foi realizada através da leitura do título e abstract (resumo). Assim, algumas referências acabaram por ser excluídas, pois não se caracterizavam como pertinentes ao tema pesquisado. Após a exclusão, o resultado ficou como o representado na Tabela 3.7. A avaliação dos anais de conferência (SBBB e e-Science) foi realizada seguindo o mesmo procedimento e ao final o resultado obtido está representado na Tabela 3.8.

Tabela 3.7. Quantidade de referências base de dados após a primeira fase de avaliação para a questão primária.

Base de dados	Nº avaliados	Excluídos (E)	Não determinado (N)	Incluídos (I)	Selecionados inicialmente	% do total
Scopus	271	230	25	16	41	69,5 %
IEEE Xplore	76	60	14	2	16	27,1 %
Compendex EI	5	4	1	0	1	1,7 %
Science Direct	8	7	1	0	1	1,7 %
TOTAL	360	301	41	18	59	100 %

Tabela 3.8. Quantidade de referências nos anais após a primeira fase de avaliação para a questão primária.

Anais de Conferências	Nº avaliados	Excluídos (E)	Não determinado (N)	Incluídos (I)	Selecionados inicialmente
SBBB	240	240	0	0	0
e-Science	12	12	0	0	0
TOTAL	252	252	0	0	0

A próxima etapa na seleção foi realizada uma leitura mais detalhada dos trabalhos a serem incluídos. Assim, foram lidos todos os artigos classificados com N e

I na etapa de seleção inicial. A Tabela 3.9 representa sinteticamente a quantidade total de artigos incluídos e aptos à extração de suas informações. Das 59 referências inicialmente incluídas, somente 16 permitiram responder a questão primária, contudo duas referências não foram obtidas, pois não estavam disponíveis (KALETAS ET AL. 2003 e KALETAS ET AL. 2004). Assim totalizando 14 referências incluídas e aptas à extração. Contudo, duas referências se tratavam de publicações prévias de outras referências: ELLISON ET AL. (2006) era um trabalho prévio de BOOSIN ET AL. (2007), e LIU ET AL. (2006) era um trabalho prévio de LIU ET AL. (2007). Desta maneira, cada uma dessas referências foi reunida com a sua continuação e contada uma única vez, pois essas tratavam dos mesmos assuntos. Então, ao final ficaram somente 12 abordagens distintas.

Tabela 3.9. Quantidade total de referências incluídas na segunda fase de avaliação para questão primária.

Base de dados	A serem avaliados	Excluídos (E)	Não determinado (N)	Incluídos (I)	Referências selecionadas	% do total
Scopus	41	28	0	13	11*	78,57 %
IEEE Xplore	16	13	0	3	3	21,43 %
Compendex EI	1	1	0	0	0	0 %
Science Direct	1	1	0	0	0	0 %
TOTAL	59	43	0	16	12**	100 %
* Duas referências não foram obtidas, pois não estavam disponíveis (KALETAS ET AL. 2003; KALETAS ET AL. 2004).						
** Duas referências são trabalhos mais antigos de outros já incluídos (ELLISON ET AL. 2006; LIU ET AL. 2006).						

Analisando o total de referências recuperadas nas bases de dados digitais (481 no total), 25,16 % das referências se tratavam de repetições. Provavelmente, isto ocorreu devido às buscas serem realizadas em bases digitais que contém superposição em seu conteúdo digital. Após a eliminação das referências repetidas nas bases de dados digitais, restaram 360 referências distintas, das quais 301 foram eliminadas por não serem pertinentes à questão de busca (83,6%), mesmo que a maioria tratasse de alguma forma o tema “*workflow científico*”. Com isso foi realizada uma leitura detalhada dos artigos restantes, 59 referências, onde somente 16 (dezesseis) foram incluídas, porém 2 (duas) referências não estavam disponíveis para os pesquisadores, por falta de direito de acesso, sendo então eliminadas, depois do contato sem sucesso com os autores. Assim ao final, do total de 360 referências, somente 14 referências foram consideradas para extração, o que representa aproximadamente 4% das referências sem repetições retornadas após a busca. Contudo, deste total, somente 12 representavam abordagens distintas, o que representa aproximadamente 3% do total. Enquanto isso, ao se analisar os anais de

conferências pesquisados, a totalidade das referências foi excluída (252 referências), pois não foram encontradas referências pertinentes à questão de busca primária.

Para responder a questão secundária de pesquisa definida no protocolo, a string de busca secundária foi formada através da modificação da string de busca primária para assim retornar referências que contivessem informações sobre atividades de revisão. Essa string secundária de busca foi aplicada no mesmo conjunto de referências incluídas recuperadas das bases de dados digitais. Portanto, as referências recuperadas através da execução da string de busca secundária estão presentes no conjunto de referências recuperadas pela string de busca primária. A Tabela 3.10 representa a quantidade inicial de referências recuperadas. Foram retornadas 63 referências, e foram encontradas 15 repetições. Assim, o total final é de 48 referências. Houve uma repetição de 23,8% das referências retornadas.

Após leitura dos abstracts (resumos) e títulos, 40 das 48 referências foram excluídas, pois essas não eram pertinentes à questão de busca secundária. Esse resumo é apresentado na Tabela 3.11. A etapa seguinte consistiu numa leitura detalhada dos artigos incluídos na primeira rodada (8 referências) e desse total, somente 4 referências foram incluídas.

No entanto, vale ressaltar que após o início da leitura detalhada das referências selecionadas para a questão primária de pesquisa, foi percebido que a referência BOOSE ET AL. (2007) apresentava informações de procedimentos de revisão e inspeção, apesar de não estar incluída inicialmente nas 63 referências obtidas das bases digitais ao se aplicar a string secundária de busca. A referência não foi incluída automaticamente, pois esta não explicitava o uso de procedimentos de garantia de qualidade no seu resumo, título ou palavras chave, que são as seções avaliadas pelas bases digitais na comparação com os termos da string. Então, optou-se por incluí-la ao final como uma referência selecionada para a questão secundária de busca, passando-se assim para 9 referências a serem avaliadas. Afinal, essa faz parte das referências selecionadas para questão primária.

Tabela 3.10. Referências recuperadas para a questão secundária.

Base de dados	Nº de recuperadas	Nº de mantidas	% do total
Scopus	41	42*	86 %
IEEE Xplore	10	7	14 %
Compendex EI	7	0	0 %
Science Direct	5	0	0 %
TOTAL	63	49*	100 %
*a referência BOOSE ET AL. 2007 foi incluída manualmente no conjunto.			

O total de referências mantidas ao final e pertinente a questão secundária de busca foi de 5 (cinco) referências, representando um total de 7,8 % (de 49 referências, 48 retornadas pela string de busca e 1 incluída posteriormente), conforme a Tabela 3.12.

Tabela 3.11. Quantidade de referências após a primeira fase de avaliação da questão secundária.

Base de dados	A serem avaliados	Excluídos (E)	Não determinado (N)	Incluídos (I)	Selecionados inicialmente	% do total
Scopus	42	37	2	4*	6*	62,5 %
IEEE Xplore	7	4	2	1	3	37,5 %
Compendex Ei	0	0	0	0	0	0 %
Science Direct	0	0	0	0	0	0 %
TOTAL	49	40	4	5*	9*	100 %

*Referência BOOSE ET AL. 2007 foi incluída manualmente.

Tabela 3.12. Quantidade final de referências incluídas e aptas à extração da questão secundária.

Base de dados	A serem avaliados	Excluídos (E)	Não determinado (N)	Incluídos (I)	Referências selecionadas	% do total
Scopus	6*	2	0	4*	4*	80 %
IEEE Xplore	3	2	0	1	1	20 %
Compendex Ei	0	0	0	0	0	0 %
Science Direct	0	0	0	0	0	0 %
TOTAL	9*	4	0	5*	5*	100

*Considerando a referência de BOOSE ET AL. 2007.

3.2.2 Segunda rodada de execução – Outubro 2010

Com o objetivo de atualizar o corpo de conhecimento obtido através da *quasi* revisão sistemática, uma nova rodada de execução foi realizada nas bases digitais no mês de Outubro de 2010 (10/2010). As *strings* de busca foram aplicadas em todas as bases digitais novamente, tanto para string primária de busca quanto para secundária. Na nova rodada de execução não houve limitação de datas a fim de minimizar a perda de alguma referência importante, pois algumas bases digitais indexam referências com defasagem de tempo e de maneira retroativa.

Como não houve restrição de data inicial, as referências encontradas na primeira rodada de execução acabaram inseridas no conjunto retornado. Entretanto, essas não foram reanalisadas, sendo considerada a classificação anterior quanto aos critérios de inclusão e exclusão. A Tabela 3.13 apresenta a quantidade de referências retornadas para string primária após a execução. Nesta segunda rodada, já

desconsiderando as duplicatas, 720 referências foram encontradas, mas considerando somente novas, foram encontradas 360.

Tabela 3.13. Quantidade de referências recuperadas e mantidas sem duplicatas para String primária, Outubro 2010.

Base de dados	Nº de recuperados	Nº de mantidos	% do total
Scopus	393	381	53 %
IEEE Xplore	513	320	44,4 %
Compendex EI	121	6	0,8 %
Science Direct	54	13	1,8 %
TOTAL	1081	720	100 %

As 360 novas referências foram avaliadas quanto aos critérios de inclusão e exclusão, conforme definido no protocolo da *quasi*-revisão. Ao final, foram encontradas 6 referências que são aderentes à questão primária de pesquisa. Contudo, 3 dessas se tratam de publicações de continuação de pesquisas de abordagens já incluídas nesta *quasi*-revisão. A Tabela 3.14 apresenta a quantidade de referências incluídas para análise nas rodadas de avaliação.

Tabela 3.14. Quantidade de referências incluídas para extração de dados questão primária, Outubro 2010.

Base de dados	Etapa de inclusão 1		Etapa de inclusão 2		Referências finais	
	Selecionados	%	Selecionados	%	Selecionados	%
Scopus	110	30,5 %	29	32 %	5	83 %
IEEE Xplore	244	67 %	59	66 %	1	17 %
Compendex EI	1	0,5 %	1	1 %	0	0
Science Direct	5	1,5 %	1	1 %	0	0
TOTAL	360	100 %	90	100 %	6	100 %

Na execução da string secundária foram recuperadas 93 referências, incluindo duplicatas, logo desconsiderando as repetições foram encontradas 74. Essa quantidade está representada na Tabela 3.15. Contudo, nesse conjunto existem referências já avaliadas na primeira execução (em Maio de 2009). Sendo assim, excluíram-se esse conjunto, ficando com 25 novas a serem avaliadas. Então, foram aplicados os critérios de inclusão e como resultado foi encontrado que nenhuma nova referência era pertinente à questão secundária de pesquisa, conforme representado na Tabela 3.16.

Tabela 3.15. Quantidade de referências recuperadas e mantidas sem duplicatas para String secundária, Outubro 2010.

Base de dados	Nº de recuperados	Nº de mantidos	% do total
Scopus	56	17	23 %
IEEE Xplore	18	57	77 %
Compendex EI	12	0	0
Science Direct	7	0	0
TOTAL	93	74	100%

Tabela 3.16. Quantidade de referências incluídas para extração de dados questão secundária, Outubro 2010.

Base de dados	Etapa de inclusão 1		Etapa de inclusão 2		Referências finais	
	Selecionados	%	Selecionados	%	Selecionados	%
Scopus	15	60 %	0	0	0	0
IEEE Xplore	10	40 %	0	0	0	0
Compendex EI	0	0	0	0	0	0
Science Direct	0	0	0	0	0	0
TOTAL	25	100 %	0	0	0	0

3.2.3 Considerações sobre a execução da revisão

Antes de partir para discussão e análise dos dados extraídos das referências incluídas, é interessante realizar a análise e discussão sobre o percentual de representatividade de cada base digital dentro do conjunto de referências retornadas. A Tabela 3.17 apresenta um resumo desse percentual por etapa de inclusão para questão primária, enquanto na Tabela 3.18 o resumo é referente à questão secundária. Analisando-se as tabelas, nota-se que a base digital Scopus apresentou um alto percentual de representatividade (valores entre 60 % a 80%), isto é, em cada etapa, a base apresentou o maior número de referências que se encaixavam nos critérios de inclusão. Não obstante ao fato já discutido de haver sobreposição nas bases digitais, isto indica que em caso de limitação de tempo para realização de uma revisão sistemática, a base Scopus é uma forte candidata a uso, pois há probabilidade de retornar uma quantidade de referências aderentes à questão de pesquisa.

Apesar do esforço para desenvolver uma string de busca que capturasse o maior número de referências pertinentes à questão de pesquisa, uma não foi capturada, mesmo apresentando conteúdo relevante. GIL ET AL. (2007b) apresentam um processo para composição e foi indicada por um revisor de artigo submetido a congresso. Por ser relevante ao tema, foi incluída no conjunto de abordagens a ter

dados extraídos. Analisando o motivo pelo qual não foi capturado, identificou-se que essa referência utiliza um termo desconhecido aos pesquisadores e do revisor dos termos utilizados para formar as strings. O termo é “Computational *workflow*” que é sinônimo de “*Scientific workflow*”, assim, numa próxima atualização, este deve ser inserido na string de busca para capturar essa e outras relacionadas.

Tabela 3.17. Percentual de representatividade das bases digitais no conjunto de referências da questão primária, por etapa de inclusão.

Base de dados	Etapa de inclusão 1		Etapa de inclusão 2		Referências finais	
	Selecionados	%	Selecionados	%	Selecionados	%
Scopus	381	53 %	70	47 %	16	75 %
IEEE Xplore	320	44 %	75	50 %	4	25 %
Compendex EI	6	1 %	2	1,5 %	0	0
Science Direct	13	2 %	2	1,5 %	0	0
TOTAL	720	100 %	149	100 %	20	100 %

Tabela 3.18. Percentual de representatividade das bases digitais no conjunto de referências da questão secundária, por etapa de inclusão.

Base de dados	Etapa de inclusão 1		Etapa de inclusão 2		Referências finais	
	Selecionados	%	Selecionados	%	Selecionados	%
Scopus	56	60 %	5	62,5 %	4	80 %
IEEE Xplore	18	19 %	3	37,5 %	1	20 %
Compendex EI	12	13 %	0	0	0	0
Science Direct	7	8 %	0	0	0	0
TOTAL	93	100 %	8	100 %	5	100 %

A seguir, na Tabela 3.19, é apresentado o resumo das informações de referência de cada um dos trabalhos aptos à extração de dados. Também está marcado se este artigo foi incluído para a questão secundária. As quinze referências incluídas foram avaliadas através dos critérios de qualidade definidos no protocolo de *Quasi-revisão sistemática*, os resultados serão apresentados na seção seguinte. Para mais informações, no anexo B se encontra a lista de todos os artigos retornados após a execução da string de busca e sua classificação quanto à inclusão.

Tabela 3.19. Informações de referência dos artigos incluídos na questão primária e secundária.

#	Informações de referência	Questão Sec.
R1	(Boose et al. 2007) Ensuring reliable datasets for environmental models and forecasts. <i>ECOLOGICAL INFORMATICS 2 (2007) 237– 247.</i>	<input checked="" type="checkbox"/>
	(Ellison et al. 2006) Analytic webs support the synthesis of ecological data sets. <i>Ecology, 87(6), 2006, pp. 1345–1358.</i>	
	(Osterweil et al., 2008) Experience in using a process language to define scientific workflow and generate dataset provenance. <i>16th SIGSOFT 2008/FSE-16, 2008, 319-329</i>	
	(Osterweil et al., 2010) A. Clear and precise specification of ecological data management processes and dataset provenance <i>IEEE Trans. Autom. Sci. Eng., 2010, 7, 189-195</i>	
R2	(Bosin et al. 2007) Applying enterprise models to design cooperative scientific environments. <i>BPM 2005 Workshops, LNCS 3812, pp. 281-292, 2006.</i>	
	(Amigoni et al., 2007) A virtual laboratory for web and grid enabled scientific experiments <i>9th ICEIS 2007, 2007, SAIC, 227-230</i>	
R3	(Chard et al. 2008) Build Grid Enabled Scientific Workflows Using gRAVI and Taverna. <i>Fourth IEEE International Conference on eScience, 2008.</i>	
R4	(Gil et al. 2007a) Wings for Pegasus: Creating large-scale scientific applications using semantic representations of computational workflows. <i>Proc. XIX IAAI-07, July 22–26, 2007.</i>	<input checked="" type="checkbox"/>
	(Gil et al., 2007b) 2007. On the black art of designing computational workflows. <i>In Proc. 2nd Workshop on WORKS '07. ACM, New York, NY, 53-62.</i>	
R5	(Kaestle et al. 1999) Sharing experiences from scientific experiments. <i>Scientific and Statistical Database Management, 1999. Eleventh International Conference.</i>	
R6	(Li et al. 2006) Flows and views for scalable scientific process integration. <i>INFOSCALE '06. May 29-June 1 2006, Hong Kong.</i>	
R7	(Liu et al. 2006) The problem-based scientific workflow design and performance in Grid environments. <i>Proc. of the 5 Int. Conference on GCCW'06.</i>	<input checked="" type="checkbox"/>
	(Liu et al. 2007) On design, verification, and dynamic modification of the problem-based scientific workflow model. <i>Simul. Model. Pract. Theory, 2007.</i>	
R8	(Medeiros et al. 2005) WOODSS and the Web: Annotating and reusing scientific workflows. <i>Sigmod Record, Vol. 34, Nº. 3, Sept. 2005.</i>	
R9	(Sorde et al. 2007) Modeling and verifying non-DAG workflows for computational grids. <i>2007 IEEE Congress on Services, SERVICES 2007.</i>	<input checked="" type="checkbox"/>
R10	(Troguer and Fernandes 2004) A language for comprehensively supporting the in vitro experimental process in silico. <i>Proc. IV IEEE Symposium BIBE'04.</i>	
R11	(Verdi et al. 2007) Conceptual-level workflow modeling of scientific experiments using NMR as a case study. <i>BMC Bioinformatics 2007, 8:31.</i>	<input checked="" type="checkbox"/>
R12	(Yang et al. 2007) Formal modeling and analysis of scientific workflows using hierarchical state machines. <i>Third IEEE Int. Conference on e-Science and Grid Computing, 2007.</i>	<input checked="" type="checkbox"/>
R13	(Jablonski et a., 2007) A conceptual modeling and execution framework for process based scientific applications. <i>1st Workshop on CIMS '07, 16th ACM CIKM 2007, 2007, 23-30</i>	
R14	(Kalyanam et al., 2007) A Web Service-Enabled Distributed Workflow System for Scientific Data Processing. <i>FTDCS '07. 11th IEEE Int. Workshop, 2007, 7-14</i>	
R15	(Lacroix et al., 2009) Reasoning on scientific workflows. <i>SERVICES 2009 - 5th 2009 World Congress on Services, 2009, 306-313.</i>	

3.2.4 Avaliação da qualidade das referências incluídas:

As referências incluídas foram avaliadas conforme os critérios de qualidade definidos no protocolo da *quasi* revisão sistemática, conforme a Tabela 3.3 e citado anteriormente tem a pontuação máxima de 15 pontos. Já a Tabela 3.20 apresenta os valores obtidos por cada referência para essas questões.

Esta avaliação de qualidade das referências não tem caráter eliminatório, somente tem valor de verificação da importância da referência para as questões de pesquisa definidas. O total de pontos que uma referência pode alcançar é de 15 pontos. Assim, dentre todas as referências avaliadas, duas se destacam por apresentarem em seu conteúdo uma representatividade maior que 75%. Os trabalhos são R6 (LI ET AL. 2006) e R11 (VERDI ET AL. 2007), pois ambos apresentam não só os instrumentos e mecanismos utilizados, mas também uma descrição do procedimento para captura do *workflow* científico.

Um ponto importante, que merece destaque, é que todas as referências incluídas apresentam uma forte característica: a descrição dos instrumentos que apoiam sua abordagem. Todas as referências foram avaliadas com valor 1 (alto grau de atendimento) neste quesito. Isto pode ser um indício que os esforços da academia estão muito focados na apresentação de ferramentas e instrumentos, mas não de métodos ou na descrição da maneira como se deve utilizar tais tecnologias.

Outro ponto de destaque é a não utilização de estudos experimentais na avaliação da abordagem proposta na referência, o que diminui o grau de generalização e evidências do que foi proposto. A maioria dos artigos apresentou exemplos ou relatos de uso como avaliação da abordagem proposta. Mesmo assim, a maioria das referências teve seus exemplos caracterizados como regulares ou menos (2 pontos ou menos), com exceção de R11 (VERDI ET AL. 2007). Isto pode indicar uma deficiência e uma fraca evidência sobre a viabilidade de utilização da abordagem apresentada pelos pesquisadores.

Tabela 3.20. Resultado da avaliação da qualidade das referências incluídas.

Trabalho	Q.1			Q.2	Q.3	Q.4		Q.5	Total
	Q.1.1	Q.1.2	Q.1.3			Q.4.1	Q.4.2		
R1	0,5	1	0,5	1	0	1	2	3	9
R2	0,5	1	0,5	1	1	1	1	3	9
R3	0,5	1	0,5	1	0	1	1	1,5	6,5
R4	1	1	1	1	1	1	2	1,5	9,5
R5	0	1	0	1	0	1	1	1,5	5,5
R6	1	1	1	1	0	1	2	1,5	11,5
R7	1	1	0,5	0	0	1	2	3	8,5
R8	0,5	1	0,5	1	0	1	2	3	9
R9	0,5	1	0	1	0	1	2	1,5	7
R10	0	1	0	1	0	1	2	1,5	6,5
R11	1	1	1	1	1	1	3	3	12
R12	0	1	0	1	0	1	1	1	5
R13	0	1	0	1	0	1	2	1,5	6,5
R14	0,5	1	0,5	1	0	1	2	1,5	7,5
R15	0,5	1	0,5	1	0	1	2	1,5	7,5

3.3 Resumo das abordagens

Nesta seção é apresentada a síntese das informações extraídas para cada uma das referências. A tabela completa é apresentada no anexo C, que inclui o trecho extraído da própria referência, além da síntese.

R1 - ELLISON ET AL. 2006, BOOSE ET AL. 2007, OSTERWEIL ET AL. 2008 e OSTERWEIL ET AL. 2010:

ELLISON ET AL. (2007) relatam que os recentes esforços no desenvolvimento de sistemas não apresentam um bom apoio a proveniência dos dados gerados (datasets), prejudicando a sua confiabilidade. A solução, portanto, seria capturar não só os dados, mas também o processo experimental que os geraram. Assim, eles propõem um sistema que suporta as fases de composição, execução e análise de um *workflow* para o processo, denominado SciWalker. ELLISON ET AL. relatam somente o uso da ferramenta no desenvolvimento de um *workflow* no domínio de Biologia, contudo, menciona que essa foi utilizada em outros projetos de pesquisa.

Durante uso do SciWalker, três diagramas são criados para representar o *workflow*. Um contém os fluxos das atividades em alto nível, outro contém os caminhos de execução do *workflow* alocado aos recursos computacionais e o último representa a proveniência dos dados gerados e das atividades executadas. ELLISON ET AL também propõem uma forma de garantir a qualidade através da verificação dos possíveis caminhos que nunca irão ocorrer *workflow*.

R.2 - BOSIN ET AL. 2007 e AMIGONI ET AL. 2007

BOSIN ET AL. (2007) apresentam uma abordagem que lida com o problema da disponibilidade de um ambiente virtual cooperativo para realizar a ciência. Eles alegam que há uma falta de apoio à integração de aplicações difundidas entre organizações e grupos locais ou distribuídos. As aplicações são utilizadas na execução dos experimentos e a falta de integração destas pode torná-la mais difícil. Considerando o contexto, foi proposto um ambiente cooperativo (*Cooperative Framework*) que utiliza um paradigma de serviços para compor o *workflow*. BOSIN ET AL. não especificam qualquer restrição sobre domínio, assim como o número de projetos em que foi aplicada. O ambiente suporta a composição e execução, utilizando o conceito de níveis de abstração. São dois modelos definidos: *Workflow*, que representa o experimento cooperativo em alto nível; e Experimento cooperativo, gerado pelo

ambiente e que contém a definição dos serviços, os papéis envolvidos e a ordem de execução.

R.3 - CHARD ET AL. 2008

CHARD ET AL. (2008) definem que o *workflow* baseado em serviços *Web* requer competências técnicas elevadas dos pesquisadores, que devem adaptar códigos para utilizar estes serviços. Logo, a abordagem proposta por eles é uma solução para reduzir a complexidade desta tarefa através do uso de um conjunto de mecanismos que permitem a composição do *workflow*: Introduce / Gravi para o desenvolvimento e implementação de serviços *Web*, utilizando técnica de “encapsulamento” (*wrapping*) de aplicações; Taverna é o SGWfC integrado para modelagem e execução. Não há abstração do *workflow* na composição, pois o mesmo é concebido diretamente no nível concreto pela adição dos serviços embrulhados.

CHARD ET AL. não especificam qualquer restrição sobre os tipos de domínio que sua abordagem pode ser aplicada. Contudo, os *workflows* são orientados a serviços, o que pode significar restrições ao uso, como por exemplo, ambientes que utilizam muitos sistemas locais ou que o acesso a recursos de rede seja restrito. O exemplo de uso descrito é no domínio da bioinformática e não há relato do número de projetos no qual foi aplicada.

R.4 - GIL ET AL. 2007a e GIL ET AL. 2007b

GIL ET AL. (2007b) argumentam que os *workflows* estão aumentando em complexidade. A justificativa está no fato desses *workflows* serem difíceis de manter e modificar. Adicionalmente, há também um aumento no custo de composição, que é assumidos pelos grupos de pesquisa e pode não haver um retorno imediato do investimento. Para solucionar estes problemas, eles propõem uma abordagem composta por um ambiente formado pela integração de componentes e o uso de um processo para composição.

O ambiente é composto por: *Wings* - para criação de *workflow*, descrito semanticamente por ontologias, permitindo a concepção de templates e instancias do *workflow*; e *Pegasus* - responsável por mapear o *workflow* (executável) para os recursos necessários na infraestrutura. Já o processo possui etapas como: a definição dos papéis envolvidos; a criação de um esboço; a geração do *workflow* como template, instância e executável; definição dos dados e metadados; e garantia da qualidade, através de execução do *workflow* com um conjunto de dados conhecido, para avaliar se o resultado está de acordo com o esperado.

GIL ET AL. não especificam qualquer tipo de restrição à abordagem sobre o seu uso em domínios diferentes do exemplo de uso, sendo a única restrição quanto à natureza do *workflow*, que deve ser tipo DAG (*directed acyclic graph*). O domínio, onde a abordagem foi usada, é a *earthquake science*, não sendo especificado o número de projetos em que a abordagem foi utilizada.

R.5 - KAESTLE ET AL. 1999

KAESTLE ET AL. (1999) afirmam que a colaboração entre times geograficamente distribuídos é um problema, pois atualmente os experimentos científicos são utilizados por cientistas em localidades distintas. Este cenário exige uma nova divisão ativa da informação, suporte a sua publicação e busca. Outro problema é o acesso aos dados criados, pois há uma preocupação com os conjuntos de dados e os tipos de informação que um cientista está interessado para que não haja manipulação indevida. A abordagem de KAESTLE ET AL., para resolver esses problemas, utiliza o conceito de documentos *hyper-media* para estruturar a solução, pois estão disponíveis na Web e contêm as informações sobre o proprietário, os experimentos, ferramentas, serviços e conjunto de dados utilizados e produzidos. Esta abordagem utiliza a integração de componentes como: Projeto ESP2Net que permite a colaboração entre os diferentes grupos de cientistas; e, OASIS que é utilizado para a definição dos experimentos científicos.

A abordagem lida com a composição, execução e análise, entretanto não há o conceito de *workflow* no nível abstrato. KAESTLE ET AL. não especifica qualquer tipo de restrição sobre o tipo de domínio que a abordagem poderia ser usada. Além disso, não é mencionado o número de projetos que a abordagem foi aplicada.

R.6 - LI ET AL. 2006

Segundo LI ET AL. (2006), a abordagem é motivada pela necessidade em atender as exigências do experimento *in silico* com a composição. Eles sustentam que a dificuldade em representar os requisitos do experimento aumentou e, além disso, a representação tradicional baseada apenas em fluxo de atividades não é apropriada. Então, LI ET AL. propõem que é necessário identificar e capturar os outros fluxos que atendam os requisitos, como: dados, controle de exceção, semânticas e segurança. Para atender esse contexto, LI ET AL. propõem uma abordagem que usa o conceito de *Web Semântica* e se concentra em experimentos científicos distribuídos entre as organizações que utilizam os seus serviços no fluxo de execução.

A abordagem utiliza o conceito de serviço *web*, pois todos os aplicativos são embrulhados como serviços. Assim, o *workflow* científico vai usar o serviço como meio

para evocar a aplicação. Para tal usam o ADOME-WFMS, que permite criar, publicar e executar suas visões do experimento, e um Editor de fluxo para especificar os serviços *web*. A visão do experimento é um conjunto estruturalmente correto de fluxos, sendo compostos por controle, semânticas, dados e exceção. LI ET AL. não especificam qualquer tipo de restrição quanto ao tipo de domínio, apresentando um estudo de caso é na Biologia e não especificam o número de projetos que a abordagem foi aplicada.

R.7 - LIU ET AL. 2006 E LIU ET AL. 2007

LIU ET AL. (2007) argumentam que poucos modelos e SGWfCs são adequados para a complexa modelagem científica, pois os *workflows* científicos têm características como: processo científico não pode ser totalmente formalizado a priori; *workflows* científicos podem mudar durante a execução; o cientista tem um papel importante. Segundo os autores, há a necessidade de propor uma perspectiva diferente das baseadas em fluxos pré-definidos, através da divisão do *workflow* em problemas, o que seria mais adequado à modelagem científica. Considerando o cenário, LIU ET AL. propõem um modelo de *workflow* científico baseado em problemas (PBSWM), no qual o *workflow* científico é decomposto em problemas, que é uma unidade básica, ao invés de atividades. Posteriormente, este modelo é executado em uma *engine*, de acordo com os recursos de rede (disponível em ambiente em grade), definidos nas descrições do *workflow*.

LIU ET AL. utilizam análise formal através de conjunto de algoritmos para identificar defeitos no modelo PBSWM. Estes são usados para verificar a corretude (solidez) e também garantir que modificações, durante a execução do modelo PBSWM, sigam os casos previstos na abordagem, evitando inserção de defeitos. Esta técnica é aplicável somente nos modelos PBSWM e dependem de apoio computacional do ambiente para sua utilização. O exemplo apresentado por LIU ET AL. é no domínio de previsão climática, mas não há descrição de restrição quanto seu uso em outros domínios e o número de projetos no qual foi utilizado. Uma restrição é que a abordagem precisa de recursos de ambientes em grade.

R.8 - MEDEIROS ET AL. 2005

MEDEIROS ET AL. (2005) defendem que os *workflows* são um bom paradigma de documentação, porque são capazes de manter o rastro de evolução e anotação sobre o experimento. Assim, MEDEIROS ET AL. propõem um *framework* para apoiar a tarefa através do uso do conceito de *Web* semântica para empacotamento (*wraps*) as aplicações utilizadas. O *framework* WOODSS permite a criação e execução dos

experimentos científicos através de *workflows* com níveis de abstração distintos e foi estendido para apoiar o uso de serviços *web* e realizar anotação semântica.

No *framework* WOODSS, o cientista define um conjunto de atividades e tipos de dados que serão utilizados na construção do *workflow* abstrato. Durante essa tarefa, o *workflow* abstrato é mapeado para o conjunto de agentes (recursos de infraestrutura) para a criação do *workflow* concreto, sendo ambos armazenados em um repositório. Posteriormente, os cientistas podem editar, consultar e anotar o *workflow* que é obtido a partir do repositório. Segundo MEDEIROS ET AL., a abordagem foi estendida para suportar outros domínios e foi aplicada no domínio agro ambiental. No entanto, o estudo de caso apresentado é no domínio da Bioinformática.

R.9 - SORDE ET AL. 2007

SORDE ET AL. (2007) afirmam que com o crescente uso de ambientes em grade na execução, a complexidade da concepção dos experimentos aumentou. Desta forma, novos meios para representar o fluxo devem ser criados, utilizando não só a forma seqüencial e paralela, mas também pontos de decisão e os padrões de iteração. Nesse contexto, SORDE ET AL. propõem uma abordagem que apoia a composição, com tarefas de verificação, e execução de *workflows*, mais especificamente na concepção no nível concreto, porque captura diretamente os programas num modelo *non-DAG*. Na abordagem, o cientista cria um *workflow*, como um grafo de *workflow non-DAG* e, depois, esse é mapeado em grafo, baseado em rede de Petri, no qual técnicas de análise rede de Petri podem ser aplicadas. A abordagem é usada para compor os *workflows* que são baseados na tecnologia de ambientes em grade (Grid).

SORDE ET AL. propõem um modelo que poderia ser utilizado para os diversos campos da ciência, como bioinformática, meteorologia, dentre outros. Porém, não especificam qualquer tipo de restrição sobre o domínio. O exemplo descrito por SORDE ET AL. é do *workflow* ENAN, no domínio da bioinformática.

R.10 - TROGER & FERNANDES 2004

Para TROGER & FERNANDES (2004), os experimentos *in silico* ainda não são compatíveis com a prática *in vitro*, pois a sua definição ainda não está de acordo com os métodos de investigação de um experimento *in vitro*. Então, TROGER & FERNANDES propõem uma forma de apoiar o biólogo na modelagem de experimentos *in silico* se aproximando dos métodos de investigação que eles estão familiarizados. Outro problema citado é o fato de ferramentas de bioinformática não suportarem a relação entre a fase de coleta de provas do experimento, a hipótese

definida e processo de validação. Esta tarefa é geralmente feita *off-line* e não são diretamente vinculadas com a coleta de provas.

TROGER & FERNANDES afirmam que as ferramentas não oferecem suporte à rastreabilidade dos conjuntos de dados utilizados, a execução de diferentes experimentos. Afirmando também que as ferramentas não suportam a evolução do experimento, ao longo do tempo. Isto parece estar relacionado com o aspecto de proveniência do experimento. Considerando os problemas citados, TROGER & FERNANDES propõem uma abordagem composta por uma infraestrutura e sua linguagem (ISXL) que agem como uma camada de alto nível. Essa camada fica acima da máquina de execução, no qual um compilador fornecido pela infraestrutura traduz os *workflows* para ela. A abordagem parece ser criada para experimentos *in silico* em Bioinformática e não foi especificado o número de projetos em que foi utilizada além deste exemplo.

R.11 - VERDI ET AL. 2007

VERDI ET AL. (2007) argumentam que os experimentos permanecem geralmente na mente do cientista e a representação em cada SGWfC, no nível concreto somente, não o expressa totalmente. O experimento pode ter variabilidade nos seus procedimentos e isto raramente é devidamente documentado. Adicionalmente, há uma falta de orientações sobre como criar um modelo de *workflow* em alto nível, mais próximo ao experimento.

A abordagem contempla três fases, e cada uma é composta por duas etapas: concepção e validação. A primeira fase, “Construir modelo de contexto” identifica o processo e sub-processo do experimento e oferece uma visão estrutural dos mesmos. A segunda fase, “Construção modelo SIPOC”, identifica as entradas e saídas (de dados) e como elas fluem em no experimento. A terceira, “Construir modelo de fluxo de controle”, que capta o fluxo e a ordenação das etapas do processo de experimentação. A abordagem utiliza técnica identificação de requisitos para conceber o *workflow*, durante cada reunião. Além disso, pode-se usar para cada fase, na etapa de validação, o *Walkthrough* (técnica de revisão), com os cientistas envolvidos no levantamento, para identificar defeitos nos ou entre modelos. Existem três papéis na abordagem: um especialista em modelagem de processos de negócios; especialista de domínio experimento e, facilitador participante. Não há especificação sobre qualquer tipo de restrição, mas a abordagem foi desenvolvida para um projeto no domínio da Bioinformática e o exemplo está no domínio da bioinformática (RMN).

R.12 - YANG ET AL. 2007

YANG ET AL. (2007) afirmam que a pesquisa científica tornou-se colaborativa, e como consequência, houve o envolvimento de cada vez mais grupos distribuídos geograficamente. Eles compartilham dados, arquivos e conhecimento. Devido esse contexto, o fluxo de informação tornou-se uma peça chave para a segurança. YANG ET AL. apresentam uma abordagem baseada no conceito de modelo formal, utilizando a notação de máquina do Estado Hierárquica (MSMSW). A abordagem estende esse conceito para propor a solução, lidando com a fase de composição (com verificação). Entretanto, a abordagem de modelagem é baseada na composição de um modelo composto por elementos ator. Para cada ator atômico (como acontece com as atividades) do *workflow* é criado e associado a um HSMSW.

A abordagem utiliza uma técnica para verificar formalmente os modelos de *workflow*, através de algoritmo descrito nesta referência. O procedimento de aplicação da técnica consiste basicamente na aplicação de um algoritmo para traduzir um ator para uma máquina de estado finito, depois é aplicada uma técnica de controle de fluxo de informações para verificar a correção do modelo. A técnica só é aplicável para modelo proposto pela abordagem.

R.13 – JABLONSKI ET AL., 2007

Para JABLONSKI ET AL. (2007), a complexidade do gerenciamento dos processos experimentais e dos dados aumentou devido ao uso de aplicações científicas mais complexas e necessidade de manipulação de grande volume de dados. Por isso, o *workflow* deveria ser representado em perspectivas distintas, como de controle ou de participantes, e não só de dados. Considerando esse contexto, eles propõem uma abordagem que utiliza três *frameworks* para realizar a composição e execução do *workflow*. O *framework Perspective Oriented Process Modeling* (POPM) é utilizado para capturar as perspectivas envolvidas nos *workflows*; o *framework Process Driven Architecture* (PDA) serve para apoiar a criação de uma infraestrutura de execução das aplicações envolvidas no *workflow*; enquanto, o *framework Data Logistic with Ontologies* (DaltOn) faz a ligação entre os recursos, fazendo as conversões necessárias

JABLONSKI ET AL. não relatam restrições explícitas sobre domínios de ciência no qual a abordagem possa ser utilizada e descrevem a sua aplicação no domínio de ecologia, no experimento de cálculo de média de chuva por ano. JABLONSKI ET AL. não citam o número de projetos em que aplicaram a abordagem além do exemplo.

R.14 – KALYANAM ET AL., 2007

KALYANAM ET AL. (2007) intencionam diminuir a complexidade da composição de *workflow* para experimentos científicos. Segundo eles, existem duas limitações nos SGWfC, que são a manipulação de código fonte e a integração das aplicações usadas na execução. Com o objetivo de solucionar estes problemas, eles apresentam uma abordagem baseada na integração de componentes como máquina de execução e interface de modelagem, criando um ambiente denominado de JOpera para que o *workflow* seja concebido e executado. A solução proposta por eles utiliza serviços *web* para embrulhar as aplicações e evitar a manipulação do código, e posteriormente estes serão utilizados no *workflow* em alto nível. O JOpera também é capaz de ligar e gerenciar os serviços à infraestrutura responsável pela sua execução.

KALYANAM ET AL. não descrevem restrições quanto ao domínio de ciência que sua abordagem pode ser utilizada e não descrevem também a quantidade exata de projetos em que a solução foi aplicada. KALYANAM ET AL. apresentam somente um exemplo de aplicação que foi no domínio meteorologia, na concepção de um *workflow* para previsão climática.

R.15 – LACROIX ET AL., 2009

LACROIX ET AL. (2009) afirmam que houve aumento da complexidade da composição por conta do grande número de atividades de avaliação de parâmetros e de simulação no *workflow*. Pois há um grande número de atividades desses tipos nos *workflows* atuais e, segundo os autores, deveria haver um ambiente que auxiliasse nessa tarefa.

LACROIX ET AL. propõem uma abordagem no qual primeiro é criado um *workflow*, depois esse é avaliado através de testes com dados de entrada já conhecidos e comparando os dados gerados com resultados reais obtidos no laboratório. Para realizar a composição utilizam o ambiente ProtoDB no qual geram o *workflow* através da busca semântica por atividades descritas em um repositório. Esse ambiente também pode executar o *workflow*, o que permite a avaliação com dados de entrada já conhecidos com os resultados reais.

LACROIX ET AL. não descrevem restrição ao domínio que sua abordagem pode ser usada e também não descrevem em quantos outros projetos já a aplicaram. Uma possível restrição está no fato que eles fazem verificações com experimentos *in vitro* para avaliar o experimento *in silico* (*workflow*) criado, o que pode não ser possível para todos os domínios. LACROIX ET AL. descrevem somente um exemplo, no domínio Medicina, na criação de um *workflow* para simular impacto do colesterol em células.

3.4 Análise dos resultados

O objetivo principal desta revisão é a caracterização das abordagens ou métodos utilizados para a composição de *workflow científico*, e se existem identificar se alguma utiliza técnica de garantia de qualidade e detecção de defeitos. Os objetivos foram sintetizados nas Tabela 3.1 e Tabela 3.2. A partir destes objetivos foram definidas questões de pesquisa que nortearam a busca, representadas na Tabela 3.21. Com base nos dados extraídos, busca-se responder as questões de pesquisa.

Tabela 3.21. Questões oriundas dos objetivos de pesquisa.

Questão primária:	Quais abordagens (ou métodos) são utilizadas para a composição de <i>workflows</i> científicos?
Questão secundária:	Quais destas abordagens (ou métodos) oferecem algum procedimento de garantia de qualidade, particularmente, para verificar (revisão) e validar dos <i>workflows</i> científicos?

Análise dos problemas na fase de composição relatados

Antes de analisar as características e indícios encontrados nos dados extraídos das quinze abordagens, cabe uma análise da sua distribuição ao longo dos anos. Considerando o ano de execução da busca por referências como 2010 (Outubro de 2010), então será considerado este o ano limite. A Figura 3.1 apresenta um gráfico de distribuição anual de referências sobre o tema composição de *workflow científico* e a quantidade de novas abordagens.

Logo, percebe-se que a quantidade de referências tratando de composição de *workflow científico* aumentou recentemente, considerando-se a época da execução da busca (ano de 2010). Existe uma referência datada no ano de 1998, depois houve um hiato até o ano de 2004, quando se iniciou um crescimento que atingiu o ápice no ano de 2006 e 2007. Isso indica que nesses dois anos houve um aumento no interesse neste tópico. Uma possível explicação para esse aumento de interesse é a publicação da série de grandes desafios para ciência da computação até o ano de 2020. Somente para listar alguns dos principais e seus respectivos anos: *UK Grand Challenges for Computer Research* (UKCRC, 2004); *Computational Science - ensuring America's Competitiveness*, 2005 (NITRD, 2005); *CRA Conference on Grand Research Challenges in Computer Science and Engineering* (CRA, 2005); *Grandes desafios da computação*, SBC (SBC, 2006).

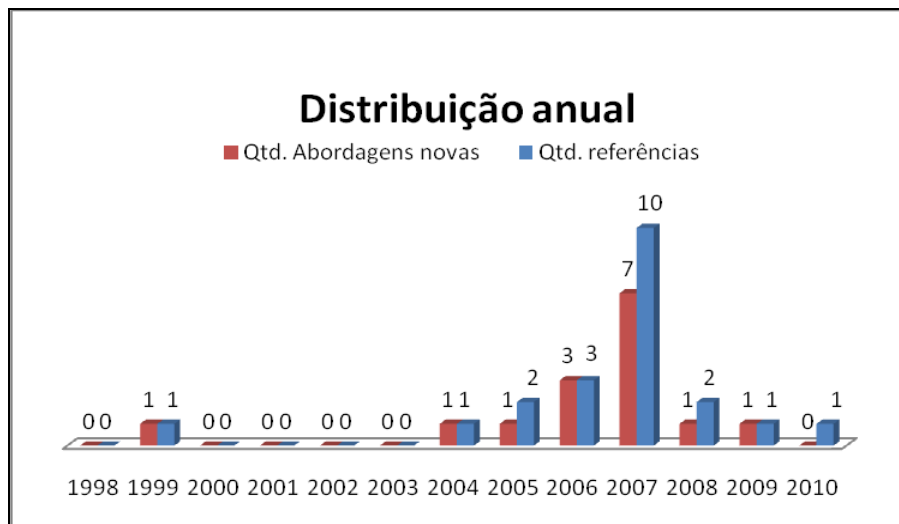


Figura 3.1. Distribuição anual de referências sobre composição e quantidade de abordagens novas.

Entretanto, a partir do ano 2008, há uma queda no número de publicações sobre composição e novas abordagens. As causas para tal queda não são claras, podendo ser desde a descontinuidade das pesquisas ou até mesmo longo tempo necessário para consolidação dessas tecnologias e abordagens. Porém é perceptível que as pesquisas não costumam executar experimentos controlados ou estudos de caso, pois, novas referências contendo avaliação dessas técnicas e/ou abordagens já descritas na literatura não são freqüentemente encontrados. Este cenário acaba por enfraquecer as próprias abordagens, pois não há relatos de uso em outros contextos diferentes e nem demonstra continuidade e evolução da pesquisa.

Análise dos problemas na fase de composição relatados

Inicialmente, uma forma de se caracterizar os esforços na área de experimentação apoiada por *workflows* científicos é entender quais são os principais problemas tratados com essas abordagens. Primeiramente, cada abordagem relata um determinado problema a sua maneira, portanto, como um meio de facilitar a discussão, estes foram sintetizados de forma a gerar categorias gerais que agregassem problemas relacionados entre si. A Tabela 3.22 apresenta as categorias gerais definidas a partir dos problemas extraídos das 15 referências, além da descrição do significado de cada uma dessas.

Tabela 3.22. Descrição e Categorias dos problemas identificados.

Categoria de problema	Descrição
Ambientes e <i>frameworks</i> deficientes	Deficiência dos ambientes ou <i>frameworks</i> que suportam a composição de <i>workflows</i> científicos. Estes ambientes deveriam permitir a colaboração entre grupos de cientistas e facilidades no uso de aplicações distribuídas.
Auxílio ao empacotamento (proveniência) deficiente	Atualmente, as abordagens para composição não suportam de maneira eficiente a proveniência, isto é o rastro, dos <i>workflows</i> científicos gerados e os dados produzidos através da sua execução.
Colaboração entre equipes distribuídas deficiente	As abordagens atuais não prevêm ou consideram a colaboração entre equipes geograficamente distribuídas, durante a composição e execução de <i>workflows</i> científicos. Isto pode acarretar em problemas diversos, tais como: falha na execução, violações de acesso, redução na confiabilidade dos dados produzidos.
Complexidade na composição de <i>workflow</i> incremental	A composição está se tornando uma tarefa complexa devido ao aumento no número e tamanho de informações a serem manipuladas, tais como: atividades, ferramentas, insumos e produtos, etc.
Documentação deficiente	As abordagens atuais para composição não geram uma documentação detalhada que suporte a experimentação. Em geral estas documentações são falhas no sentido de permitir uma replicação destes estudos/experimentos por outros cientistas ou grupos de pesquisa.
Garantia da qualidade deficiente	As abordagens atuais não se preocupam com aspectos de qualidade do <i>workflow</i> científico gerado e os dados que serão utilizados. Existe uma falta de preocupação com tarefas de validação e verificação do <i>workflow</i> .
Método para composição inexistente ou deficiente	Não existe uma abordagem baseada em métodos, com definição do <i>workflow</i> científico baseado nos requisitos necessários para sua realização. Falta definição do que deve ser identificado em termos de: atividades, ferramentas, papéis e restrições.

Com as categorias de problemas gerais definidos, a Tabela 3.23 foi criada, com o intuito de se obter uma visão da distribuição mais detalhada desses problemas pelas abordagens. Nessa tabela estão representadas as categorias de problemas gerais e quais são os problemas tratados/relatados por cada uma das abordagens.

Analisando-se mais detalhadamente a distribuição dos problemas pelas abordagens, como apresentado na Tabela 3.23, percebe-se que duas categorias de problemas foram mais exploradas: (i) Auxílio ao empacotamento (proveniência) deficiente, com três citações, e (ii) Complexidade na composição de *workflow* incremental, com seis citações. Para obter um panorama mais detalhado, essas duas categorias são avaliadas separadamente abaixo.

Tabela 3.23. Problemas tratados pelas abordagens.

Categoria de problema	Referências														
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15
Ambientes e frameworks deficientes		X					X								
Auxílio ao empacotamento (proveniência) deficiente	X				X					X					
Colaboração entre equipes distribuídas deficiente					X							X			
Complexidade na composição de <i>workflow</i> incremental			X	X					X				X	X	X
Documentação deficiente								X			X				
Garantia da qualidade deficiente											X				
Método para composição inexistente ou deficiente						X					X				
Modelos de representação deficientes							X			X					

- i. *Complexidade na composição de workflow incremental*: Este problema foi abordado pelas abordagens R.3, R.4, R.9, R.13, R.14 e R.15. Analisando-se as soluções propostas, nota-se que todas apresentaram, de uma maneira geral, infraestruturas e modelos que permitissem ao cientista (ou modelador) modelar seu experimento aumentando o nível de abstração do modelo. Todas as 6 abordagens seguem a linha de um modelo de mais alto nível (por exemplo, escondendo detalhes de chamadas a recursos computacionais, como *web services*) derivadas, em um segundo momento, para uma infraestrutura como *grids*.
- ii. *Auxílio ao empacotamento (proveniência) deficiente*: Este problema foi tratado pelas abordagens R.1, R.5 e R.10. Em R.1, argumenta-se que os “datasets” gerados (resultados ou dados) não são confiáveis, pois não há ligação e registro do processo experimental que os produzem. Assim, propõe uma infraestrutura que suporte a captura e relacionamento entre esses. Novamente, mas agora em R.5, aborda-se o problema de manter os tipos e resultados ou dados ligados a uma descrição do processo

experimental e para tal é proposta uma nova linguagem, que pode ser compilada e executada, capaz de resolver tal problema. Por último, em R.10, também se propõe uma nova linguagem e infraestrutura para se executar, a fim de tratar o problema da rastreabilidade entre processo experimental e dado gerados. Em comum entre as três abordagens é que as soluções apresentadas envolvem a utilização de uma nova infraestrutura para composição, execução e armazenamento destas informações. A escolha por tais soluções pode ser explicada pelo fato de que a manutenção da rastreabilidade entre artefatos (no caso, processo experimental e dado gerados) é de difícil gestão, já que a quantidade de ligações pode ser grande e incremental.

Interessante notar que os problemas mais citados são resolvidos através da criação de infraestruturas e ferramentas computacionais. Mas ao olhar mais atentamente para a Tabela 3.20, percebe-se que quase todas as abordagens citadas (R.1, R.3, R.5, R.9, R.10, R.13, R.14 e R.15) não apresentam métodos ou processos bem descritos (todos abaixo de 0,5, numa escala de 0, 0,5 e 1). Isto é, as ferramentas são apresentadas como solução para um determinado problema, porém não é explicitado detalhadamente como e quando essa deve ser utilizada, assim como o quê o usuário deve se preocupar em capturar, identificar ou modelar.

Analisando-se os problemas menos citados, foi detectado que se trata da Garantia da qualidade (no caso a categoria é a Garantia da qualidade deficiente), área de pesquisa muito abordada na Engenharia de *Software* e Desenvolvimento de Sistemas. Na Engenharia de *Software*, o objetivo desta área é aumentar a qualidade dos produtos gerados nos processos de *Software*, de forma que estes apresentem menos defeitos e, conseqüentemente, menores custos de construção e correção. Assim, poder-se-ia supor que mais abordagens tratassem da garantia da qualidade, porém isto não é confirmado na Tabela 3.23. Mesmo havendo seis abordagens que definem procedimentos de garantia da qualidade, somente R.11 declara claramente sua preocupação com a qualidade dos *workflows* científicos gerados.

Análise da organização tecnológica de construção das abordagens

Após a caracterização do cenário de problemas, é tratada aqui a organização e tecnologias de construção das soluções propostas nas quinze abordagens. Uma primeira questão passível de análise é como estruturam suas soluções, isto é, se descrevem como organizam tecnologias existentes ou se constroem uma nova. Isso é importante para caracterizar se as soluções propostas constroem novas tecnologias,

ferramentas, sistemas ou se reutilizam uma dessas já criadas. Assim, para responder a isso foram extraídas informações e sintetizadas na Tabela 3.24.

Das quinze abordagens, somente quatro dessas (R.1, R.7, R.9 e R.11) não descreveram como a solução foi constituída, sendo representado como '-' na Tabela 3.24. Entretanto, as demais revelam algumas informações interessantes, tais como o uso de integração de componentes, arquitetura orientada a serviços, descrição semântica e modelos formais. Cada um desses itens é analisado a seguir.

A **integração de componentes** foi um princípio de construção da solução para sete abordagens (R.2, R.3, R.4, R.5, R.10, R13 e R14). Em geral, estas abordagens partem da união de tecnologias já existentes, como por exemplo, *engines* de execução, e então desenvolvem alguma nova tecnologia que agrega essas tecnologias. Contudo, não é somente realizada uma união de tecnologias, mas nos casos apresentados, as abordagens também desenvolveram alguma ferramenta que explorava ou unia todas as outras. Por exemplo, em R.4, foi proposto um ambiente de modelagem que permite a definição de *workflow* científico e depois sua instanciação para recursos mapeados em ambientes em grade (*grids*). Esta solução tem como objetivo tratar a dificuldade e complexidade de criação de *workflows* científicos concretos. Em alguns casos (R.3, R.4), essa integração foi motivada por se tratar da continuação do trabalho de um grupo de pesquisa, que detectou algum problema em seu trabalho anterior e desenvolveu uma nova tecnologia que suprisse tal necessidade. Enquanto nas demais abordagens, a integração de componentes também é feita para solucionar algum problema detectado, mas que não é necessariamente uma continuação de um trabalho de pesquisa anterior.

Arquitetura orientada a serviços também foi utilizada como paradigma na construção de algumas das abordagens propostas (R.2, R.3, R.6 e R.8). Uma arquitetura orientada a serviços permite que os recursos que serão utilizados na execução do *workflow* científico sejam encapsulados e utilizados como serviços web. Isto também permite uma utilização de diversos recursos computacionais distribuídos em diferentes locais e infraestruturas computacionais. Então, como estas abordagens focam no suporte a **Composição e Execução** de *workflows* científicos, se explica a utilização deste paradigma na construção de suas tecnologias.

Outra tecnologia de construção utilizada foi **descrição semântica** (R.4, R.8 e R15), que visa o suporte a descrição dos relacionamentos entre os diversos elementos que compõem o *workflow* científico. Por exemplo, tais descrições podem ser o relacionamento entre atividades do *workflow* científico, entre atividades e recursos computacionais, entre produtos gerados durante a execução e *workflows* científicos. Assim, ao se analisar as características das abordagens, percebe-se que ambas

propõem ambientes para composição de *workflow* científico e instanciação destes para determinadas infraestruturas e recursos computacionais, também permitindo recuperação através de busca semântica na base. Esta busca pode ser, por exemplo, através da comparação de qual recurso computacional suporta uma determinada atividade *X* com entradas *a* e *b*, descritas através de relações semânticas. Vale destacar que as abordagens tendem a lidar com nível concreto, pois são *workflows* científicos compostos por atividades executáveis e que representam somente uma possível combinação de recursos e atividades.

Tabela 3.24. Tecnologias de construção utilizadas por cada abordagem.

Ref.	Tecnologias de construção utilizadas	Categoria
R.1	Não foi possível identificar.	-
R.2	Possui paradigma para composição de experimentos que utiliza tarefas cooperativas, denominado de e-process , baseado em SOA – Service Oriented Architecture e e-service .	Arquitetura orientada a serviço
	O protótipo da abordagem é construído usando o conceito de multi-tier .	Integração de componentes
R.3	Possui um paradigma para composição de experimentos que usa Service-Oriented (É baseado em arquitetura orientada a serviço, Service-oriented architecture – SOA).	Arquitetura orientada a serviço
	A abordagem usa integração de componentes para representar a solução.	Integração de componentes
R.4	A abordagem usa integração de componentes para representar a solução.	Integração de componentes
	A abordagem usa o conceito de descrição semântica para assistir a composição de <i>workflows</i> . Utiliza ontologia para descrever os seus elementos.	Descrição semântica
R.5	Usa o conceito de documentos Hyper-media para estruturar a solução.	Modelos formais
	A abordagem usa a integração de componentes para representar a solução.	Integração de componentes
R.6	A abordagem usa o conceito de web services para construir seu ambiente. No qual as aplicações são disponibilizadas como <i>web services</i> .	Arquitetura orientada a serviço
R.7	Não foi possível identificar.	-
R.8	A abordagem usa o conceito de Semantic Web para encapsular a aplicação que é usada no <i>workflow</i> .	Arquitetura orientada a serviço & Descrição semântica
R.9	Não foi possível identificar.	-
R.10	De fato, a abordagem propõe uma infraestrutura e linguagem que age como uma camada em alto nível . Esta camada se mantém acima da engine de execução, compilando os <i>workflows</i> traduzidos para a infraestrutura.	Integração de componentes
R.11	Não há tecnologia de construção descrita no artigo. Mas, a abordagem é parte de um projeto maior, no qual o foco é a construção de um framework para suportar o processo experimental no contexto do experimento NMR.	-
R.12	A abordagem é baseada no conceito de modelo formal, Hierarchical State Machine .	Modelos formais
R.13	Utiliza a integração de componentes (frameworks) para apresentar sua solução.	Integração de componentes
R.14	Utiliza a integração de componentes para apresentar sua solução.	Integração de componentes
R.15	Descrição semântica do domínio. Descrição das atividades e os relacionamentos entre seus artefatos e ferramentas.	Arquitetura orientada a serviço & Descrição semântica

E, por último, um paradigma também utilizado foi o de **modelos formais**, que é a utilização de modelos baseados em formalismo, buscando vantagens como correção ou estruturação dos modelos. As abordagens (R.5 e R.12) são propostas de novos modelos de representação para *workflows* científicos. Contudo é interessante que uma possível justificativa para escolha deste paradigma pode ser o fato que ambas se preocupam com a colaboração entre equipes e recursos computacionais distribuídos. Desta forma, há uma preocupação maior com a correção dos modelos gerados, realizada através da garantia de modelos formais que possam ser verificados para evitar desperdícios ou erros.

Análise das características das atividades de concepção das abordagens

Novamente considerando os dados extraídos das quinze abordagens, agora é analisado como e quais são as características das atividades de concepção descritas nessas abordagens. Uma breve análise já foi realizada na seção sobre a avaliação da qualidade das abordagens, na qual foi apontado um indício de que estas tendem a descrever com detalhes os instrumentos utilizados em detrimento das atividades de concepção. Então, percebe-se que há uma preocupação em descrever a tecnologia, isto é, suas ferramentas, seus sistemas, incluindo sua organização, enquanto as atividades, tarefas ou procedimentos que explicam como essas tecnologias devem ser utilizadas, por quem e em que etapa do ciclo de vida, não são muito exploradas.

Porém, nesta seção o foco não está mais na qualidade dos relatos, mas sim no que foi relatado, tentando explorar e identificar tendências de soluções, assim como características que permitam responder a questão de pesquisa definida. Tendo em vista a análise das características, a Tabela 3.25 traz um resumo das informações extraídas de todas as abordagens.

Ao se analisar quais são as etapas do ciclo de vida do experimento suportado por *workflow* científico que estas abordagens contemplam, constata-se que a maioria das abordagens não só apoiam a fase de composição, mas também de Execução (citado por 80% dos trabalhos). Uma possível explicação para este fato pode ser a própria característica das soluções dadas, pois todas propõem modelos de representação novos e *engines* de execução, que são capazes de interpretar esses modelos, assim executando o *workflow* científico e simulando o experimento. Ou seja, o modelo é próprio para uma determinada *engine* ou máquina de execução, não podendo ser utilizado em outra. Este fator pode ter guiado a abordagem para que também apoiasse a fase de Execução.

Tabela 3.25. Características das abordagens identificadas.

Ref.	Etapas do ciclo de vida	Organização das abordagens	Papéis/ Responsável
R.1	Composição Execução e Análise.	Três atividades gerais: Composição; Execução; Análise	Cientista (caso biólogo).
R.2	Composição e Execução.	Três tarefas gerais: Definição do experimento; Decomposição em serviços; Disponibilização para uso.	Cientista chefe e Gerente usuário experiente.
R.3	Composição e Execução.	Três atividades gerais: Criação dos serviços; Geração do <i>workflow</i> concreto; Execução <i>workflow</i> concreto.	Não é descrito.
R.4	Composição e Execução.	Processo de composição: Criação do template do <i>workflow</i> científico; Criação da instância do <i>workflow</i> científico executável;	Não é descrito.
R.5	Composição Execução e Análise.	Não há atividades descritas, somente uma descrição geral da forma como conceber os documentos em SEML.	Não é descrito.
R.6	Composição e Execução.	Dois fases: Criação das visões de seus serviços; Composição de serviços.	Organizações e parceiros.
R.7	Composição e Execução.	Dois atividades gerais: Modelar o modelo de <i>workflow</i> baseado em problemas; Ligação aos recursos e execução do modelo.	Não é descrito.
R.8	Composição e Execução.	Dois fases: Concepção das atividades e mapeamento aos recursos computacionais; Busca e composição de novos experimentos.	Cientista.
R.9	Composição e Execução.	Dois tarefas gerais: Criação do <i>non-DAG workflow</i> ; Mapeamento em <i>WF-Petri</i> e Verificação.	Não é descrito.
R.10	Composição, Execução e Análise.	Dois fases: Definição do modelo de experimento em IXSL; Compilação e execução do modelo.	Não é descrito.
R.11	Composição.	Três fases, cada uma composta de duas etapas. Construção e verificação do modelo de contexto; Construção e verificação do modelo entrada e saída; Construção e verificação do fluxo de controle.	Modelador de processos de negócio; Cientista experiente no domínio; Facilitador.
R.12	Composição.	Não há descrição detalhada. Uma atividade identificada: Composição do modelo baseado em atores.	Não é descrito.
R.13	Composição e Execução.	Não é muito bem descrito. Mas primeiro é realizado uma modelagem alto nível, depois é instanciado para infraestrutura.	Não é descrito.
R.14	Composição e Execução.	Não são descritas em detalhes. Primeiro é construída a atividade (<i>service web</i>) para ser utilizada na concepção do <i>workflow</i> abstrato (alto nível). Posteriormente, essa atividade é ligada a um recurso computacional pelo JOpera.	Não é descrito.
R.15	Composição.	São duas fases. Projeto do <i>workflow</i> o experimento é expresso como <i>workflow</i> em alto nível. Análise do <i>workflow</i> : o experimento é expresso como <i>workflow</i> executável.	Não é descrito

Agora, avaliando-se a organização das abordagens, isto é, quais são os procedimentos ou as atividades para que sejam seguidos durante a composição, execução e/ou análise do *workflow* científico de um experimento, observa-se um

interessante aspecto. Além da pouca importância dada por estes artigos à descrição de métodos ou procedimentos, conforme explicado anteriormente, também se percebe que, em geral, são divididas em macro etapas correspondentes às fases do ciclo de vida que suporta. Somente duas abordagens não se caracterizam desta maneira, R.4 e R.11, pois nelas realmente são propostas e descritas processos para composição de *workflows* científicos, em nível abstrato. Outro ponto a destacar é a quantidade de abordagens que não explicitam quais são os papéis responsáveis pela composição do *workflow* científico, no caso das quinze (15) abordagens incluídas, dez (10) não apresentaram descrição dos papéis, suas responsabilidades e perfil técnico necessário. Isto é um indício ruim, pois a não definição ou descrição destes papéis pode acarretar em problemas e dificuldades na utilização da abordagem.

Análise das características das atividades de garantia da qualidade

Após avaliar as características das abordagens buscando caracterizar a forma como é realizada a concepção dos *workflows* científicos, os dados extraídos são analisados a fim de caracterizar e entender o cenário relacionado à sub-questão de pesquisa desta revisão. A sub-questão trata da caracterização dos procedimentos de garantia da qualidade, que por ventura as abordagens propostas utilizem durante sua execução. A Tabela 3.26 apresenta o resumo dos dados extraídos. Nesta tabela estão listados os procedimentos de garantia de qualidade encontrados e quais são as abordagens que os utilizam.

Apesar de cada abordagem denominar a atividade de inspeção e garantia de qualidade de maneira distinta, algumas delas puderam ser agrupadas na mesma categoria de procedimento. A seguir é discutida cada uma das categorias.

- A categoria “*Análise de caminhos de execução*” é a aplicação de técnicas para verificação de caminhos de execução, por exemplo, verificando a presença de caminhos que nunca serão executados ou convergem em *deadlocks*. Somente uma das abordagens citou a utilização deste procedimento.
- A categoria “*Análise estrutural do modelo*” é a aplicação de técnicas para verificação de possíveis problemas estruturais. A categoria ligada à verificação sintática, procurando capturar defeitos ligados à própria linguagem de modelagem ou padrão de representação. Três abordagens utilizam procedimentos desta natureza.
- A categoria “*Revisão dos modelos gerados*” é caracterizada pela aplicação de procedimentos de avaliação dos modelos gerados, mas não necessariamente em busca de defeitos estruturais. Também pode ser

executada a fim de encontrar defeitos de origem semântica. Duas abordagens aplicaram procedimentos dessa natureza.

Tabela 3.26. Categoria de procedimentos de verificação listados por cada referência.

	Referência:	R1	R4	R7	R9	R11	R12
Procedimento de garantia da qualidade	<i>Análise de caminhos de execução</i>	X					
	<i>Análise estrutural do modelo</i>			X	X		X
	<i>Revisão dos modelos gerados</i>		X			X	

É interessante notar que quatro das seis abordagens (R1, R7, R9 e R12) apresentaram explicitamente procedimentos de garantia da qualidade ligados à identificação de problemas de origem sintática, enquanto uma avaliação considerando também aspectos semânticos só foi citada por duas abordagens (R4 e R11). Um provável cenário é que, dependendo da abordagem aplicada, é possível que se obtenha um modelo de *workflow* estruturalmente correto, mas que semanticamente pode não estar. Também é interessante destacar que quatro (R1, R7, R9 e R12) dessas abordagens baseiam seus modelos em formalismo, o que pode explicar a escolha por procedimentos mais formais de verificação e avaliação.

3.5 Considerações finais do capítulo

A execução de uma *quasi*-revisão sistemática mostrou-se satisfatória, pois através dela foi possível identificar um conjunto de indícios interessantes a cerca do tema “composição de experimentos científicos baseados em *workflow* científico”. Um primeiro indício notado foi a existência de poucos trabalhos relacionados ao tema, depois de duas execuções do protocolo, em Maio de 2009 e Outubro de 2010. De um total de 720 referências distintas, retornadas pela string primária de busca, somente 20 foram selecionadas após a aplicação dos critérios de inclusão, o que representa um percentual de aproximadamente 4% do total, podendo ser considerado muito baixo. Uma possível explicação para este número de trabalhos encontrados pode estar relacionado ao interesse recente neste tema (composição de *workflows*). Destaca-se que a partir de meados de 2000 que a modelagem computacional de experimentos passou a figurar como um dos grandes desafios da computação. Importante ressaltar que a referência GIL ET AL. (2007b) foi incluída manualmente no conjunto, pois se tratava de importante fonte e não poderia ser ignorada. Além disso, representava uma

continuação de uma abordagem já capturada pela *quasi-revisão* (totalizando assim 21 referências).

Um indício interessante foi constatar que nenhuma das 21 referências encontradas apresentava algum estudo primário associado ou alguma forma de avaliação mais estruturada. Neste conjunto, a grande maioria dos trabalhos apresentava uma prova de conceito ou exemplo de uso dentro de um contexto bem específico, em geral, pouco descrito. Assim, considerando este cenário, pode-se afirmar que se torna difícil fazer qualquer tipo de generalização acerca da possibilidade de uso em outros contextos a não ser aqueles descritos. Também, não há indício sobre a continuidade da pesquisa sobre uma determinada abordagem, visto que do conjunto de 15 abordagens identificadas, somente quatro apresentaram duas ou mais referências retornadas após a execução das strings de busca. Isto pode indicar que esforços estão mal direcionados ou as abordagens falharam em solucionar os problemas propostos.

Ao se observar especificamente os dados retornados por essas quinze abordagens, percebe-se que todas as abordagens apresentam uma descrição detalhada das ferramentas ou sistemas que são apresentadas como solução para determinado problema. Entretanto, doze delas apresentavam uma descrição não muito detalhada de como a ferramenta deveria ser utilizada dentro da fase de composição, isto é, não apresentavam métodos ou processos descrevendo como um pesquisador ou usuário deveria utilizá-la. Este cenário indica que os esforços da academia estão muito focados na apresentação de ferramentas e instrumentos, ou seja, uma tentativa de automatização de uma solução. Contudo, a automatização é realizada sem o entendimento do processo no qual esta ferramenta está inserido, isto é, as abordagens não apresentam guias ou descrições de como se devem utilizar tais tecnologias dentro das atividades de composição de experimentos científicos suportados *workflows* científicos.

Neste contexto, um interessante aspecto diz respeito às atividades de garantia da qualidade de *workflows*. Do conjunto de quinze abordagens encontradas, foi identificado que somente seis possuíam informações sobre procedimentos para garantir a qualidade dos seus modelos de *workflow*. Contudo, somente uma abordagem deste conjunto explicitamente citava a questão da garantia da qualidade como problema a ser resolvido. Isto pode indicar que a garantia da qualidade ainda não é foco das pesquisas do domínio de *workflow* científico. Outro ponto diz respeito ao tipo de procedimento utilizado, que na sua maioria (três abordagens) têm o foco no aspecto sintático do modelo, isto é, na verificação/validação de que o modelo de *workflow* está estruturalmente correto com relação a um padrão. Mesmo sendo

importante, destaca-se que somente a correção sintática do modelo não garante a qualidade, pois nestas abordagens ainda é possível construir um *workflow* sintaticamente correto, mas semanticamente equivocado.

Uma possível ameaça à validade deste estudo está relacionada à completude da busca, isto é, se a *quasi*-revisão sistemática realmente capturou as referências que tratavam sobre o tema composição de *workflows* científicos. A fim de mitigar essa possibilidade, foram utilizadas referências de controle, de onde foram extraídos os principais termos utilizados para busca e, após isso, foi solicitado ajuda de um especialista no domínio de *workflow* científico para validação destes termos. Mesmo assim uma referência não foi capturada pela string de busca, pois utilizava um termo não conhecido pelos pesquisadores. Isto pode ser explicado pelo fato que por ainda se tratar de uma área de pesquisa recente, a nomenclatura do domínio ainda não é única, podendo apresentar variações de grupo para grupo de pesquisa. A não utilização da Base de dados ACM também pode caracterizar uma ameaça à validade, já que alguma referência pode não ter sido identificada. Contudo a base não pode ser utilizada, devido a limitações de execução da string de busca, optando-se assim por aceitar este risco.

Por último, esta *quasi* revisão sistemática auxiliou na observação de um aspecto interessante da área de *workflow* científico, em especial na fase de composição, que é a falta de uma agenda de pesquisa. A Tabela 3.23 apresenta uma distribuição esparsa, nela se percebe que cada trabalho aborda um problema distinto, e alguns casos mais de um trabalho trata de um mesmo problema. A falta de direcionamento de pesquisa em longo prazo pode acarretar em redundância de soluções para um mesmo problema, possibilidade de haver problemas que nunca serão tratados ou debatidos, dentre outros. Isto é preocupante, pois sem organização não há como conhecer a real necessidade desta área e o que de fato necessita de esforço de pesquisa.

CAPÍTULO 4 - ABORDAGEM PARA CONCEPÇÃO DE *WORKFLOW* CIENTÍFICO ABSTRATO

Neste capítulo é apresentada versão inicial da abordagem para concepção de workflow científico em nível abstrato. A abordagem foi composta inicialmente por um procedimento descrito em tarefas e papéis a serem desempenhados, utilizando instrumentos específicos desenvolvidos neste trabalho. Neste capítulo também é apresentada a primeira avaliação realizada que consistiu na aplicação da abordagem como uma prova de conceito no domínio de Evolução de Software (Engenharia de Software).

4.1 Introdução

Conforme revisado no Capítulo 3, a fase de Composição, mais especificamente a Concepção de *workflow* científico para experimentos *in virtuo* e *in silico*, ainda apresenta problemas como: (i) necessidade de documentação e representação em níveis mais altos de abstração; (ii) métodos que auxiliem na concepção do *workflow* científico. Assim no primeiro semestre de 2009, uma abordagem foi desenvolvida para a concepção de *workflows* científicos, considerando esses problemas (i) e (ii). A organização geral da abordagem é apresentada na Figura 4.1, sendo inicialmente composta por duas partes: procedimento para concepção e instrumentos para captura do *workflow*.

O procedimento para concepção apresenta tarefas relacionadas à identificação, documentação, verificação e avaliação de *workflow* científico em nível abstrato. As tarefas são realizadas em ordem pré-definida e estão organizadas basicamente em duas fases: especificação e validação. Para a execução das tarefas, os cientistas devem utilizar técnicas oriundas da Engenharia de *Software* para identificação de requisitos de sistemas e inspeção de documentação de projeto. Como parte da abordagem, os papéis essenciais para a aplicação do procedimento foram definidos, descrevendo as responsabilidades e o conhecimento necessário para seu desempenho.

Durante a aplicação do procedimento para concepção, os instrumentos devem ser usados para capturar os requisitos do *workflow* graficamente, através de modelos, e textualmente, através de formulários. Estes instrumentos foram criados

especificamente para apoiar as atividades do procedimento para concepção. Os instrumentos devidamente preenchidos formam um documento denominado Especificação de *workflow* científico.

Para representar os modelos, usa-se a extensão da notação do Diagrama de Atividades, UML 2 (OMG, 2009), e esta foi especializada para representar elementos como atividade manual, automatizada e sub-*workflow* (identificados na seção 2.2). As descrições textuais são representadas através de formulários que são compostos por um conjunto pré-definido de dados que caracterizam os elementos de modelos de *workflow* científico, sendo estes de três possíveis tipos: atividade, ferramenta ou artefato. A UML 2 foi escolhida para uso na abordagem proposta nesta dissertação, pois se trata de uma linguagem de modelagem comumente utilizada na indústria de desenvolvimento de *software* em seus projetos. Outro motivo para esta escolha está associado ao fato que essa linguagem é independente de sistema, máquina de execução ou linguagem de programação.

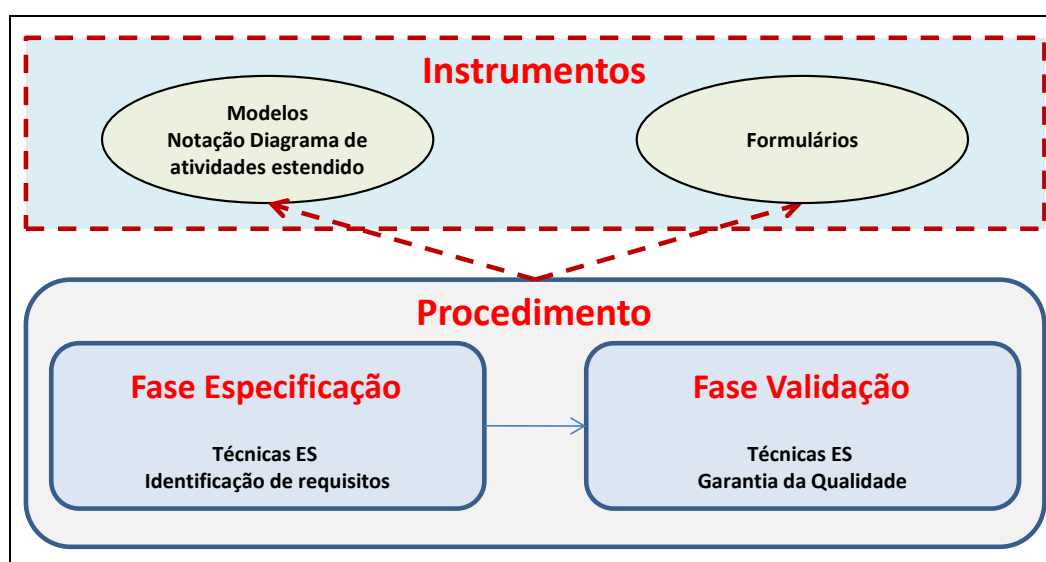


Figura 4.1. Abordagem para concepção de *workflow* científico em nível abstrato na versão inicial.

Nas próximas seções são descritos com detalhes as tarefas do procedimento e a estrutura dos instrumentos. Também é apresentada uma prova de conceito da abordagem. Este Capítulo organiza-se da seguinte forma: na seção 4.2 está a descrição do procedimento para composição de *workflow* científico abstrato; na seção 4.3 estão descritos os formulários para captura dos elementos do *workflow* científico e a extensão realizada na notação do diagrama de atividades; na seção 4.4 está descrita a prova de conceito efetuada no domínio de evolução de *software*, que teve como intuito de avaliar a aplicabilidade da abordagem e identificar oportunidades de melhorias; na seção 4.5 estão as considerações finais do capítulo.

4.2 Descrição do procedimento de concepção de *workflow* abstrato

No procedimento de concepção existem quatro papéis definidos, que devem ser desempenhados durante a execução das tarefas. Contudo não é obrigatório que esses papéis sejam realizados por indivíduos distintos. Sendo assim, um mesmo indivíduo pode desempenhar mais de um papel. Na distribuição de responsabilidades, devem-se levar em conta os recursos humanos disponíveis dentro do grupo, assim como os conhecimentos desejáveis para a realização de um papel. Os quatro papéis são:

- **Analista:** responsável por conduzir as reuniões com o objetivo de especificar o *workflow* científico, isto é, criar os modelos do *workflow* representados através de diagramas de atividades e preencher os formulários pré-definidos pela abordagem. Responsável também por conduzir a inspeção sobre esses artefatos e sua validação. O indivíduo que desempenha esse papel deve possuir conhecimento em modelagem de diagramas de atividade da UML 2 e na concepção de *workflow* científico.
- **Especialista do domínio:** responsável por fornecer as informações necessárias ao analista para que este possa executar a especificação. É desejável que o indivíduo desempenhando esse papel tenha conhecimento na notação de diagrama de atividades.
- **Inspetor:** responsável por realizar uma inspeção da Especificação do *workflow* científico, relatando defeitos e problemas nesse documento. Sua principal função é encontrar defeitos na especificação. É desejável que o indivíduo desempenhando esse papel tenha algum conhecimento na notação de diagrama de atividades.
- **Avaliador:** responsável por avaliar a Especificação do *workflow* científico, definido se esse documento realmente contém as informações necessárias para identificar o experimento e sua representação como *workflow* científico abstrato. Sua principal função é avaliar se a especificação representa o experimento científico.

Na Figura 4.2, o procedimento para concepção de *workflow* científico em nível abstrato está representado, seguindo a notação de diagrama de atividades. Esse é composto por duas fases: especificação, como “Especificar *workflow* científico”, e a validação, como “Validar *workflow* científico”. O principal artefato produzido após a execução satisfatória de ambas as fases é a “Especificação de *workflow* científico”,

documento contendo os modelos de *workflow* científico abstrato e os formulários de atividades, artefatos e ferramentas. Nas próximas subseções são detalhadas todas as fases e as tarefas que compõem o procedimento para concepção de *workflow* científico abstrato.

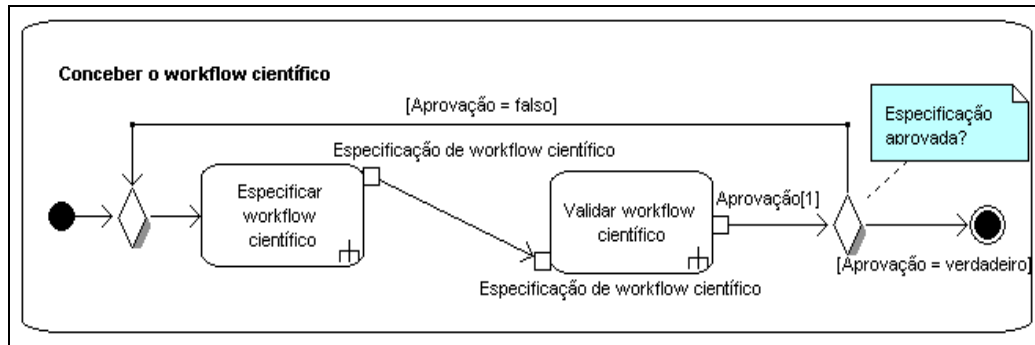


Figura 4.2. Método para concepção de experimentos baseados em *workflow* científico, em nível abstrato. Representado na notação de diagrama de atividade, UML 2.

4.2.1 Detalhamento da fase “Especificar *workflow* científico”

A primeira fase é “Especificar *workflow* científico”, o que corresponde à realização das seguintes tarefas: “Definir contexto do *workflow* científico” e “Identificar e modelar requisitos do *workflow* científico”. A Figura 4.3 apresenta o modelo que representa a ordem de execução dessas duas tarefas e os atores que as executam, sendo o Analista e os Especialistas do domínio.

Definir contexto do *workflow* científico

“Definir contexto do *workflow* científico”, primeira tarefa da Figura 4.3, é realizada através de uma reunião livre, cujo objetivo é o entendimento em alto nível do contexto e do experimento. Por alto nível, define-se o conjunto inicial de atividades do *workflow* científico, identificando um fluxo de atividades mais importantes, sem a preocupação em detalhar cada uma das atividades. Também deve se procurar identificar os papéis desempenhados pelos participantes do experimento científico, tentando já definir as responsabilidades. Considerando o objetivo dessa reunião, possíveis perguntas são:

- Papéis envolvidos: “Quem são os donos desses dados?”; “Quem executa essa tarefa/atividade?”.
- Atividades: “Qual é a primeira tarefa/atividade que vocês desempenham?”; “Esse dado é insumo de outra atividade?”.

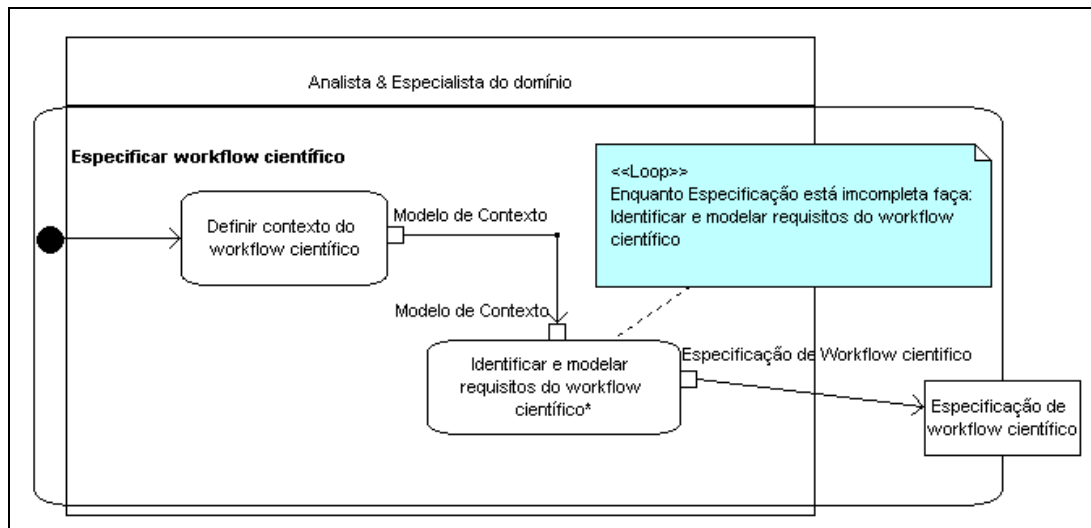


Figura 4.3. Detalhamento da fase “Especificar *workflow científico*”, composta por duas tarefas: “Definir modelo inicial de *workflow científico*” e “Identificar e modelar requisitos do *workflow científico*”.

Vale destacar que essas perguntas representam exemplos; outras questões podem ser utilizadas, ficando a critério do Analista. . Dois papéis estão envolvidos nesta tarefa: analista que é responsável pela concepção do modelo inicial; e o especialista do domínio, responsável por fornecer as informações sobre as atividades do experimento. A primeira reunião proporciona ao analista a oportunidade de se familiarizar com o domínio, pois pode desconhecê-lo, e também introduzir o especialista do domínio nos conceitos envolvidos no procedimento, como a notação. O artefato gerado é o “modelo de contexto” que é um esboço do *workflow*.

Identificar e modelar requisitos do *workflow científico*

A tarefa “Identificar e modelar *workflow científico*” (Figura 4.3) é realizada através de reuniões semi-estruturadas, com o propósito de capturar os requisitos do *workflow científico*. Neste contexto, os requisitos são as restrições, a ordem das atividades e as condições do ambiente para o experimento. A tarefa tem como insumo o “modelo de contexto”, que é o modelo base sendo refinado até se tornar aquele que será o *workflow científico* abstrato.

A reunião conduzida é semi-estruturada, seguindo um roteiro pré-definido. O primeiro passo é a identificação de novos elementos do *workflow científico* abstrato, usando como base o modelo de contexto. Para isso, utiliza-se a técnica de entrevista semi-estruturada (LAUESEN, 2002), o Analista prepara perguntas sobre o domínio e da execução das atividades do experimento. O segundo passo é o preenchimento dos formulários (atividade, artefato e ferramenta) com as informações identificadas sobre os elementos do *workflow*. Através do preenchimento dos formulários, novos

elementos no *workflow* podem ser indiretamente identificados, pois os três formulários contêm campos que se relacionam entre si, como por exemplo, cada atividade lista o conjunto de artefatos.

A técnica de entrevista semi-estruturada é utilizada, pois existem indícios que esta apresenta resultados melhores quando comparada com outras utilizadas no contexto da identificação de requisitos, na Engenharia de *Software* (DAVIS ET AL., 2006). A entrevista é semi-estruturada porque as perguntas empregadas são geradas a partir das descrições textuais dos campos presentes nos formulários, contudo o analista pode, a critério próprio, criar suas perguntas, não se limitando às informações dos formulários.

Nessa tarefa, dois papéis estão envolvidos: um deles é o analista, responsável por conceber o *workflow* científico abstrato e o outro, o especialista do domínio, que provê as informações necessárias para o analista. O artefato final gerado é a “Especificação de *workflow* científico” em nível abstrato; esse documento é composto por: modelos do *workflow* – representados na notação de diagrama de atividades, e formulários – tabelas de texto livre de três tipos atividades, resultados e ferramentas.

4.2.2 Detalhamento da fase “Validar *workflow* científico”

A segunda fase, “Validar *workflow* científico”, corresponde à verificação e avaliação do artefato gerado em “Especificar *workflow* científico”. É composta por quatro tarefas: “Inspeccionar Especificação”, “Avaliar discrepâncias relatadas”, “Corrigir especificação” e “Avaliar especificação”. A Figura 4.4 apresenta o modelo que representa a “Validar *workflow* científico”, estando descritos os papéis envolvidos, a ordem de execução das tarefas e os artefatos transmitidos entre elas.

Nas práticas de garantia da qualidade aplicadas na Engenharia de *Software*, a inspeção é um meio eficaz de detecção de defeitos, pois permite que eles sejam encontrados mais cedo no processo de desenvolvimento, evitando sua propagação para as fases posteriores e, por conseqüência, reduzindo o custo da correção (MAFRA & TRAVASSOS, 2005). Considerando o princípio de que todo artefato pode conter defeitos, foi detectada a necessidade de realizar atividades de garantia da qualidade na especificação criada. Então, para evitar que possíveis defeitos permaneçam e sejam propagados para as próximas fases da composição, realiza-se a inspeção da especificação de *workflow* científico.

Antes de descrever estas tarefas, os termos defeito e discrepância devem ser definidos, pois estes são necessários para o entendimento desta seção. Na Engenharia de *Software*, o termo defeito é definido como uma manifestação concreta de um erro em um artefato de *software* (IEEE 610.12, 1990). Assim, neste

procedimento, quando encontrados, os defeitos estão sempre presentes no principal artefato, que é a “Especificação de *workflow* científico”.

Entretanto, durante as tarefas de verificação ou avaliação, os participantes relatam supostos defeitos identificados na especificação. Estes relatos são denominados de discrepâncias. Destaca-se que as discrepâncias podem ser classificadas como defeitos reais ou não na especificação, sendo que quando uma destas não representa um defeito no artefato, denomina-se falso positivo. Uma discrepância é considerada defeito quando os avaliadores avaliam, discutem e chegam a um consenso sobre essa condição (SAUER ET. AL., 2000). A seguir é descrito como as tarefas de garantia da qualidade são realizadas.

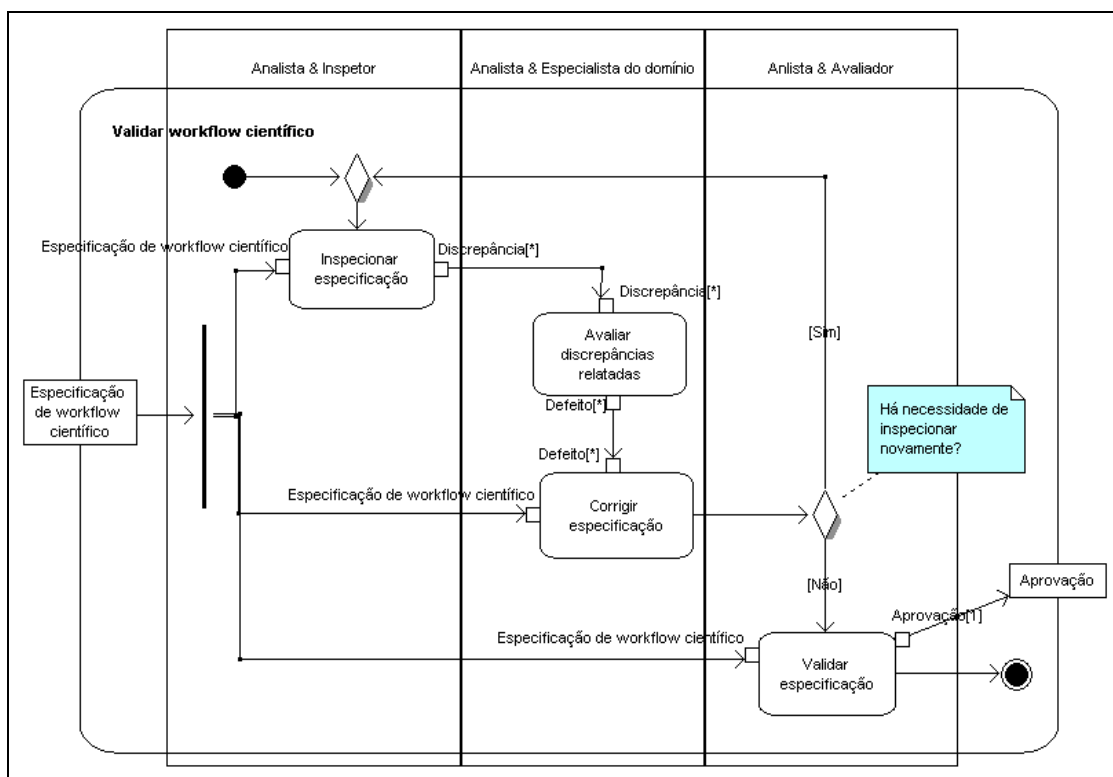


Figura 4.4. Detalhamento da fase “Validar *workflow* científico”, contendo tarefas de inspeção e validação da Especificação de *workflow* científico abstrato.

Inspeccionar especificação

A primeira tarefa da Figura 4.4 é “Inspeccionar especificação”, cujo objetivo é revisar a especificação do *workflow* científico a fim de identificar defeitos no documento. A revisão é feita através de inspeção (TRAVASSOS ET AL., 2002) aplicada nos formulários preenchidos e nos modelos durante uma ou mais reuniões. A inspeção foi escolhida, pois quando comparada a outras técnicas de revisão, apresenta vantagens como um mecanismo efetivo para a detecção de problemas, apresentando alto percentual de defeitos encontrados, assim como permite um retorno

qualitativo e quantitativo do projeto e seus artefatos (MELO ET AL., 2001). A inspeção aplicada é *ad hoc*, ou seja, consiste numa verificação livre do documento, sem a utilização de um protocolo ou guia de revisão específico.

Na reunião, o analista deve apresentar os modelos, formulários e deve descrever o objetivo da reunião. Deve-se também apresentar aos inspetores a notação do modelo e os formulários, antes da aplicação da inspeção, caso estejam realizando-a pela primeira vez. Depois o material da inspeção é distribuído, sendo esse composto pela especificação de *workflow* científico e uma planilha para relato das discrepâncias encontradas.

Então, o analista determina um tempo para que os inspetores realizem a inspeção, e não é necessário que seja realizada em uma única reunião. Contudo, um formulário ou planilha deve ser fornecido aos inspetores para que estes relatem as discrepâncias. Ao final, tem-se um conjunto de discrepâncias relatadas por todos os inspetores, contendo para cada discrepância encontrada, a sua descrição e o seu tipo.

Para facilitar a classificação e detecção de possíveis fontes de problemas no procedimento para concepção do *workflow* científico, utiliza-se uma taxonomia para classificação das discrepâncias e dos defeitos encontrados. A taxonomia é apresentada na Tabela 4.1 e foi baseada na taxonomia de defeitos em requisitos proposta por SHULL (1998). Essa taxonomia é utilizada em inspeções aplicadas em diferentes processos de *software* da indústria. A taxonomia é usada pelo Analista, Especialista do domínio e Inspetores.

Tabela 4.1. Classificação das discrepâncias, adaptado de [SHULL, 1998].

Tipo de discrepância	Definição
Omissão	Informação necessária não incluída. (a) atividade, artefato ou ferramenta não identificada. (b) termos ou conceitos não identificados. (c) alguma relação entre artefato, atividade e ferramenta não identificada. (d) falta de seções no documento. (e) algum campo de informação dos formulários não preenchido.
Ambigüidade	Informação passível de ter múltiplas interpretações. (a) conceito do domínio, atividade, artefato pode conter várias interpretações devido a diferentes termos utilizados. (b) vários significados para um termo, dependendo do contexto.
Inconsistência	Informações conflitantes. (a) termos ou conceitos conflitantes. (b) modelagem e especificação conflitantes.
Informação estranha	Informação desnecessária. (a) informações fornecidas na especificação ou modelagem que não são necessárias ou não são utilizadas de alguma forma.
Fato incorreto	Informação que não é verdadeira para as condições especificadas. (a) termo, conceito que não descreve um fato que não é verdadeiro. (b) atividade, resultado ou ferramenta que não pertence ao contexto ou domínio.

Avaliar discrepâncias relatadas

A averiguação das discrepâncias relatadas pelos Inspetores é realizada durante a tarefa “Avaliar discrepâncias relatadas” (vide Figura 4.4), sendo responsabilidade do Especialista do domínio, do Analista e dos Inspetores realizar a

avaliação. Na Engenharia de *Software*, a avaliação de discrepância é importante porque nem todos os relatos de discrepâncias feitos pelos inspetores são realmente defeitos da especificação, como já descrito por HUMPHREY, em 1989 (HUMPHREY, 1989). Falsos positivos devem ser eliminados do conjunto de defeitos, pois a mudança do documento com base em falsos positivos pode acarretar na introdução de novos defeitos (KALINOWSKI, 2004).

Baseado nas práticas de inspeção propostas por SAUER ET. AL. (2000), o primeiro passo consiste no agrupamento das discrepâncias relatadas pelos inspetores em uma lista. Este passo é de responsabilidade do Analista. O objetivo é agrupar os relatos de discrepâncias relacionadas aos mesmos defeitos, que foram encontradas por mais de um inspetor, mantendo um registro único para cada discrepância. O segundo passo é a discriminação de defeitos, no qual participam o analista e o especialista do domínio, discutindo os relatos de discrepâncias, buscando classificá-las como falso positivo ou defeito. A avaliação das discrepâncias relatadas deve ser executada com atenção e qualquer dúvida no relato deve ser averiguada com o inspetor, antes de assumir que se trata ou não de um defeito real. Ao final, os falsos positivos são descartados e uma lista final com os defeitos presentes na especificação é construída.

Corrigir especificação

A lista de defeitos é um insumo para a tarefa “Corrigir especificação” (Figura 4.4). O objetivo disso é, através dos defeitos reais identificados, fazer sua correção na “Especificação de *workflow* científico”, tanto no modelo como nos formulários. Esta tarefa é realizada pelo Analista e o Especialista do domínio.

Depois de corrigidos os defeitos encontrados na Especificação de *workflow* científico, o Analista e o Especialista do domínio podem tomar a decisão conjunta de disponibilizar aos inspetores os formulários e os modelos para uma nova rodada de inspeção. Essa decisão é tradicional nos procedimentos de inspeção realizados no desenvolvimento de *software* (FAGAN, 1976). Entretanto, os critérios a serem utilizados para tomar esta decisão variam projeto para projeto, ficando a cargo do Analista e Especialista do domínio. Como exemplo pode-se citar: número de defeitos encontrados é abaixo da média histórica; número de defeitos graves encontrados é abaixo da média; número de alterações de severidade grave abaixo da média; dentre outros.

Validar especificação

A última tarefa da Figura 4.4 é “Validar especificação”, na qual é avaliado se a especificação representa o experimento através do *workflow* científico identificado. A validação é realizada em uma reunião, consistindo na apresentação da especificação. A validação da especificação é feita através de revisões *walkthrough* (MELO ET AL., 2001) nos modelos e nos formulários.

A técnica de *walkthrough* é um processo de revisão por pares, podendo variar no seu nível de formalidade desde o formal, como a avaliação de requisitos pela alta gerência, até informal, como realizado por grupos de programadores durante o desenvolvimento de *software* (YOURDON, 1989). No procedimento proposto nesta dissertação, o nível de formalidade do *walkthrough* é intermediário, pois os avaliadores (participantes) são outros especialistas do domínio (pares), como no nível informal, mas ao mesmo tempo a tarefa tem como objetivo a aceitação, como no nível formal. O foco desta técnica está no consenso entre os participantes e através disso, possivelmente, eliminar problemas ainda existentes (MELO ET AL., 2001). Durante uma reunião de *walkthrough*, o produtor do artefato repassa e explica todo o material sob avaliação, enquanto os avaliadores apontam questões identificadas por eles, caso existam (PRESSMAN, 2010).

Nesta tarefa participam: o Analista, cuja responsabilidade é apresentar e moderar a reunião, também anotar defeitos identificados; os Avaliadores, cuja responsabilidade é avaliar toda a especificação; e o Especialista no domínio, cuja responsabilidade é auxiliar o Analista em questões técnicas sobre o domínio do experimento. A tarefa, com uso de *walkthrough*, é realizada da seguinte forma:

1. O Analista prepara a documentação e distribui para todos os Avaliadores participantes para que esses possam ler e fazer anotações prévias de sugestões e defeitos.
2. O Analista prepara uma apresentação que serve para guiar a revisão de todos os modelos e formulários durante a reunião.
3. Durante a reunião, ao mesmo tempo que o Analista apresenta os modelos e formulários, os Avaliadores acompanham a leitura do documento, podendo fazer sugestões, comentários e críticas. Como consequência adicional, podem-se detectar defeitos ainda presentes na Especificação, que por ventura não foram capturados na inspeção. Estes devem ser anotados e posteriormente avaliados e corrigidos.
4. Ao final, os Avaliadores são responsáveis por decidir se o documento contém as informações que realmente representam o experimento como um *workflow* científico abstrato.

Importante destacar que, caso a especificação não seja aprovada, deve haver uma nova iteração no procedimento, partindo desde a fase de Identificação do *workflow* científico abstrato até a validação.

4.3 Especificação de *workflow* científico em nível abstrato

A especificação de *workflow* científico foi desenvolvida para ser apresentada sob duas perspectivas: uma textual e outra gráfica. O principal motivo para essa divisão é o fato que essas duas representações são complementares. A representação gráfica através de modelos permite descrever, por exemplo, a *ordem de execução* das atividades, os dados consumidos e produzidos pelas atividades e os pontos de tomada de decisão do *workflow* científico, enquanto a representação textual fornece mais detalhes sobre os elementos do modelo. Por exemplo, uma atividade no *workflow* tem um formulário associado e nele há um campo chamado “descrição” no qual detalha como se deve realizá-la. Esta divisão é justificável, pois a inserção de todas as informações na representação gráfica poderia comprometer a sua legibilidade, uma vez que o modelo se tornaria sobrecarregado e a leitura dos modelos mais complexa.

Os formulários são representações textuais de três elementos da modelagem: atividade, artefato e ferramenta. Cada formulário reúne as características destes elementos, mas também traz referências a outros elementos do próprio *workflow*. Com isso, definem-se de forma explícita as associações presentes no modelo com o propósito detalhar o experimento científico, complementando as informações ao modelo de *workflow* científico. Adicionalmente, a descrição textual está em linguagem natural, permitindo ao especialista do domínio discutir e entender a especificação durante a concepção, pois a representação visual é uma notação mais formal, com a qual o especialista do domínio pode não estar totalmente familiarizado.

O modelo de *workflow* científico é expresso através de diagrama de atividades, sendo a representação mais próxima da suportada pelos SGWfC. Entretanto, as notações não são específicas para representar o domínio de *workflow* científico, pois seu foco é na modelagem de comportamentos de sistemas. Logo, torna-se necessário estender a notação do diagrama de atividades para que seja capaz de representar os elementos do domínio.

4.3.1 Descrição textual através de Formulários

A idéia de utilizar formulários para descrição textual iniciou-se no projeto HP-COPPE⁴, no qual o Grupo ESE⁵ teve participação. Os formulários criados para tal representação textual eram atividade e artefato. Estes também foram utilizados em CHAPETTA (2006), no qual as atividades de um estudo experimental foram descritas através desses formulários. Os formulários descreviam atividades genéricas e seus artefatos para qualquer categoria de estudo experimental e não eram específicos para experimentos *in virtuo* e *in silico*. Os formulários continuaram a ser utilizados no Grupo ESE para descrever seus estudos experimentais. Logo, a idéia de utilizar formulários como complemento a descrição de estudos experimentais não é inédita, sendo aproveitada na abordagem proposta nessa dissertação.

Nesta pesquisa, os dois formulários (atividade e artefato) foram adaptados e expandidos para o novo contexto de experimentos *in virtuo* e *in silico*. Como exemplo desta alteração, pode-se citar a introdução do campo “Custo computacional”, que captura informação relativa ao tempo gasto na execução de uma atividade, fato importante quando são utilizados recursos computacionais compartilhados entre diversos experimentos. Adicionalmente, mais um formulário foi desenvolvido, que permite descrever as informações capturadas relativas às ferramentas utilizadas no experimento científico. As alterações foram motivadas pela característica inerente a experimentos *in virtuo* e *in silico* da alta dependência de recursos computacionais para realização de suas atividades. Cada formulário é distinto no seu conjunto de informações e cada campo conta com uma descrição textual do seu significado e pode estar acompanhado de exemplos.

Destaca-se que as informações dos formulários de atividade e artefato são complementares aos elementos presentes nos modelos de *workflow* na notação de diagrama de atividades em nível abstrato, isto é, existem representações visuais para esses elementos nos modelos. Contudo, o formulário de ferramentas não apresenta representação visual no diagrama de atividades. Durante o mapeamento optou-se pela não representação de ferramentas no modelo de *workflow*, pois não se encontrou na notação de diagrama de atividades elemento compatível com o conceito de ferramentas. Além disso, o fato de haver variadas opções de ferramentas para uma atividade específica, quando o *workflow* científico se encontra descrito no nível abstrato, acarretaria em diversos elementos visuais (ferramentas) conectados a

⁴ <http://lens.cos.ufrj.br:8080/hpcoppe>

⁵ <http://lens-ese.cos.ufrj.br/ese/>

determinada atividade, o que poderia resultar em poluição visual do modelo, tornando sua leitura mais difícil.

Formulário de atividades

Conforme definido anteriormente, uma atividade se caracteriza como a representação de um passo dentro do fluxo de execução do *workflow* do experimento científico. No formulário atividade, algumas características gerais foram baseadas em campos do formulário apresentado no projeto HP-COPPE e em CHAPETTA (2006), sendo eles: nome, descrição, obrigatoriedade, ferramenta, insumo, produto, papel, pré-condição, pós-condição, pré-atividades, sub-atividades e outros comentários. Do formulário original, somente foram adaptadas as explicações sobre o que significa cada campo para refletir o contexto de experimentos *in virtuo* e *in silico* que utilizam *workflow* científico. Todos estes campos foram avaliados e mantidos, pois em grande parte são descrições textuais de representações gráficas de características da atividade presentes no diagrama de atividades.

Exceto pelos campos ferramenta e obrigatoriedade, todos os outros podem aparecer nos modelos na notação de diagrama de atividade. Os campos ferramenta e obrigatoriedade foram mantidos, pois estes são pertinentes à experimentação baseada em simulação. O campo ferramenta lista todos os programas (ou sistemas) utilizados numa atividade, e como experimento *in virtuo* e em especial *in silico* são altamente dependente de ferramentas, continua justificada sua manutenção. O campo obrigatoriedade é justificado pelo fato que o experimento pode ter variações no conjunto de atividades que serão executadas, porém algumas são obrigatórias. De fato, o experimento científico é um conjunto de diferentes fluxos de atividades, que variam entre si em alguns pontos específicos.

Entretanto, o formulário foi adaptado e novas informações foram inseridas nele, visando capturar informações características de atividades realizadas em *workflows* para experimentos *in silico*. Assim, foram inseridos os seguintes campos:

- **Tipo de atividade:** Devido à natureza dos experimentos *in silico* quanto a sua dependência aos recursos computacionais, as atividades de um *workflow* foram divididas em três classificações possíveis, da maior para menor intensidade, sendo elas: Automatizada, Semi-automatizada e Manual.
- **Capacidade de paralelismo:** Algumas atividades permitem que seus passos sejam divididos e executados separadamente através de recursos computacionais (e.g. ambientes em grade ou *clusters*), assim aumentando seu desempenho. Essas atividades podem ser realizadas por diversos

indivíduos ou recursos computacionais ao mesmo tempo. Entretanto, se a atividade não possuir capacidade de paralelismo, então há uma restrição quanto a sua execução.

- **Risco:** A atividade, quando sujeita a fontes de perigo, possibilidade de perda ou infortúnio, está vulnerável a falhas. Essas situações são denominadas risco e estão associadas à execução do *workflow* científico e suas atividades. Com a captura das situações e dos riscos que podem ocorrer, o pesquisador ao criar seu *workflow* concreto deve considerar meios para mitigar sua ocorrência.
- **Frequência de utilização:** A atividade pode ser utilizada mais de uma vez no mesmo experimento *in silico*, pois o fluxo de execução pode conter decisões ou repetições que obrigam a sua re-execução diversas vezes. Nesta abordagem, isso se caracteriza como frequência de utilização da atividade. A frequência de utilização pode ser classificada como: alta, média ou baixa.
- **Custo computacional:** A atividade pode necessitar de apoio de recursos computacionais para que seja realizado. Quando isto acontece torna-se interessante registrar qual o seu custo computacional e o tempo de execução. Caso haja uma fórmula que calcule o tempo de execução, esta deve ser explicitada. Nesta abordagem, também se adota uma classificação qualitativa para o custo computacional: alta, média e baixa. Com o custo computacional das atividades registradas, pode-se estimar o tempo de execução do experimento.

O formulário de **Atividade** e seus campos estão representados na Tabela 4.2. Nela também estão presentes as descrições de cada campo e, em alguns casos, apresentando exemplos de respostas esperadas.

Formulário de Artefatos

O formulário descrito na Tabela 4.3 apresenta as características do elemento **Artefato**, gerado no modelo de *workflow* científico. Equivalente ao realizado para Atividade, os campos do formulário de Artefato foram baseados em instrumentos utilizados no Grupo ESE para especificar estudos experimentais de qualquer categoria. Todos os campos baseados nos formulários foram analisados, e dado as características do elemento artefato identificados na revisão (Capítulo 2 e Capítulo 3), se mostraram necessários para representar com detalhes produtos e insumos do *workflow* científico de um experimento. Novamente, somente as descrições desses

campos que explicavam seu significado foram alteradas. Os campos do formulário de artefato original são nome, descrição, origem, utilização, formato, ferramenta, sinônimos e outros comentários. O campo utilização é composto por três sub-campos: atividade - que é referência à atividade consumidora ou produtora do artefato; insumo / produto – indica se o artefato é consumido ou produzido, respectivamente; condições de obrigatoriedade – indica se o consumo ou produção de um artefato é obrigatório para uma atividade.

Tabela 4.2. Formulário de Atividade e seus campos.

Atividade	[NOME DA ATIVIDADE]
Descrição	Descreva as etapas de execução da atividade de forma detalhada. Também deve descrever o objetivo da atividade.
Tipo de atividade	Informe o tipo de atividade. Os possíveis tipos de atividades são: <ul style="list-style-type: none"> • Manual: uma atividade na qual não há ou existe pouca utilização de ferramentas que auxiliem em sua execução, sendo essencial a participação do pesquisador/cientista. • Semi-Automatizada: uma atividade ou ação na qual há uma interação do pesquisador e apoio de algum recurso computacional necessário. • Automatizada: uma atividade ou ação que não requer a participação essencial do cientista, pois existe recurso computacional responsável por apoiar sua execução completamente.
Obrigatoriedade	Informe se essa atividade sobre a sua condição de obrigatoriedade. As possíveis classificações são: obrigatória ou opcional. Caso a atividade seja classificada como opcional, cite as situações em que a sua execução é desejada.
Ferramenta	Liste os <i>softwares</i> , sistemas ou simuladores utilizados durante a execução da atividade.
Insumo	Liste os artefatos consumidos pela atividade. Esses artefatos representam a “matéria-prima” da atividade, ou seja, os dados de entrada necessários para executar a atividade.
Produto	Liste os artefatos produzidos pela atividade. Essa produção pode ser de um artefato novo ou a transformação de um artefato existente. Esses artefatos são os dados de saída produzidos pela atividade. Registre a quantidade gerada para cada produto.
Papel	Liste os papéis desempenhados por um ou mais indivíduos para a realização da atividade. Responsáveis pela execução da atividade (e.g., geólogo).
Pré-condição	Liste as pré-condições que devem ser atendidas para iniciar a execução da atividade (e.g., atividade X executada, recurso financeiro disponível e alocado, etc.).
Pós-condição	Liste as pós-condições que são geradas ao final da execução da atividade (e.g., dado gravado no repositório, recurso computacional liberado).
Pré-atividade	Liste as atividades que devem ter sido concluídas para que seja possível a execução dessa atividade.
Sub-atividade	Liste as atividades que compõem esta atividade (se existirem). Neste caso, cada sub-atividade listada deverá ter uma ficha, equivalente a essa, preenchida.
Capacidade de paralelismo	Informe se a atividade pode ser paralelizada ou distribuída. Caso positivo, cite as situações em que essa atividade pode ser paralelizada.
Risco associado	Liste os riscos associados à atividade. Os riscos são fontes de perigo, possibilidade de perda ou infortúnio. Como exemplo: entrada de dados manual, transformação de dados ou entradas.
Freqüência de utilização	Classifique a freqüência de uso da atividade. Pode ser utilizada uma classificação qualitativa (Alta, Média ou Baixa). Também pode ser especificado um valor numérico.
Custo computacional	Informe o tempo médio despendido na execução da atividade. Caso haja uma fórmula que calcule o tempo de execução, explicite-a.
Outros comentários	Se necessário, utilize este espaço para fazer outros comentários que julgar pertinente.

Entretanto, houve a necessidade em acrescentar um campo a este formulário, este é denominado **metadado**. Este campo descreve a estrutura interna de um artefato, pois este pode ser composto por um conjunto de outras informações. Essa estrutura deve ser capturada, pois representa a forma como o conhecimento está

organizado. A justificativa está no fato que isto permite, ao se criar o *workflow* científico concreto, determinar quais são as informações a serem utilizadas, em quais artefatos estão presentes, como estão organizadas e em qual parte esta se encontra internamente.

Tabela 4.3. Formulário de Artefato e seus campos.

Artefato	[NOME DO ARTEFATO]		
Descrição	Descreva resumidamente o artefato.		
Origem	A classificação quanto à origem é: interna ou externa. Interna se é produzido por uma das atividades do <i>workflow</i> científico. Externa se é produzido por alguma atividade que não faz parte do <i>workflow</i> científico que está sendo modelado.		
Utilização	Informe na tabela abaixo as atividades na qual esse artefato é gerado/usado. Se o artefato é usado pela atividade, esse é um Insumo. Se o artefato é gerado pela atividade esse é um Produto. Para cada caso em que o artefato é usado/gerado por uma atividade, descreva sua condição de obrigatoriedade. Faça isso para cada atividade com a qual esse artefato se relaciona. As classificações possíveis são:		
		<ul style="list-style-type: none"> • Obrigatório: Artefato é um insumo obrigatório de uma atividade, isto é, sempre é consumido; Artefato é produto obrigatório de uma atividade, isto é, sempre é produzido. • Opcional: Artefato é um insumo opcional de uma atividade, isto é, pode ser consumido, dependendo de uma determinada condição; Artefato é produto opcional de uma atividade, isto é, pode ser produzido, dependendo de uma determinada condição. 	
	Atividade	Insumo/ Produto	Condição de obrigatoriedade
	Nome da Atividade	Insumo ou Produto	Descrição da obrigatoriedade: Obrigatório; Opcional.
Formato	Informe o formato do artefato (digital, digitalizável ou físico). Digital se é arquivo de computador. Digitalizável se é um documento em papel, que pode ser transformado em um arquivo de computador (através de um scanner). Físico se é uma amostra de material (ex.: rocha, areia, etc.). Caso o formato do artefato seja digital, pode-se listar a extensão digital do formato. Descreva as extensões digitais nas quais o artefato pode ser encontrado no experimento.		
Metadado	Descreva informações relacionadas ao artefato, quando pertinentes. Informações como campo, nº de colunas, disposição dos campos (e.g., colunas de 1 a 4 representam o id, colunas de 5 a 10 representam o peso).		
Ferramenta	Se o artefato for um arquivo de computador, diga que ferramenta(s) lida(m) com este arquivo.		
Sinônimos	Liste aqui os sinônimos deste artefato.		
Outros comentários	Se necessário, utilize este espaço para fazer outros comentários que julgar pertinente.		

Formulário de ferramentas

Diferente dos outros dois formulários desta abordagem, **Ferramenta** foi desenvolvido completamente do seu início. Este está representado na Tabela 4.4, apresentando seus campos e as descrições explicativas. Além dos campos nome, descrição e outros comentários que capturam informações gerais da ferramenta, foram definidos outros campos como:

- **Tipo de aplicação:** Ferramentas possuem características próprias quanto ao seu uso, podendo ser: acessadas através de uma interface com usuários; acessada por outra aplicação como um serviço; ou executada

automaticamente através de comandos ou chamadas remotas. Assim, foi definido que ferramentas podem ser classificadas em três tipos básicos: *Batch*; *Serviço*; *Interface*.

- **Versão:** Ferramentas podem evoluir com o passar do tempo e por necessidade podem adicionar ou modificar ou retirar funcionalidades, e conseqüentemente isso pode acarretar versões novas.
- **Sistema Operacional:** Ferramentas, em geral, são executadas sobre determinado sistema operacional, devendo-se capturar esta informação, pois pode constituir em restrição para execução do *workflow* científico.
- **Local de execução:** Ferramentas podem ser executadas em locais distintos. Foi definido que o local de execução deve ser classificado como: Local, Remota, *Cluster*, *Grade (Grid)*.
- **Forma de disparo:** Ferramentas estão armazenadas em algum local e, em geral, podem ser invocadas por um caminho (*path*) ou método de invocação. A forma de disparo tem como objetivo capturar a maneira como a ferramenta é disparada/invocada durante a execução do *workflow*.
- **Extensões digitais suportadas:** Ferramentas são capazes de ler e escrever artefatos. Por isso deve listar as extensões que ela compreende. Destaca-se que os artefatos podem ser digitais e, conseqüentemente, possuem extensões digitais. Ao se listar as extensões digitais, cria-se uma ligação entre os possíveis artefatos e ferramentas no *workflow*.

Tabela 4.4. Formulário de Ferramenta e seus campos.

Ferramenta	[NOME DA FERRAMENTA]
Descrição	Descreva a ferramenta de forma detalhada.
Tipo de aplicação	Informe o tipo de aplicação. Alguns tipos de ferramentas são: - <i>Batch</i> : Arquivos utilizados pra automatizar tarefas; - <i>Serviço</i> : Aplicação que provê funcionalidades para o uso por outra; - <i>Interface</i> : Aplicação na qual há uma interação com o usuário através de telas, formulários ou comandos.
Versão	Cite a versão da ferramenta utilizada.
Sistema Operacional	Liste os sistemas operacionais nos quais a ferramenta é executada.
Extensões digitais suportados	Liste os formatos que a ferramenta suporta, consome ou manipula.
Local de Execução	Liste os locais de execução que a ferramenta suporta. Alguns tipos são:- • Local – se é executado no mesmo local que o <i>Workflow</i> ; • Remota – se é invocado e é executado em infraestrutura diferente do <i>Workflow</i> ; • Cluster – se é executado em ambientes de alto desempenho; • <i>Grade (Grid)</i> – se é executado em ambientes computacionais dispersos para alto desempenho.
Forma de disparo	Liste a forma de disparo da ferramenta. Esta é a forma como a ferramenta será ativada. Por exemplo: Linha de comando "H:\Movimento\CMD09.EXE".
Outros comentários	Se necessário, utilize este espaço para fazer outros comentários que julgar pertinente.

4.3.2 Representação gráfica do *workflow* científico

Nesta abordagem, a representação gráfica do *workflow* científico segue a notação do diagrama de atividades da UML 2.2 (OMG, 2009). A notação permite a representação de fluxo de controle (ordem de execução) e fluxo de dados (transferência de artefatos, como produtos ou insumos) entre atividades do *workflow* científico. Contudo, esta notação é genérica para modelagem de *software*. Assim, com o intuito de especializar para representação de modelos de *workflow* científico, elementos da notação de diagrama de atividades foram estendidos a fim de agregar significado específico a eles.

Extensão da notação de diagrama de atividades para domínio do *workflow* científico

A UML permite que seus modelos sejam estendidos através de um mecanismo genérico para representar modelos de domínios ou plataformas específicos. Esse mecanismo é denominado de **Perfil** (*Profile*) (SINAN, 2002; OMG, 2009). O Perfil UML permite o refinamento da notação pela adição de novos elementos e conceitos semânticos ao seu padrão, através do uso de estereótipos, *tagged values* e restrições aos elementos de seus modelos. Os estereótipos são marcações de extensão de conceitos; a sua notação gráfica é marcada pelo símbolo “<< >>” com o nome do estereótipo escrito meio dos símbolos << >>. *Tagged Value* são atributos associados aos estereótipos criados, enquanto restrições são regras que devem ser obedecidas no perfil, para que o modelo gerado a partir deste seja verdadeiro e correto.

Logo, foram selecionados elementos presentes no metamodelo da UML (Figura 2.8) que pudessem ser refinados e passassem a representar semanticamente os elementos do domínio de *workflow* científico, descritos na seção 2.2. A Tabela 4.5 apresenta o resultado desse mapeamento e refinamento. A primeira coluna apresenta o elemento do domínio de *workflow* científico; a segunda coluna apresenta o elemento do diagrama de atividades para qual o primeiro foi mapeado; e a terceira coluna apresenta a descrição do mapeamento.

Nesta abordagem, escolheu-se o elemento Atividade (*Activity*), da notação UML, para representar o elemento de *Workflow* no modelo. A justificativa é que semanticamente ambos são similares, pois representam fluxos de execução que contém outros elementos internamente. O elemento Atividade da UML foi estendido através de estereótipo <<*Workflow*>>.

No domínio de *workflow* científico, uma atividade é um passo atômico dentro do experimento científico. Logo, para fins de modelagem o elemento Ação Opaca (*Opaque Action*) do diagrama de atividades foi escolhido para representar uma atividade atômica no modelo do *workflow*. Entretanto, uma atividade atômica do *workflow* pode ser especializada em três tipos: Manual, Semi-automatizada e Automatizada. Assim seguindo essa definição foram criados três estereótipos, sendo cada estereótipo descrito como:

- O estereótipo <<Manual>> é atribuído a Ação Opaca na qual não há ou é pouca a utilização de ferramentas que auxiliem na execução, sendo essencial a participação do pesquisador.
- O estereótipo <<Semi-automatizada>> é atribuído a Ação Opaca na qual é necessária a utilização de ferramentas que auxiliem na execução, sendo fundamental a participação do pesquisador.
- O estereótipo <<Automatizada>> é atribuído a Ação Opaca na qual existe a utilização de ferramentas que auxiliam na execução, sendo quase nenhuma a participação do pesquisador.

Tabela 4.5. Mapeamento entre elementos do domínio de *workflow* científico para elementos do diagrama de atividades da UML 2.2.

Metamodelo <i>workflow</i> científico	Metamodelo UML	Descrição Mapeamento
<i>Workflow</i>	Atividade (<i>Activity</i>)	<i>Workflow</i> do experimento contendo fluxos de execução. Atividade na UML estereotipada como <i>Workflow</i> .
Manual	Ação Opaca (<i>Opaque Action</i>)	Atividade atômica cuja principal característica é a intervenção do cientista. Ação opaca na UML estereotipada como Manual.
Semi-automatizada	Ação Opaca (<i>Opaque Action</i>)	Atividade atômica cuja característica é a intervenção do cientista com auxílio de uma ferramenta. Ação opaca na UML estereotipada como Semi-automatizada.
Automatizada	Ação Opaca (<i>Opaque Action</i>)	Atividade atômica cuja característica é automação através de uma ferramenta. Ação opaca na UML estereotipada como Automatizada.
Structured	<i>Call Behaviour Action</i>	Sub- <i>workflow</i> utilizada em um <i>Workflow</i> . <i>Call Behaviour Action</i> da UML estereotipado como <i>Structured</i> .

Para representar um *sub-workflow* dentro de um *workflow* foi escolhida a especialização do elemento Ação denominado *CallBehaviour Action*, já que a notação do diagrama de atividades define que para se representar a evocação de um comportamento composto, deve-se utilizar esse elemento. No entanto, como a própria UML já define um estereótipo para este elemento, inicialmente preferiu-se utilizá-lo, tentando através desse manter a consistência com a notação. Assim, a Ação do tipo *CallBehaviour Action* é estereotipada como <<*Structured*>> para representar um *sub-workflow*.

Os elementos que representam o início e término de fluxos, tanto de dados como de controle, não foram estereotipados, mantendo a mesma semântica da notação de diagrama de atividades. Entretanto, esses elementos agora atuam sobre o fluxo de execução do *workflow* e não mais do processo ou comportamento de um sistema ou *software*. O nó de início é utilizado para representar o início de um fluxo de execução do *workflow*. O nó Final de processo é utilizado para representar o termino do fluxo de execução do *workflow*, isto é, se um desses elementos é alcançado durante a execução, então o fluxo de execução de todo *workflow* termina. O nó final de fluxo também determina o término, mas não é o final da execução de todo *workflow*, mas sim de um caminho de execução.

A Figura 4.5 apresenta o perfil definido a partir do mapeamento descrito nesta seção, a representação é através da notação de diagrama de classes da UML 2. As classes representadas na cor azul são os estereótipos (<<stereotype>>) definidos, enquanto as classes representadas na cor amarela são as metaclasses da notação do diagrama de atividades, estereotipadas com <<metaclass>>, que serviram como base para a extensão.

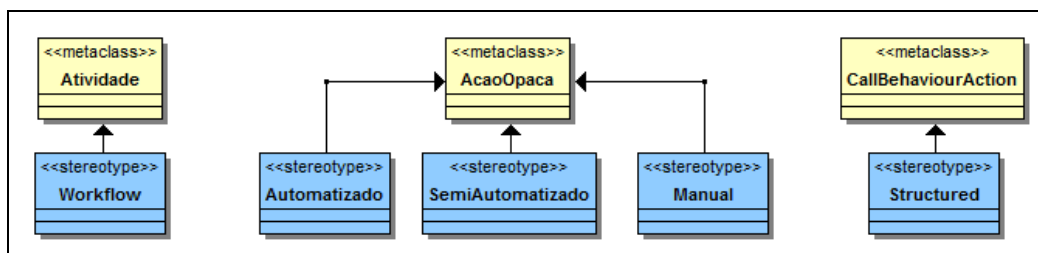


Figura 4.5. Perfil UML - Estereótipos e elementos do diagrama de atividades UML 2.2, versão um.

4.4 Prova de Conceito – Domínio de Evolução de Software

4.4.1 Objetivo da Avaliação

O principal objetivo da avaliação foi realizar uma primeira utilização da abordagem apresentada nesta dissertação e assim verificar possíveis problemas em sua usabilidade, construção e viabilidade. Desta forma, a avaliação consistiu na aplicação de uma prova de conceito e a partir dessa aplicação foi possível corrigir os problemas e identificar oportunidades de melhoria.

A prova de conceito foi realizada utilizando-se como fonte um experimento *in virtuo* no domínio de Evolução de Software, uma área de pesquisa da Engenharia de Software cujo objetivo é estudar o comportamento do software através de diversos ciclos de evolução. Mesmo não sendo o experimento completamente baseado em

simulação, ou seja, *in silico*, trata-se de experimento *in virtuo* e por isso ele foi utilizado, pois existe uma dependência de ferramentas e sistemas para executar certas atividades, como, por exemplo, a simulação da evolução do *software*. Dada essa característica, a representação através de *workflow* científico já se faz viável, tornando desta forma o experimento sobre Evolução de *Software* um candidato aceitável. Outro ponto se dá pelo fato de haver a disponibilidade de especialista de domínio para consulta e auxílio durante a aplicação da abordagem, o que influenciou na escolha deste experimento como prova de conceito.

4.4.2 Descrição do Domínio

A pesquisa sobre Evolução de *Software* tem como objetivo entender como sistemas evoluem e se modificam ao longo do seu ciclo de vida e como essa evolução pode influenciar no decaimento de sua qualidade. Para isso, podem-se construir modelos de simulação para ajudar a observar a evolução do *software* ao longo de sucessivos ciclos de manutenção. Em sua tese, ARAÚJO (2009) apresenta um modelo, baseado nas Leis de Evolução de *Software* (*LSE – Laws of Software Evolution*) (LEHMAN ET AL., 1998), que permite a observação do processo de decaimento do *software* ao longo do tempo através da simulação do comportamento de determinadas características do *software*.

O modelo de Araújo baseia-se em premissas descritas através de formulações lógicas das LSE. Essas formulações representam as tendências do comportamento de determinadas características de *software* ao longo do tempo (e.g., características da fase de codificação: esforço, tamanho, periodicidade, complexidade, confiabilidade, manutenibilidade). Entretanto, as premissas não permitem a observação direta do processo de decaimento da qualidade do *software*, pois não representam as influências que uma Lei exerce sobre outra. Assim, utilizam-se ferramentas para simular as características do *software*, através das equações definidas, que representam as influências entre as LSE e como estas afetam as características. A Figura 4.6 apresenta os procedimentos para o estudo sobre Evolução de *Software* (Araújo, 2009).

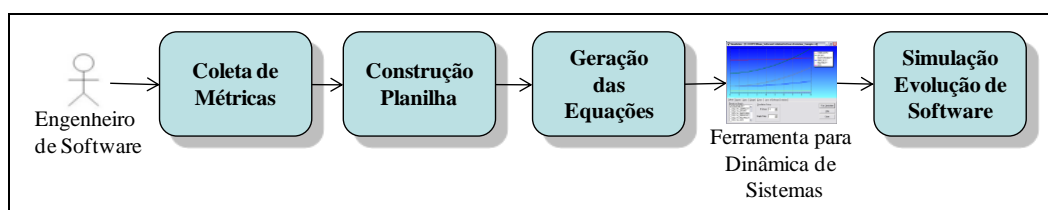


Figura 4.6. Processo para Observação de Evolução de *Software* (Araújo, 2009).

O modelo da Figura 4.6 e a descrição textual do procedimento são a especificação do estudo experimental, presentes no trabalho de ARAÚJO (2009). Contudo, a especificação, da forma como é documentada, pode apresentar problemas para sua utilização em outros contextos e por outros indivíduos. Por exemplo, na representação do modelo (Figura 4.6), estão misturadas informações como o perfil do pesquisador (Engenheiro de *Software*), a ferramenta utilizada (Ferramenta para Dinâmica de Sistemas) e as atividades do experimento. Essas informações distintas, sem uma organização de representação que explicita qual é o significado de cada uma dentro do modelo, podem gerar confusão para um pesquisador que pretende repetir o experimento *in virtuo*. Além disso, existem informações (e.g., descrição das atividades, insumos e produtos, dentre outras) que estão descritas em formato textual, o que provoca risco do pesquisador ao replicar o experimento, esquecer-se de utilizar alguma informação, pois não há um conjunto característico de informações pré-definidas para cada elemento atividade, ferramenta e artefato que deva sempre ser identificado.

Considerando o contexto, esse experimento foi escolhido para a aplicação da abordagem de concepção de *workflows* científicos, com intuito de avaliar a aplicabilidade do procedimento para composição e possíveis problemas nos instrumentos utilizados (Modelos UML e Formulários).

4.4.3 Aplicação do procedimento para composição

Etapa de identificação

Primeiramente, o modelo de contexto do experimento foi criado, como representado na Figura 4.7, a partir de ARAÚJO (2009) e de uma reunião com o especialista de domínio. Na reunião foram identificadas, inicialmente, três atividades, retiradas da Figura 4.6, presentes no trabalho de ARAÚJO (2009), e durante “Definir contexto do *workflow* científico”, identificou-se uma nova atividade que não aparece explicitamente na Figura 4.6. As atividades identificadas foram:

- 1) *Preparar dados para simulação*: na qual as métricas extraídas do processo real de desenvolvimento são padronizadas, avaliadas e excluídas caso apresentem comportamento incomum;
- 2) *Gerar as equações para simulação*: na qual são criadas as equações baseadas nas formulações da LSE e que servirão como modelo para simulação das características do *software*;
- 3) *Simular a evolução do software*: na qual ocorre a simulação da evolução das características do *software* para um determinado tempo definido;

- 4) *Analisar o resultado da simulação*: na qual o objetivo é gerar uma análise do resultado da simulação executada. Atividade inicialmente não descrita explicitamente.

Ainda durante a definição do contexto, também se identificou o papel do *Engenheiro de Software* e sua responsabilidade, que é garantir a qualidade dos dados escolhidos, das equações geradas e análise do resultado da simulação.

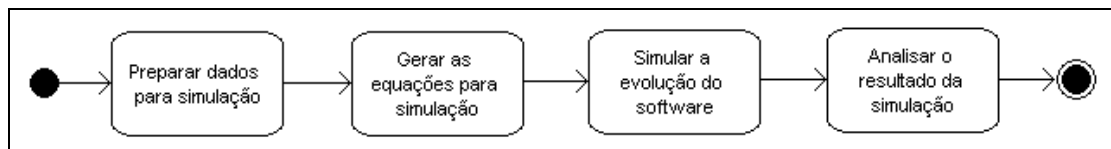


Figura 4.7. Modelo de contexto para o estudo “*Simulação da Evolução de Software*”.

O modelo de contexto (Figura 4.7) serviu como insumo para a tarefa de “Identificar e modelar requisitos do *workflow* científico”, sendo refinado e como consequência acabou-se por identificar pontos de decisão no conjunto de atividades (Figura 4.8). No modelo de *workflow* abstrato resultante foram representados tanto os fluxos de controle como os de dados, permitindo ao pesquisador visualizar as dependências entre as atividades do estudo.

Durante essa tarefa de “Identificar e Modelar requisitos do *workflow* científico”, inicialmente identificou-se que as três primeiras atividades eram compostas, como representado na Figura 4.8. Na época da realização da prova de conceito, as atividades compostas ainda eram estereotipadas com <<structured>> para representar o conceito de *sub-workflows*. Contudo, o termo *structured* não é comum no domínio de *workflow* científico, sendo os termos mais corriqueiros *sub-workflow* ou sub-atividade. Isto motivou, posteriormente, a reorganização do conjunto de estereótipos para representar os elementos do *workflow* científico identificados na literatura técnica.

A Figura 4.9 contém um modelo retirado da especificação de *workflow* que representa o *sub-workflow* “Gerar as equações para simulação”. Todas as atividades compostas foram organizadas de forma hierárquica e representadas em modelos separados. A escolha por essa organização em sub-modelos foi motivada pelo fato que a representação em um único modelo se mostrou complexa, por apresentar muitos elementos como atividades, artefatos, pontos de decisão, sincronização, dentre outros, resultando em um visual muito carregado e de difícil entendimento. Ao final, percebeu-se que essa prática poderia originar uma heurística para concepção, cujo conceito é a utilização de sub-modelos para representar hierarquias de *workflow* com o objetivo de tornar os modelos mais legíveis.

Durante a aplicação dos formulários não foi detectado nenhum problema no conjunto de informações que inviabilizasse sua utilização. Contudo identificou-se uma oportunidade de melhoria no formulário de Artefato relativo ao seu uso e às possíveis extensões digitais que ele pode se apresentar, em muito motivado pelo fato de haver “classes” de artefatos, que possuem um conjunto em comum de informações, e a noção de “instância” de uso de um artefato numa determinada atividade. Isso posteriormente acarretou na inclusão de mais dois campos no formulário de Artefato e na utilização de diagrama auxiliar para armazenar as “classes” de artefatos.

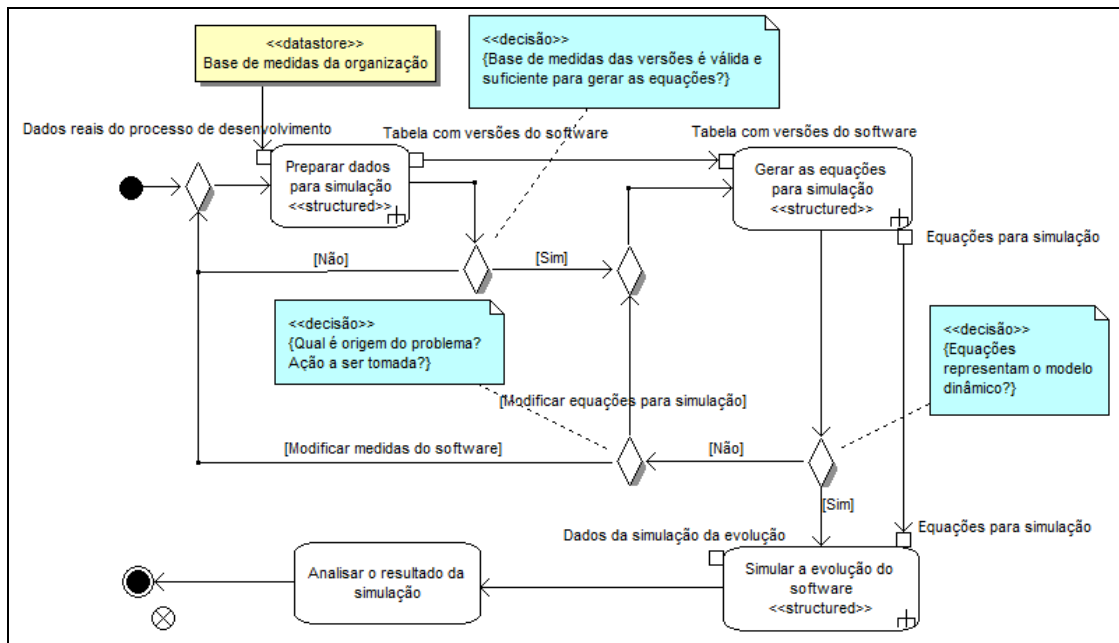


Figura 4.8. Workflow abstrato inicial para o estudo de *Simular a Evolução de Software*.

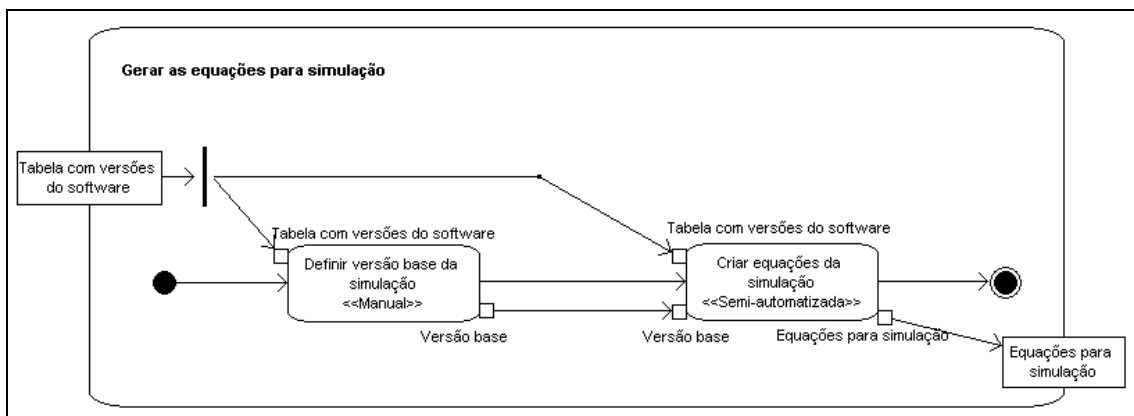


Figura 4.9. Sub-workflow para atividade composta “Gerar equações para simulação”.

Os formulários foram preenchidos ao decorrer das reuniões (no caso, duas) e houve identificação de novos elementos dos modelos de workflows científicos.

Entretanto, durante a aplicação, percebeu-se que certas informações eram repetidas nos modelos e nos formulários, o que poderia acarretar em defeitos por redundância. Por exemplo, as pré-atividades, insumos e produtos da atividade estão descritas como campos nos formulários, mas também estão representadas graficamente nos modelos. Para evitar isso, concluiu-se que era viável inserir as informações sobre atividades, artefatos dentro do próprio modelo UML. Isto é possível através do uso de *tagged values* nos estereótipos definidos para os elementos do *workflow* científico, permitindo que as informações sejam inseridas ao mesmo tempo em que é realizada a modelagem e evitando redundâncias desnecessárias.

Com a possibilidade de inserção de informações sobre os elementos do *workflow* dentro do próprio modelo, tornou-se necessário definir uma forma para armazenar as ferramentas identificadas. Inicialmente, estes elementos eram os únicos não expressos diretamente nos modelos de *workflow* (considerando a notação do diagrama de atividades), sendo somente descritos através de formulários textuais. Então, para armazená-los, posteriormente optou-se pela utilização de um diagrama auxiliar.

Na Tabela 4.6 é apresentado um formulário de atividade “*Criar equações da simulação*” retirado da especificação. Já na Tabela 4.7 é apresentado o formulário do artefato “*Equações para simulação*”, que contém as equações geradas em “*Criar equações da simulação*”. Na Tabela 4.8 é apresentado formulário da ferramenta “*Tabela_Excel*”, que foi utilizada na atividade “*Criar equações da simulação*”. No total, foram especificados 13 atividades, 3 ferramentas e 13 artefatos.

Como foi necessário apresentar a especificação para outros pesquisadores, foi definido um modelo de documento (*template*), que é composto por uma seção de introdução, descrição dos papéis envolvidos, apresentação dos modelos (diagramas de atividades) e formulários preenchidos. O *template* de especificação obriga ao analista a preencher todos os itens e organiza a apresentação das informações no documento. Esse modelo está disponível no anexo D.

Tabela 4.6. Formulário atividade *Criar equações da simulação* retirado da especificação.

Atividade	CRIAR EQUAÇÕES DA SIMULAÇÃO
Descrição	As equações combinadas (referentes à formulação lógica pra cada Lei de Evolução de <i>Software</i>) e os valores-base das características são definidos nas equações para simulação e serão efetivamente utilizados na simulação da evolução do <i>software</i> . Para tal utilizam-se duas técnicas: regressão linear e método de mínimos quadrados. A aplicação da técnica de regressão linear, apesar da possibilidade de aumento do erro, é condizente com a análise semiquantitativa dos dados, pois neste estudo é a tendência do comportamento de uma variável que deve ser considerada, mais do que seus valores individuais. A aplicação do método de mínimos quadrados, que é uma técnica de otimização matemática, procura encontrar o melhor ajustamento para um conjunto de dados, tentando minimizar a soma dos quadrados das diferenças entre a curva ajustada e os dados, cujas diferenças são chamadas de resíduos.
Tipo de atividade	Semi-automatizada.
Obrigatoriedade	Obrigatória.
Ferramentas	Tabela Excel
Insumos	TABELA COM VERSÕES DO SOFTWARE; VERSÃO BASE
Produtos	EQUAÇÕES PARA SIMULAÇÃO
Papéis	Engenheiro de <i>Software</i>
Pré-condições	Nenhuma.
Pós-condições	Nenhuma.
Pré-atividades	DEFINIR VERSÃO BASE DA SIMULAÇÃO
Sub-atividades	Nenhuma.
Capacidade de paralelismo	Não há.
Riscos associados	Nenhum
Frequência de utilização	Baixo.
Custo computacional	Baixo.
Outros comentários	Nenhum.

Tabela 4.7. Formulário artefato *Dados da simulação da evolução* retirado da especificação.

Artefato	DADOS DA SIMULAÇÃO DA EVOLUÇÃO		
Descrição	Este artefato é composto pelos dados gerados pela simulação da evolução do <i>software</i> , considerando as equações geradas, a versão base do <i>software</i> e o tempo determinado.		
Origem	Interna.		
Utilização	Atividade	Entrada/Saída	Condição de obrigatoriedade
	SIMULAR EVOLUÇÃO	Saída.	Obrigatória.
	ANÁLISE DO RESULTADO DA SIMULAÇÃO	Entrada.	Obrigatória.
Formato	Digital.		
Metadado	Valores simulados para as características do <i>software</i> . Também podem ser apresentados em forma de gráfico de tendência.		
Ferramenta	Illium <i>Software</i> Evolution		
Sinônimos	Nenhum		
Outros comentários	Nenhum		

Tabela 4.8. Formulário ferramenta *Tabela_Excel* retirado da especificação.

Ferramenta	Tabela_Excel
Descrição	Tabela no formato “.xls” onde já estão pré-definidos campos para o cálculo da regressão linear e do método dos mínimos quadrados. Também gera as equações para simulação a partir dos dados das versões do <i>software</i> e permite a definição dos valores das características do <i>software</i> para versão base.
Tipo de aplicação	Interface
Versão	Não há.
Sistema Operacional	Windows XP SP3 com Office Excel
Formatos Suportados	Formato: .xls
Local de Execução	Local.
Forma de disparo	-
Outros comentários	Nenhum.

Etapa de verificação e validação

Depois de realizada a tarefa “*Identificar e modelar workflow científico*”, a etapa de verificação começou. Foi solicitado a dois inspetores que verificassem e relatassem as discrepâncias encontradas em uma planilha. Os inspetores participaram da especificação do *workflow* científico abstrato. Para auxiliar os inspetores, foram cedidas algumas perguntas que poderiam ser utilizadas como guias auxiliares na inspeção, entretanto isto não estava previsto na abordagem inicialmente. Este fato motivou a criação de heurísticas para serem utilizadas na inspeção, servindo como uma melhoria na abordagem.

Destaca-se que a inspeção foi realizada separadamente: um dos inspetores se encontrava remotamente em outro *site*. Toda comunicação foi apoiada por troca de mensagens eletrônicas. Portanto, não é necessário que os inspetores estejam localmente para efetuar a inspeção. Como resultado da inspeção foram encontradas 20 discrepâncias no documento pelo inspetor 1 e 29 discrepâncias pelo inspetor 2. O inspetor 1 gastou 45 minutos para efetuar a inspeção, enquanto o inspetor 2 gastou 230 minutos para realizar a mesma atividade. A Figura 4.10 apresenta exemplo de discrepâncias relatadas pelos inspetores.

Nro	Tipo de discrepância	Descrição do Problema Encontrado	Local onde aparece?	Problema se repete?	Locais nas quais se repete
1	- Omissão	Na figura 1, mesmo descrito posteriormente, já deveriam aparecer as atividades de coleta de dados e de análise dos resultados	Seção 1	- Não se repete -	
2	- Ambigüidade	Na figura 2, após a atividade Gerar as equações para simulação, passa pela decisão se as equações representam o modelo dinâmico. Se não, chega-se na decisão sobre alterar as equações ou medidas do software. Essa decisão, primeiramente, precisa ter seus caminhos rotulados por Sim e Não. No caso de não, volta-se para a atividade Gerar as equações para simulação e, aparentemente, entra em loop, pois não fica explícita nenhuma mudança nas equações.	Seção 2.1	- Não se repete -	
3	- Omissão	Faltou o sub-workflow de análise dos resultados, como apresentado na seção 2.2.1.4	Seção 2.1	- Não se repete -	
4	- Omissão	Quando se fala de análise do desvio padrão, também pode-se complementar com análise de outliers	Em todo o documento	Sim em todo documento	
5	- Omissão	Algumas ferramentas podem ser utilizadas para a coleta de algumas das métricas consideradas, principalmente a nível de código fonte	Em todo o documento	Sim em todo documento	

Figura 4.10. Exemplo da planilha para relato de discrepâncias encontradas na inspeção.

O próximo passo foi discriminar as discrepâncias para descartar aquelas que não se tratavam de defeitos reais. Na lista do inspetor 1 foram encontrados dois falsos positivos, enquanto na lista do inspetor 2 foram encontrados cinco falsos positivos. Os inspetores possuíam diferentes níveis de conhecimento sobre o domínio. O inspetor 1 era especialista no domínio, enquanto o inspetor 2 tinha pouco conhecimento sobre o tema. Isto se refletiu diretamente nos tipos de defeitos encontrados: enquanto o inspetor 1 detectou 12 defeitos relacionados ao domínio do experimento, o inspetor 2 não achou defeitos no documento sobre o domínio, cuja grande parte foi relacionada a omissão de descrições de termos.

Por exemplo, o inspetor 1 relatou que a atividade “Analisar o resultado da simulação” é na verdade uma atividade composta e deveria ser representada como um *sub-workflow*, o que não foi identificado inicialmente. Esse tipo de defeito é grave, pois representa uma omissão de informações que são usadas na construção dos *workflows* concretos e sua execução. Posteriormente, este fato motivou a criação de heurística de organização da inspeção, que foi realizada durante a evolução da abordagem (descritas no próximo capítulo). A Figura 4.11 representa como o modelo de *workflow* abstrato ficou após a correção desse problema, a atividade modificada se encontra destacada através do retângulo tracejado em vermelho.

Após a correção dos defeitos, o documento foi validado em conjunto pelo analista e dois especialistas do domínio, sendo um deles um dos inspetores. Não houve reunião presencial de validação; somente foi distribuído o documento aos participantes, que ao final aceitaram a especificação sem restrições. Destaca-se que o *workflow* científico abstrato não foi implementado em nenhum SGWfC, pois não era o foco inicial da prova de conceito.

4.5 Considerações finais do capítulo

Os relatos após o uso da abordagem para concepção descreveram-na como uma forma viável de capturar a especificação de *workflow* abstrato incrementalmente para experimentos *in silico*, apoiando sua garantia de qualidade. O analista e os especialistas de domínio foram capazes de identificar as atividades e seus objetivos, os artefatos produzidos e consumidos, e as ferramentas, o que conduziu a uma especificação ***workflow*** científico disponível para a validação. Os formulários e modelos de *workflow* foram utilizados para auxiliar na identificação das informações necessárias para o experimento *in virtuo* e *in silico* e para representar os detalhes não habitualmente explicitados quando abordagens *ad hoc* são utilizadas.

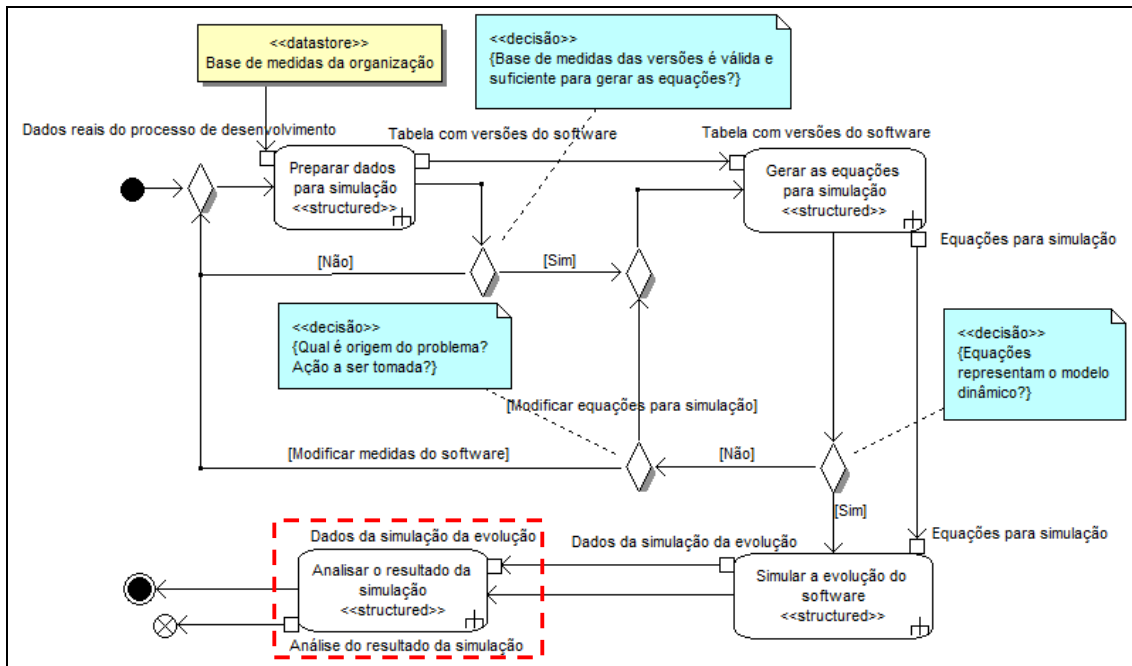


Figura 4.11. Workflow abstrato corrigido para o estudo de “Simular a Evolução de Software”.

A formalização do experimento por meio de modelos e formulários permite uma maior exploração da especificação como um importante recurso para *implementar* experimentos *in silico* (e *in virtuo*) utilizando SGWfC ou infraestruturas computacionais diferentes. Tendo sido a abordagem projetada para ser aplicado pelos engenheiros de *software* ou especialistas de outros domínios (cientistas ou pesquisadores), somente há a restrição de compreender a notação de diagrama de atividades da UML 2. O modelo descrito como um diagrama de atividades permite que a seqüência de atividades e os dados transmitidos sejam observados, enquanto os formulários são capazes de sintetizar a informação, o que possibilita o acesso rápido aos detalhes do experimento. No entanto, observou-se que ainda é necessário conhecimento sobre modelagem de diagramas de atividade UML para que a abordagem funcione. Entretanto isto não foi problema, pois todos os participantes eram engenheiros de *software* e possuíam conhecimento no uso da notação de diagrama de atividades.

Com relação à prova de conceito, alguns comentários devem ser realizados. Esta avaliação encontrou indícios interessantes com relação à viabilidade e às vantagens do uso da abordagem para concepção de *workflow* científico, contudo não se podem generalizar estes indícios para todos os domínios e possíveis contextos, dado que não há como se determinar todos os fatores que influenciaram durante sua aplicação. Adicionalmente, o ambiente no qual a prova de conceito foi aplicada não era controlado, pois não houve sua caracterização e não se mapeou as possíveis variáveis que influenciaram o resultado obtido com uso da abordagem. Por isso, não há como afirmar completamente que os indícios encontrados são totalmente válidos a

outros contextos. Outro ponto é que esta prova de conceito, apesar de representar um experimento *in virtuo* real, nunca teve como propósito a concepção de um *workflow* concreto a partir da especificação produzida. O foco sempre foi a avaliação da viabilidade da abordagem para concepção de *workflow* abstrato e não da criação de uma aplicação real, *workflow* concreto, a ser utilizada em projeto.

A partir da aplicação algumas possíveis melhorias foram identificadas. O uso de heurísticas para a inspeção do conteúdo dos formulários, para a organização das tarefas de garantia de qualidade e para o uso de sub-modelos na modelagem. Outra oportunidade de melhoria identificada é relativa ao uso de modelos auxiliares (diagrama de classes com notação estendida) para captura de ferramentas e suas informações, com os seguintes objetivos: evitar redundância de informações; separar a noção de “classe” e “instância” do elemento artefato; e armazenar em modelo os elementos do tipo ferramenta. Também se identificou a necessidade de reformular a extensão na notação, para representar mais precisamente os elementos do modelo de *workflow*. Adicionalmente, identificou-se a oportunidade de inserir as informações dos formulários como *tagged values* dos novos estereótipos. Assim, surgiu a necessidade de desenvolver uma nova versão da abordagem, que é apresentada no próximo capítulo.

CAPÍTULO 5 - EVOLUÇÃO DA ABORDAGEM PARA CONCEPÇÃO DE *WORKFLOW* CIENTÍFICO

ABSTRATO

Neste capítulo é apresentada a evolução da abordagem para concepção de workflow científico em nível abstrato, decorrente das modificações oriundas das oportunidades de melhoria identificadas durante a aplicação de prova de conceito. Como resultado desta evolução, heurísticas para o uso na abordagem foram criadas, além da adaptação dos instrumentos já existentes. Posteriormente, a abordagem foi aplicada em campo, em um projeto real de pesquisa, no domínio de modelagem offshore para exploração de petróleo (Engenharia Civil). Os resultados obtidos indicaram que a abordagem auxilia na explicitação de detalhes do experimento e na implementação em diferentes SGWfC e infraestruturas.

5.1 Introdução

Conforme apresentado no Capítulo 4, após a aplicação da prova de conceito foram identificadas oportunidades de melhoria em pontos específicos da abordagem, resultando em modificações em sua estrutura. Como uma das consequências, a organização geral da abordagem foi alterada com a introdução do uso de heurísticas para garantia da qualidade. Assim, a abordagem foi reorganizada, resultando na estrutura representada na Figura 5.1. A estrutura é decomposta em três partes, sendo elas: procedimento para composição, instrumentos para captura do *workflow* e conjunto de heurísticas para garantia da qualidade.

O procedimento para concepção não apresentou problemas, contudo devido à introdução de novos instrumentos, duas de suas tarefas sofreram modificações quanto à forma como são executadas. Já nos instrumentos definidos para a abordagem, modificaram-se tanto os formulários quanto o modelo. Na verdade, somente o formulário de artefato foi reorganizado, no qual novos campos foram adicionados ao seu conjunto inicial. Já as modificações nos modelos se deram pela criação de uma nova extensão da notação do Diagrama de Atividades, inserindo-se novos estereótipos e o uso de *tagged values*, e pela inclusão do uso de modelos auxiliares na representação de artefatos e ferramentas, através do Diagrama de Classes.

A nova parte da abordagem é representada pelo conjunto de heurísticas para garantia da qualidade. Estas foram definidas a partir da experiência adquirida na aplicação da abordagem na prova de conceito, além de práticas utilizadas em processos de desenvolvimento da Engenharia de *Software*. As heurísticas são agrupadas em dois grupos, um relativo à organização da tarefa de inspeção e outro à verificação do conteúdo da especificação do *workflow* científico abstrato.

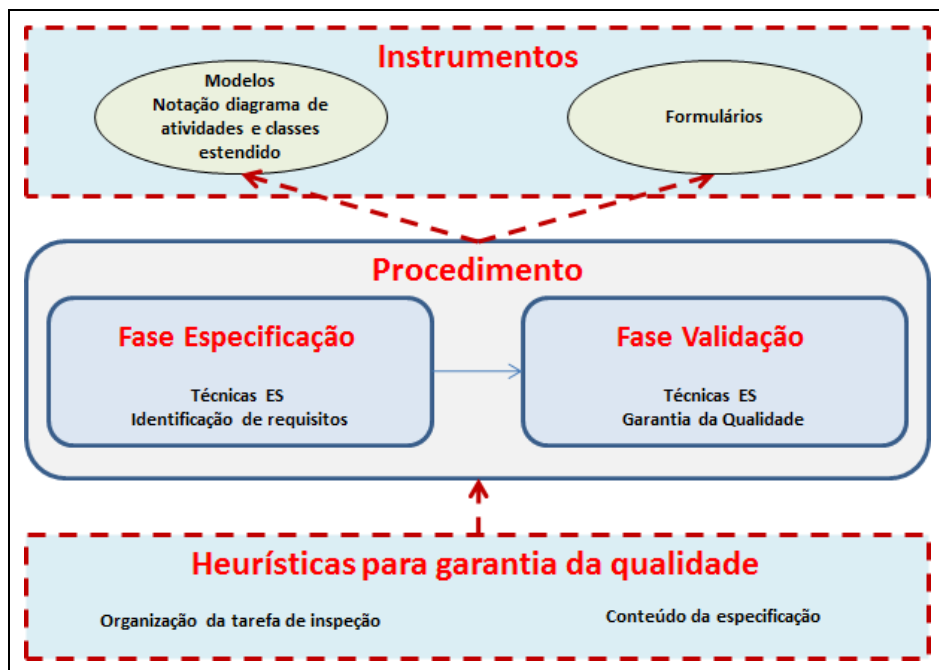


Figura 5.1. Estrutura da abordagem para composição de *workflow* científico

Além desta introdução, este Capítulo é composto pelas seguintes seções: na seção 5.2 estão descritas as modificações no procedimento e nos instrumentos utilizados na concepção; na seção 5.3 estão descritas as heurísticas de garantia da qualidade da especificação de *workflow* científico em nível abstrato; na seção 5.4 é apresentada a aplicação em campo desta versão da abordagem e são discutidos os indícios extraídos desta utilização; na seção 5.5 estão as considerações finais do capítulo.

5.2 Modificações na abordagem para concepção

5.2.1 Modificações no procedimento de concepção

As modificações no procedimento para concepção do *workflow* científico abstrato foram motivadas pela introdução de novos instrumentos. O uso destes durante uma tarefa tem por consequência a alteração na forma como esta é realizada. Contudo, como a introdução de novos instrumentos foi pontual, não foram necessárias

alterações em todas as tarefas, mas somente “Identificar e modelar *workflow* científico” e “Inspeccionar especificação”.

Tarefa “Identificar e modelar *workflow* científico”

A modificação realizada nesta tarefa diz respeito à introdução do uso de diagrama de classes como modelo auxiliar na representação gráfica do *workflow*. Na verdade, o especialista do domínio além de representar as atividades e artefatos na notação estendida do diagrama de atividades, vai representar os artefatos e ferramentas como classes no diagrama de classes. Agora, o analista deve criar dois outros diagramas e anexá-los à especificação de *workflow* científico abstrato. Destaca-se que as informações devem ser extraídas do conhecimento transmitido pelo especialista do domínio também durante as reuniões.

O uso do diagrama de classes se justifica por dois motivos. Em primeiro lugar, com o intuito de evitar redundância, pois existiam informações de artefatos que eram repetidas em várias partes do modelo. Por exemplo, nos elementos *Pin* (notação UML 2) para cada associação a uma ou mais atividades eram inseridas informações como metadado, sinônimo e ferramentas. Estas informações dizem respeito ao artefato como “classe” e não sua “instância” de uso. Em segundo lugar, com o objetivo de registrar também em modelos as informações de ferramentas, já que estas não eram capturadas pelos diagramas de atividades, que não possuem notação gráfica adequada para tal.

Tarefa “Inspeccionar especificação”

A modificação realizada nesta tarefa diz respeito à introdução do uso de heurísticas a serem utilizadas pelos inspetores. Estas heurísticas são na verdade um conjunto de perguntas que guiam o inspetor na verificação de detalhes considerados importantes pela abordagem. Desta maneira, o material distribuído para a inspeção deve ser composto pela especificação de *workflow* científico, um conjunto de heurísticas para guiar a inspeção e uma planilha para relato das discrepâncias encontradas. O analista pode despende um tempo da reunião para apresentação das heurísticas, a fim de mitigar problemas de entendimento.

5.2.2 Modificações na representação textual

Nesta nova versão da abordagem, somente o formulário de artefato sofreu modificação no seu conjunto de campos. A Tabela 5.1 apresenta as modificações efetuadas no formulário artefato e apresenta também as descrições textuais para cada um desses campos.

Tabela 5.1. Novos campos e suas descrições do formulário de Artefato.

Utilização	<p>Informe na tabela abaixo as atividades na qual esse artefato é gerado/usado. Se o artefato é usado pela atividade, esse é um Insumo. Se o artefato é gerado pela atividade esse é um Produto.</p> <p>Para cada caso em que o artefato é usado/gerado por uma atividade, descreva sua condição de obrigatoriedade. Faça isso para cada atividade com a qual esse artefato se relaciona. As classificações possíveis são:</p> <ul style="list-style-type: none"> • Obrigatório: Artefato é um insumo obrigatório de uma atividade, isto é, sempre é consumido; Artefato é produto obrigatório de uma atividade, isto é, sempre é produzido. • Opcional: Artefato é um insumo opcional de uma atividade, isto é, pode ser consumido, dependendo de uma determinada condição; Artefato é produto opcional de uma atividade, isto é, pode ser produzido, dependendo de uma determinada condição. <p>[Informe o formato do artefato (digital, digitalizável ou físico). Digital se é arquivo de computador. Digitalizável se é um documento em papel, que pode ser transformado em um arquivo de computador (através de um scanner). Físico se é uma amostra de material (ex.: rocha, areia, etc.). Caso o Formato do artefato seja digital, pode-se listar a extensão digital do formato.</p>				
	Atividade	Insumo/ Produto	Condição de obrigatoriedade	Formato	Extensão digital
	Nome da Atividade	Insumo ou Produto	Descrição da obrigatoriedade: Obrigatório; Opcional.	Físico, digitalizável, Digital.	Extensão do artefato
Tipos de extensão digital	A lista de todas as extensões digitais possíveis para o artefato.				
Temporalidade	Artefato é um insumo/produto de uma atividade cujo tempo de vida é limitado, isto é, após a execução do experimento científico é descartado. Ex.: O artefato A é Insumo (entrada) da atividade Atv1 e Atv2, mas artefato A é insumo Obrigatório de Atv1 e Opcional de Atv2.				

Os campos nome, descrição, origem, metadado, ferramenta, sinônimos e outros comentários mantiveram-se inalterados, diferente dos campos formato e utilização. Na verdade, percebeu-se que o formato estava relacionado ao uso de artefato numa atividade, por exemplo, Artefato “documento de requisitos” pode estar no formato digital na atividade “especificar requisitos” e no formato físico na atividade “inspecionar requisitos”. A classificação para o campo formato continua a mesma, isto é: digital; digitalizável e físico.

Todavia, as modificações não consistiram somente de reorganização de campos, mas também na definição de novos, sendo eles:

- **Temporário:** Um artefato pode possuir, associado a ele, um tempo de vida determinado ou finito. Existem artefatos que permanecem armazenados mesmo ao final da execução da atividade ou *workflow*. Quando isto ocorre, são considerados permanentes. Entretanto, existem artefatos que só existem temporariamente e acabam descartados ou destruídos ao final da execução da atividade ou *workflow*.
- **Tipos de extensão digital:** Os tipos de extensão digital de um artefato, quando produto/insumo de uma atividade, são registrados neste atributo.

Essa informação permite que posteriormente seja associado o produto/insumo às ferramentas que a suportam, caracterizando possíveis restrições no *workflow* científico concreto.

- **Extensão digital:** se o produto / insumo está no formato digital, este possui extensão digital associada. Pode haver mais de um possível formato para esse Produto / Insumo, e esse se relaciona à Instância do Artefato em uso numa Atividade.

5.2.3 Modificações na representação gráfica do *workflow* científico

As modificações na representação gráfica foram motivadas por duas observações oriundas da aplicação em campo: nem todos os estereótipos representavam semanticamente os elementos do *workflow* científico; e a necessidade de evitar redundância das informações no modelo. Como consequência, foi realizada uma redefinição da extensão da notação do Diagrama de Atividades, inserindo-se novos estereótipos e o uso de *tagged values* para representar as mesmas informações dos formulários. Também, foi proposta a extensão da notação do Diagrama de Classes para representar de artefatos e ferramentas com o intuito de evitar redundância, através da definição de estereótipos e a incorporação de *tagged values* para representar informações estruturais desses elementos. Adicionalmente, outro objetivo foi representar através de modelos os elementos do tipo ferramenta, que inicialmente eram descritos somente através de formulários. Novamente, um perfil UML foi definido para representar a extensão da notação gráfica dos modelos, representado na Tabela 5.2.

Inicialmente, os estereótipos já definidos (vide Tabela 4.5, no Capítulo 4) foram reanalisados, visto que, poderiam não ser aderentes à terminologia usualmente utilizada na área de *workflow* científico. Todavia, do conjunto inicial somente o estereótipo <<*structured*>> se mostrou insatisfatório e, assim, foi alterado, enquanto os demais foram mantidos. Então, para representar uma atividade composta presente num *workflow*, manteve-se o uso do elemento *CallBehaviour Action*, substituindo-se o estereótipo <<*structured*>> por <<*Sub-workflow*>>.

A fim de identificar se um artefato no modelo de *workflow* científico é Produto e Insumo, indicando explicitamente o seu sentido, estes elementos foram mapeados para os elementos *Input pin*, *Output pin* e Parâmetro (*Parameter node*) na notação do diagrama de atividades. Na UML, o elemento Ação Opaca se relaciona com o elemento *Input pin* quando recebe um objeto (dado) como entrada, assim o elemento *Input pin* foi escolhido para representar um Insumo de uma Atividade do *workflow*. O

estereótipo definido foi <<Insumo>>. O elemento Ação Opaca se relaciona com o elemento *Output pin* quando disponibiliza um objeto, no caso como saída da sua execução. Por este motivo o elemento *Output pin* (UML) foi escolhido para representar Produto da atividade no *workflow*. O estereótipo definido foi <<Produto>>. Entretanto, na UML, para representar tanto dados de entrada ou de saída de Atividade é utilizado o elemento Parâmetro (*Parameter node*). Logo, para manter a consistência com a notação, se utiliza esse elemento para representar os Insumo e Produto de um *workflow*. Os estereótipos possíveis para o Parâmetro são <<Insumo>> e <<Produto>>.

Tabela 5.2. Mapeamento final entre elementos do domínio de *workflow* científico para elementos do diagrama de atividades da UML 2.2.

Metamodelo <i>workflow</i> científico	Metamodelo UML	Descrição Mapeamento
<i>Workflow</i>	Atividade (<i>Activity</i>)	<i>Workflow</i> do experimento contendo fluxos de execução. Atividade na UML estereotipada como <i>Workflow</i> .
Manual	Ação Opaca (<i>Opaque Action</i>)	Atividade atômica cuja principal característica é a intervenção do cientista. Ação opaca na UML estereotipada como Manual.
Semi-automatizada	Ação Opaca (<i>Opaque Action</i>)	Atividade atômica cuja característica é a intervenção do cientista com auxílio de uma ferramenta. Ação opaca na UML estereotipada como Semi-automatizada.
Automatizada	Ação Opaca (<i>Opaque Action</i>)	Atividade atômica cuja característica é automação através de uma ferramenta. Ação opaca na UML estereotipada como Automatizada.
Sub- <i>Workflow</i>	<i>Call Behaviour Action</i>	Atividade composta utilizada em um <i>Workflow</i> . <i>Call Behaviour Action</i> da UML com estereótipo Sub- <i>workflow</i> .
Produto / Insumo	Parâmetro (<i>Parameter node</i>)	Produto ou Insumo para um <i>Workflow</i> . Representados como Parâmetro (<i>Parameter node</i>) na UML com estereótipo Produto ou Insumo.
Insumo	Porta (<i>Input Pin</i>)	Insumo de uma atividade. Representados como Input pin da UML com estereótipo Insumo.
Produto	Porta (<i>Output Pin</i>)	Produto de uma atividade. Representados como Output pin da UML com estereótipo Produto.
Artefato	Classe	Artefato é mapeado para elemento Classe na UML, com estereótipo Artefato. Artefatos não aparecem no diagrama de atividades.
Ferramenta	<i>Tagged Value</i> em Ação opaca (<i>Opaque Action</i>)	Ferramenta utilizada em uma atividade. Listada na <i>tagged value</i> nos estereótipos Manual, SemiAutomatizado e Automatizado.
Ferramenta	Classe	Ferramenta é mapeada para elemento Classe na UML, com estereótipo Ferramenta. Ferramentas não aparecem no diagrama de atividades.

Os elementos Ferramenta e Artefato do domínio de *workflow* científico não são mapeados diretamente para elementos do diagrama de atividades da UML. Para o caso de Ferramenta, a justificativa é pelo fato que não existe elemento na UML que represente completamente, já que para uma atividade pode haver zero ou muitas ferramentas passíveis de utilização. Assim, a solução foi criar uma *tagged value* nos estereótipos (Manual, Semi-automatizada, Automatizada) de Ação opaca, para que o relacionamento com as ferramentas seja explicitado. Para completar a modelagem, as

Ferramentas são mapeadas em classes, em um diagrama de Classes (OMG, 2009), com o estereótipo <<Ferramenta>>. Para o caso de Artefato, também não há elemento que represente no diagrama de atividades. Por isso foi mapeado também para classes, em um diagrama de Classes, com o estereótipo <<Artefato>>.

Para representar o transporte de artefatos entre atividades do *workflow*, é usada nesta abordagem a aresta Fluxo de dados da notação UML. Essa aresta foi estendida com estereótipo para caracterizar a conversão de uma instância de Artefato (produto ou insumo), sendo definido como <<Convertido>>. Um Fluxo de dados é estereotipado como <<Convertido>> quando durante o transporte de um determinado Produto ou Insumo, este sofre conversão de extensão digital, mas o seu conteúdo se mantém inalterado. Por exemplo, um determinado Artefato está no formato digital '.xpt' e é convertido para o formato '.xll', mas seu conteúdo se mantém inalterado, esse fluxo dado é estereotipado como <<Convertido>>. Enquanto o estereotipo <<Transformado>> já é definido na própria notação da UML, e é utilizado quando Produto ou Insumo sofre modificações em seus dados, isto é, altera-se a informação contida nele. Por exemplo, quando um arquivo digital tem uma coluna excluída após sua criação, quando um dado científico sofre mudança de unidade, ou se utiliza somente um subconjunto dos dados desse arquivo. Ressalta-se que o uso destes estereótipos para Fluxo de dados não é obrigatório, podendo haver Fluxo de dados no modelo de *workflow* que não estejam estereotipados.

Assim, na Figura 5.2 está representado o perfil resultante a partir do mapeamento dos elementos do domínio de *workflow* científico (vide seção 2.2, no Capítulo 2) para elementos da notação UML 2. Para se organizar este perfil, os relacionamentos e elementos do domínio do *workflow* foram estabelecidos como um modelo estrutural. O resultado desta organização é descrito no Anexo E deste documento, mas este modelo não é parte integrante dos instrumentos da abordagem, sendo somente uma organização da visão desta pesquisa sobre o domínio de *workflow* científico.

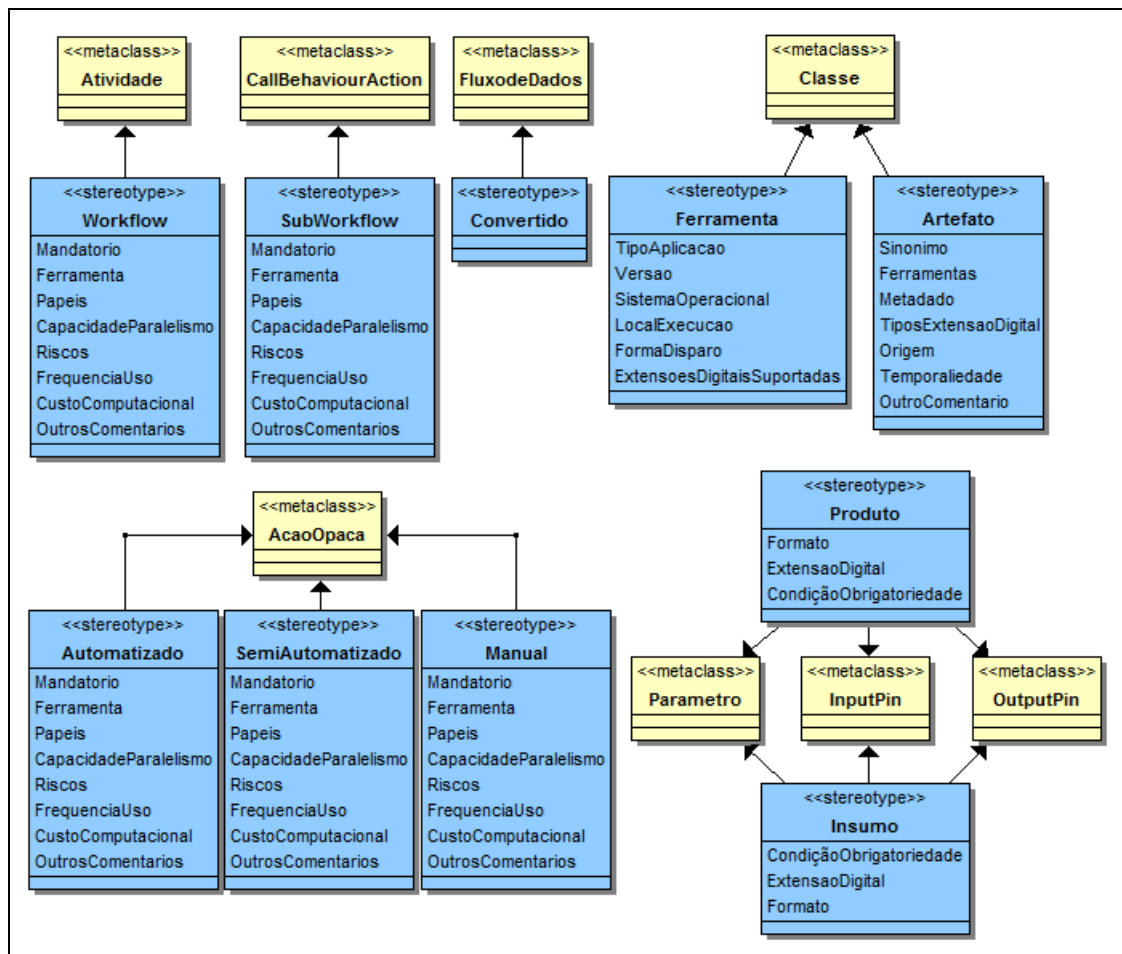


Figura 5.2. Estereótipos e elementos principais do diagrama de atividades UML 2.2.

5.3 Heurísticas para garantia da qualidade da especificação de *workflow* científico

As heurísticas propostas foram identificadas a partir da aplicação da abordagem numa prova de conceito (vide seção 4.4, no Capítulo 4) e, posteriormente, da adaptação de práticas já utilizadas nos processos de *software*, e podem ser, portanto, consideradas como lições aprendidas. Para fins de organização desta seção, o conjunto de heurísticas é dividido em dois (2) grupos: o primeiro relaciona-se à organização da tarefa de inspeção, tratando de alguns assuntos como perfis dos inspetores e organização das tarefas; o segundo grupo relaciona-se ao conteúdo que deve ser inspecionado durante a execução desta tarefa e aborda o uso de modelos hierarquizados de *workflow*.

5.3.1 Heurísticas para organização das tarefas de detecção de defeitos

As tarefas que visam detectar defeitos estão presentes na fase “Validar *workflow* científico” do procedimento para concepção. Esta fase está representada

novamente na Figura 5.3. Durante a aplicação em campo, foram identificadas oportunidades de melhoria para a organização das seguintes tarefas “Inspeccionar especificação”, “Avaliar discrepâncias relatadas” e “Corrigir especificação”, e para o uso de inspetores com perfil distinto dentro destas. Considerando estas oportunidades, heurísticas foram definidas a serem aplicadas nestas tarefas, e tiveram como fonte de conhecimento práticas da Engenharia de *Software*. A seguir as heurísticas são apresentadas e descritas.

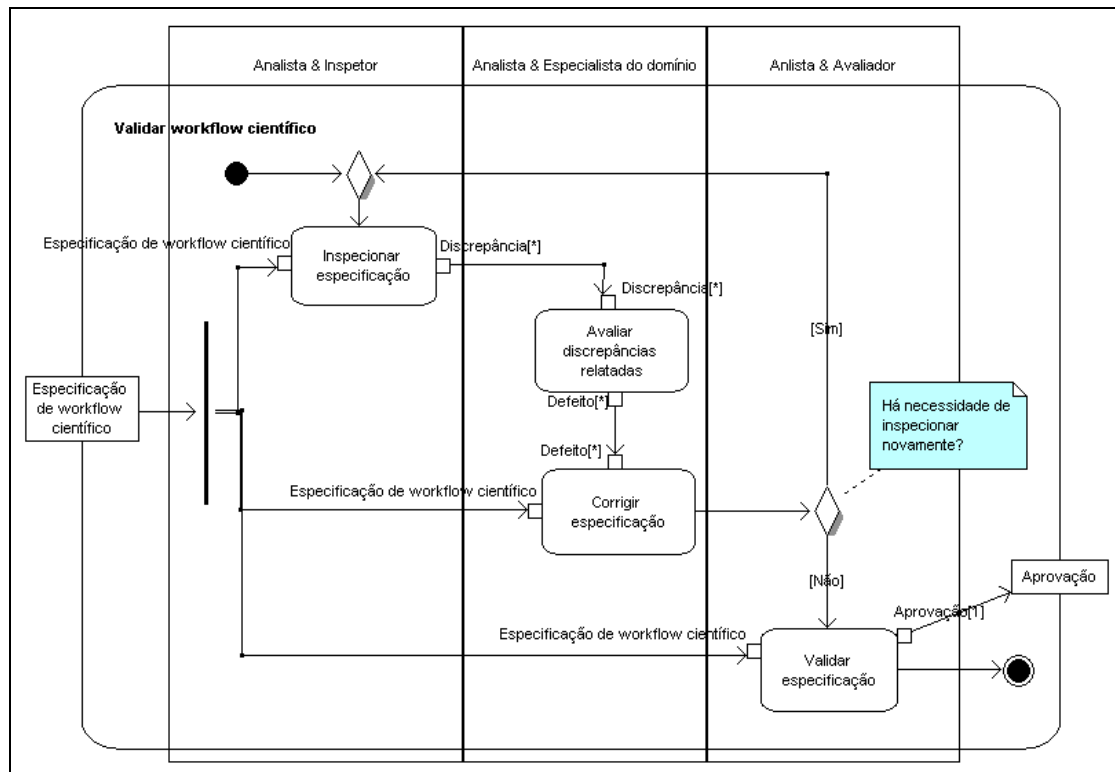


Figura 5.3. Fase de “Validar *workflow científico*” do procedimento proposto nesta dissertação.

Relativo à execução da inspeção no procedimento de concepção:

1. Tarefa “Inspeccionar especificação”:

As inspeções de requisitos de *software*, em geral, não são realizadas pelos criadores da especificação, pois isso pode diminuir o desempenho no seu resultado. Essa premissa de organização é frequente em diversas técnicas, processos e procedimentos para a garantia da qualidade de artefatos de *software*, como exemplo, na revisão técnica por pares (IEEE, 1998), no processo inspeção proposto por SAUER ET. AL. (2000) ou na técnica de *walkthrough* (MELO ET. AL., 2001), dentre outras. Normalmente, os inspetores são analistas que não participaram na identificação e uma de suas metas é verificar se os requisitos são inteligíveis por outros participantes do processo de desenvolvimento. Isso se justifica pelo fato que não são sempre

necessariamente os mesmos indivíduos a utilizarem este documento nas fases posteriores, como por exemplo, de projeto ou teste. O indivíduo que criou o documento pode apresentar maior dificuldade para detectar problemas, pois o documento reflete seu conhecimento. Assim, podem-se utilizar inspetores que não participaram da criação da especificação, buscando-se evitar que defeitos passem despercebidos na inspeção pelos possíveis vieses de leitura.

Outro ponto a ser considerado é o uso de inspetores com níveis de conhecimento distintos sobre o domínio na tarefa de inspeção. Isto é proposto com base na observação realizada durante a aplicação da Prova de Conceito (vide Capítulo 4), na qual os inspetores com diferentes níveis de conhecimento no domínio identificaram diferentes tipos de defeitos. O inspetor mais experiente identificou em sua maioria problemas de omissão, fato incorreto e informação estranha, enquanto o menos experiente relatou descrições deficientes de conceitos e ambiguidade entre termos do domínio do experimento. Isso aconteceu, pois os termos eram familiares para o inspetor experiente no domínio, enquanto o menos experiente não conhecia bem os termos e sentiu necessidade de mais detalhes para compreendê-los. Considerando estas premissas de organização, a heurística para esta tarefa é:

a) Selecionar um ou mais indivíduos do grupo de pesquisa que não participaram da criação da especificação de workflow científico.

a.1) Caso não haja especialistas do domínio disponíveis, selecionar outros indivíduos disponíveis para participar da inspeção.

a.2) Em último caso, não havendo nenhum outro indivíduo disponível, o especialista do domínio que participou da especificação do workflow abstrato pode realizar a inspeção, contudo há risco de perda de eficácia da inspeção.

b) Se houver mais de um especialista disponível, selecionar um mais experiente no domínio e outro menos experiente, a fim de, possivelmente, identificar defeitos de categorias diferentes.

2. Tarefa “Avaliar discrepâncias relatadas”:

Nesta tarefa, cada item da lista de discrepâncias é analisado e é classificado como falso positivo ou defeito pelo Analista e especialista do domínio. Em alguns casos, um mesmo defeito possivelmente é relatado por mais de um inspetor. Quando isto ocorre, para esta tarefa se tornar ágil, pode-se adotar a premissa que estas discrepâncias sejam agrupadas e classificadas como defeitos. Esta premissa é baseada em estudo experimental que indica que o esforço na avaliação das discrepâncias deve ser apenas nos relatos efetuados por um inspetor (LANUBILE e MALLARDO, 2003). A Heurística definida é:

a) Durante a avaliação das discrepâncias, se houver mais de um relato de discrepância efetuado por diferentes inspetores que apontam para um mesmo defeito, separe todos esses relatos relacionados entre si e classifique-os como defeito.

a.1) Agrupe todos os relatos em um único defeito e coloque-o na lista de defeitos a serem corrigidos.

b) Para os relatos de discrepâncias efetuados por um único inspetor, deve-se avaliar se este é um falso positivo ou defeito real.

b.1) Se for classificado como defeito real, anexá-lo a lista de defeitos a serem corrigidos.

b.2) Se for classificado como falso positivo, retirar o relato e eliminar do conjunto discrepâncias.

3. Tarefa “Corrigir especificação”:

Durante a tarefa “Corrigir especificação”, o Analista e Especialista de domínio realizam a correção da especificação de *workflow* científico, devendo registrar as mudanças efetuadas na especificação. Essa prática é utilizada na Engenharia de Software, consistindo no registro de toda a mudança efetuada, a fim de manter rastreabilidade para as alterações (PRESSMAN, 2010). A vantagem obtida é a manutenção do histórico das mudanças e a possibilidade de identificação de causas para os defeitos. Outra prática utilizada na Engenharia de Software é criar versões da especificação a cada alteração completada. Com isso, pode-se também manter a rastreabilidade do documento, conseguindo identificar quais foram as mudanças efetuadas a cada versão (PRESSMAN, 2010). A heurística baseada nestas práticas consiste em:

a) Gerar uma nova versão da especificação para cada rodada de inspeção e validação, que houver necessidade de realizar modificações. Deve-se sempre descrever nela qual foi o motivo do versionamento.

b) Manter uma tabela com cada defeito identificado e as alterações provocadas por sua correção.

5.3.2 Heurísticas para verificação do conteúdo da especificação de *workflow*

Heurísticas para inspeção:

Nesta abordagem, foi definido o uso de inspeção do tipo *ad hoc* com foco na detecção de discrepâncias na especificação. Para realizar a inspeção *ad hoc* não é necessário um treinamento prévio em inspeção e o inspetor está livre para inspecionar o documento da maneira que julgar necessário. Contudo, devido a essa característica,

as inspeções *ad hoc* têm seu desempenho influenciado fortemente pela experiência do inspetor. Neste caso, a verificação do documento é livre, sem direcionamento do inspetor sobre o que deve ser avaliado, portanto algumas questões ou perspectivas podem ser ignoradas. Para tentar minimizar essas influências, podem-se utilizar instrumentos para auxiliar os inspetores na inspeção *ad hoc*. Então, visando um melhor aproveitamento da inspeção, heurísticas oriundas da aplicação de prova de conceito foram definidas (vide seção 4.4, no Capítulo 4), sendo essas descritas a seguir:

a) Verificar se todos os elementos dos tipos: atividade, ferramenta e artefato, presentes nos modelos de workflow científico, diagrama de artefatos e diagrama de ferramentas, possuem um formulário correspondente na especificação.

a.1) Se não existir formulário para um elemento, verificar se esta omissão está justificada na especificação. Não existindo justificativa para omissão, relatar esta falta como defeito.

b) Verificar se o modelo de workflow científico está completo e correto.

b.1) Verificar se nas atividades estão descritos todos seus insumos ou produtos. Se houver algum insumo e/ou produto não descrito neste conjunto, descrever esta falta como um defeito.

b.2) Verificar se estão representadas todas as atividades, manual ou semi-automatizada, do experimento científico. Se houver alguma atividade, manual ou semi-automatizada não descrita na especificação, descrever esta falta como um defeito.

b.3) Verificar se existe algum fluxo de execução que não faça sentido. Se houver fluxos que não fazem sentido ao experimento científico, descrever um defeito.

b.4) Verificar se as atividades citam as ferramentas que utilizam. Se houver alguma ferramenta não descrita, relatar esta falta como um defeito.

b.5) Verificar se todas as ferramentas utilizadas no experimento científico estão presentes no diagrama de ferramentas. Se houver alguma ferramenta faltando no diagrama, descrever esta falta como um defeito.

c) Verificar o conteúdo de cada formulário da especificação.

c.1) Verificar se no formulário de atividades estão descritas as informações de ferramentas associadas, papéis desempenhados, riscos associados, capacidade de paralelismo, tempo de execução e

opcionalidade. Se houver algum campo faltando no formulário, descrever como defeito.

c.2) Verificar se algum artefato, opcional ou com tempo de vida definido, não foi descrito na especificação. Se houver a falta de descrição de algum artefato, opcional ou com tempo de vida definido, descrever um defeito.

O conjunto de heurísticas acima deu origem a 3 (três) categorias de questões a serem utilizadas na inspeção: geral, modelo e formulários. As categorias e suas questões estão representadas na Tabela 5.3. As categorias foram definidas a partir da estrutura da especificação, cuja composição possui uma descrição textual através de formulários e uma representação gráfica através de modelos. Assim, a primeira categoria é composta por questões gerais, que podem envolver tanto os formulários quanto os modelos ou ambos ao mesmo tempo, com objetivo de realizar uma verificação entre as duas formas de representação da especificação. Já a segunda categoria relaciona-se aos modelos utilizados na abordagem, no caso o diagrama de atividades e diagrama de classes. A terceira categoria é composta por questões ligadas aos formulários utilizados pela abordagem.

Tabela 5.3. Heurísticas para inspeção e suas categorias.

Questões para inspeção			Heurística
Geral	G1.	Todos os elementos (atividade, artefato e ferramenta) presentes no modelo possuem um formulário associado? E todos os formulários representam elementos presentes no modelo?	a)
	M1.	Alguma ordem no fluxo de execução de atividades não faz sentido ou não pode ser realizada como a representada no modelo?	b.3)
Modelo	M2.	Existe algum artefato (insumo ou produto) importante e essencial para o estudo que não está mapeado e identificado no modelo? Ele tem uma classe correspondente no diagrama de artefatos?	b.1)
	M3.	Existe algum ponto de decisão ou paralelismo de atividades e artefatos não expresso no documento e que seja importante e essencial para execução do estudo?	b.3)
	M4.	Existe alguma atividade manual ou semi-automatizada necessária para que o estudo experimental seja executado corretamente?	b.2)
	M5.	Existe alguma atividade automatizada que não cita pelo menos uma ferramenta? A ferramenta possui uma classe no diagrama de ferramentas?	b.4) e b.5)
	F1.	Os riscos inerentes a execução de uma atividade estão identificados e documentados?	c.1)
F2.	Os papéis responsáveis pela execução, auxílio ou coordenação das atividades estão identificados?		
F3.	As ferramentas que suportam as atividades automatizadas e semi-automatizadas estão identificadas e possuem formulários associados?		
F4.	A possibilidade de paralelismo da execução de uma atividade está documentada ou explícita?		
F5.	O tempo de execução de uma atividade está descrito e identificado?		
F6.	A opcionalidade de uma atividade está descrita no seu respectivo formulário?		
Formulários	F7.	A opcionalidade ou temporalidade de um artefato está descrita no seu respectivo formulário?	c.2)

Uso de Hierarquia de *workflows*

Alguns autores destacam que na modelagem de *workflow* científico é importante utilizar o conceito de hierarquia de *workflow*. Em VERDI ET AL. (2007) um dos documentos gerados pelo seu processo é o modelo de contexto, cuja característica é apresentar todos os *workflows* do experimento de uma maneira hierárquica. Os *workflows* são dispostos como uma estrutura hierárquica, onde o nó de nível acima é pai dos nós mais baixos, isto é, os *workflows* acima são compostos por outros *sub-workflows* ou atividades. Ao decompor o *workflow* em sub-partes, VERDI ET AL. conseguem criar um *workflow* em mais alto nível, composto por sub-partes representadas através de diagramas mais detalhados, tais como o fluxograma. A modelagem dessa hierarquia é a primeira tarefa a ser realizada na abordagem proposta por VERDI ET AL.

Já LI ET AL. (2006) propõem um método para criação de modelos de *workflow* onde o principal foco é na decomposição hierárquica de serviços. Primeiro, são identificados os principais serviços e, num segundo passo, para cada serviço principal, são identificados os sub-serviços e depois são representados os diferentes fluxos de execução. Uma das vantagens citadas por LI ET AL. é a melhor organização dos modelos gerados, pois são menores e contém menos elementos descritos.

Através da decomposição e hierarquização de *workflow* pode-se diminuir a quantidade de informação nos diagramas significativamente, pois os *workflows* podem apresentar dezenas de atividades interligadas e dezenas de artefatos gerados ou consumidos (insumos e produtos). Assim, na abordagem, ao se utilizar o estereótipo de *Workflow*, dividindo o experimento em várias partes, como consequência os diagramas tendem a ficar mais sucintos e legíveis. Outro ponto positivo é a facilidade em reutilizar o *Workflow* em outros experimentos. Afinal, este *Workflow* pode ser utilizado em outros experimentos que utilizem os mesmos procedimentos que ele representa, somente indicando a Atividade estereotipada *Workflow* e adicionando sua especificação a este novo experimento. Tudo isso é facilitado pela própria notação escolhida, no caso o diagrama de atividades, que possui elementos próprios para representação de atividades e sub-atividades. A heurística para organização hierárquica do *workflow* é:

a) Se o modelo apresentar muitas atividades (12 ou mais atividades), deve-se reorganizar as atividades de forma a criar níveis.

a.1) Agrupar todas as atividades relacionadas entre si por um conceito em um sub-workflow (e.g., atividades que representem passos de um mesmo método ou que possam ser reutilizadas em outros workflows).

a.2) Representar este sub-workflow no nível superior como um único elemento.

a.2) Criar um novo modelo, com todas as atividades agrupadas, e nomeá-la com um nome representativo do conceito.

5.4 Aplicação em Campo – Projeto GALILEU

5.4.1 Domínio de Exploração de Petróleo em Águas Profundas

A avaliação através da aplicação em campo foi realizada em um importante problema do domínio de Engenharia Civil, mais precisamente a exploração de petróleo em águas profundas. Essa área é de intensa investigação, pois se trata de uma importante atividade econômica, na qual o petróleo é elevado de poços na profundidade até milhares de metros na superfície marítima para plataformas, através de estruturas tubulares, denominada como *risers*. Entretanto, essas estruturas podem sofrer avarias ao longo do tempo devido à ação de cargas ambientais, como ondas, ventos e correntes marítimas, e aos efeitos do estresse cíclico ao qual são submetidos. Desta maneira, torna-se crítico analisar a fadiga a qual cada *riser* é submetido, a fim de avaliar quanto tempo a estrutura resistirá e assim planejar para evitar problemas, como rupturas e interrupções na extração do petróleo. Como resultado da complexidade de análise de fadiga, o desenvolvimento e a aplicação rigorosa das ferramentas de análise estrutural se tornou uma tarefa relevante aos engenheiros.

Atualmente, a análise de *risers* requer o uso de diversos programas independentes, que são executados em atividades específicas durante o procedimento de análise. Adicionalmente, a análise lida com uma grande quantidade de arquivos de dados gerados a cada rodada de execução, tornando-a complexa. Além disso, os dados de entrada para um programa geralmente são saída de outro, mas não estão necessariamente no mesmo formato ou local; são necessárias muitas manipulações de dados e transformações através de instruções de linha de comando. Portanto, isso pode tomar muito tempo e predispor a erros, caso seja feito manualmente. Esse cenário torna o experimento difícil de ser executado manualmente e gerenciado.

Considerando o cenário descrito, os projetistas envolvidos na análise de fadiga decidiram utilizar a tecnologia de *workflow* científico e SGWfC para apoiar a execução de seus procedimentos. Inicialmente, eles criaram uma representação, isto é, abstração, como representada na Figura 5.4. Entretanto, essa representação somente contempla parcialmente os procedimentos, pois não apresenta informações como:

- A descrição, pré e pós-condições de cada atividade.

- A identificação dos insumos e produtos para cada atividade.
- A identificação de fluxos alternativos e pontos de decisão.

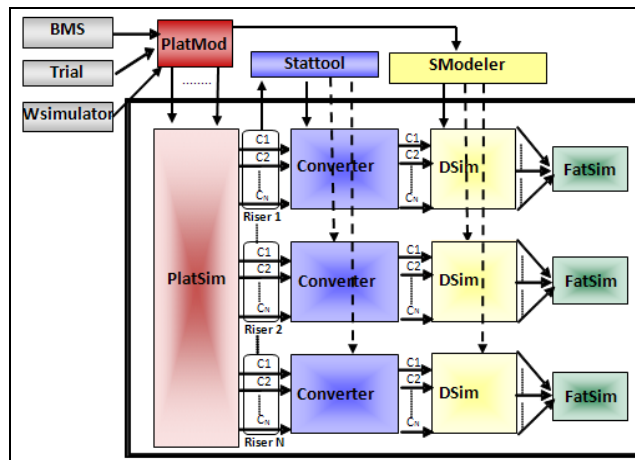


Figura 5.4. Representação dos procedimentos de “Analisar fadiga de risers” (CORREA, 2003).

Além dos problemas citados, ainda há o fato que a representação é ambígua e não é baseada em nenhum padrão. Por exemplo, entre DSim e FatSim, na Figura 5.4, é difícil saber se as arestas ligando-os significam relações de sincronismo ou paralelismo. Todo esse contexto justifica a necessidade de conduzir uma concepção mais detalhada, usando um procedimento mais adequado e que foque na qualidade do *workflow* gerado.

A aplicação em campo faz parte de um projeto de pesquisa denominado Galileu⁶, que estuda novas tecnologias a serem implantadas em uma grande empresa de energia. Por motivos de confidencialidade, algumas informações, como nomes de ferramentas, artefatos e atividades serão modificados para não haver vazamento de informações.

5.4.2 Aplicação do procedimento

Etapa de identificação

O procedimento foi aplicado para a concepção do *workflow* científico abstrato do experimento de “Analisar fadiga de risers”. Para a tarefa de definição do contexto, um documento explicando o problema foi cedido pelo especialista de domínio e a partir dele foi criado o modelo de contexto. O modelo de contexto está representado na Figura 5.5. Esse modelo é similar à Figura 5.4, pelo fato que o especialista do domínio ainda estava focado apenas em explicar o encadeamento de ferramentas, em vez de

⁶ <http://galileu.arquitetura.nacad.ufrj.br/>

descrever o fluxo de atividades e seus conceitos, tanto que no modelo aparece uma atividade “Converter dados da análise acoplada”, que basicamente só transforma um artefato (arquivo digital) de um formato para outro. Desta forma, ainda era necessário subir o nível de abstração e desvincular a descrição do experimento científico dos recursos computacionais utilizados naquele contexto específico. Afinal, o objetivo é descrever uma especificação de *workflow* abstrato, que contém todas as possíveis variações no conjunto de atividades e não somente uma possível combinação de execução.

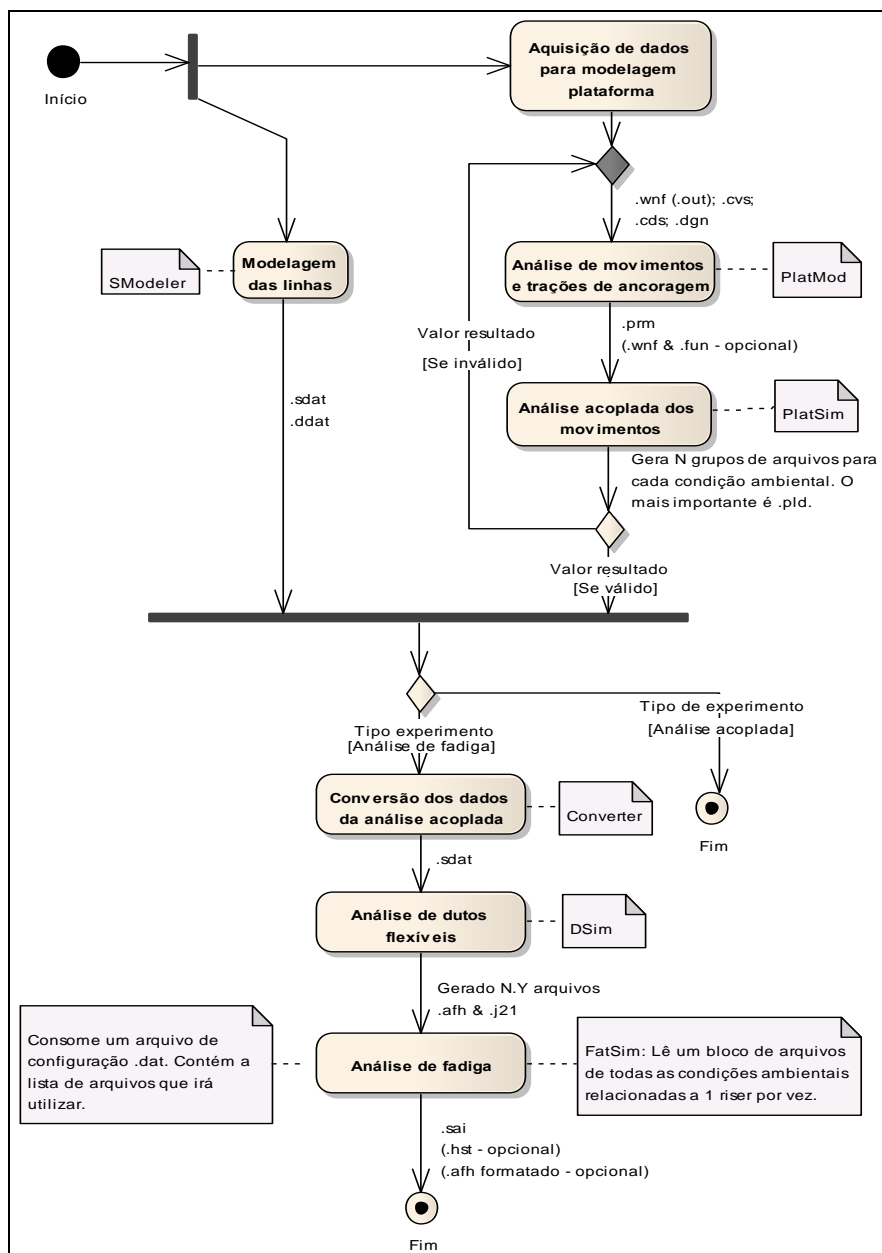


Figura 5.5. Modelo de contexto para experimento in silico Análise de fadiga de risers.

Entretanto, a reunião inicial se mostrou oportuna, pois, ao mesmo tempo em que o analista aprofundou seu conhecimento no domínio, também houve o convencimento do especialista de que o objetivo era capturar o *workflow* abstrato e não somente as ferramentas. De início, o modelo de contexto já apresentava alguns pontos de decisão identificados a partir do documento cedido pelo especialista do domínio.

Com o modelo de contexto criado, foi iniciada a tarefa “Identificação e Modelagem de *workflow* científico abstrato”. Foram realizadas oito reuniões periódicas (utilizando-se a técnica de entrevistas) entre o analista e o especialista do domínio, para que fossem gerados os modelos de *workflow* abstrato e fossem preenchidos os formulários. A Figura 5.6 apresenta dois modelos de *workflow* extraídos da especificação e novamente houve a necessidade de utilizar hierarquia de *workflow* com o objetivo de simplificar a leitura do modelo. Visto que o modelo aumentou seu tamanho e sua capacidade de leitura começou a ser afetada. A Figura 5.6 (a) apresenta o *workflow* abstrato de mais alto nível para “Realizar análise de fadiga de risers” e a Figura 5.6 (b) apresenta um *sub-workflow*, “Analisar movimento da plataforma”.

Seguindo as modificações realizadas na abordagem, mais dois modelos foram gerados, como demonstrado na Figura 5.7. Na Figura 5.7 (a) é representado o extrato do diagrama de ferramentas. Nele foram criadas classes estereotipadas <<Ferramenta>> e foram inseridas todas as informações capturadas pelos seus respectivos formulários. Também foram feitas algumas associações entre essas Ferramentas, como na figura, no qual *PlatMod* utiliza os recursos do *PlatSim* para simulação. Já na Figura 5.7 (b) está representado o extrato do diagrama de artefatos. Nele, os artefatos foram representados como classes estereotipadas com <<Artefato>>, também as informações que diziam respeito a todas as instâncias de um determinado artefato foram inseridas, conforme apresentado na seção 5.2.3.

No total a especificação contou com 12 (doze) atividades, 38 (trinta e oito) artefatos e 7 (sete) ferramentas. Também foram identificados três diferentes *workflows* científicos, que representam variações no fluxo de execução do experimento *in silico*. Um fato de destaque foi a identificação de dois papéis responsáveis pela execução do experimento: o projetista de *risers* e o projetista de ancoragem, que possuíam atribuições distintas e precisavam se comunicar intensamente durante o experimento. O interessante é que isso não estava claro no início, pois não era descrito nem mesmo no documento cedido pelo especialista do domínio. Entretanto para motivos de proveniência é importante, pois define a responsabilidade pela informação gerada ou cedida ao longo da execução do experimento.

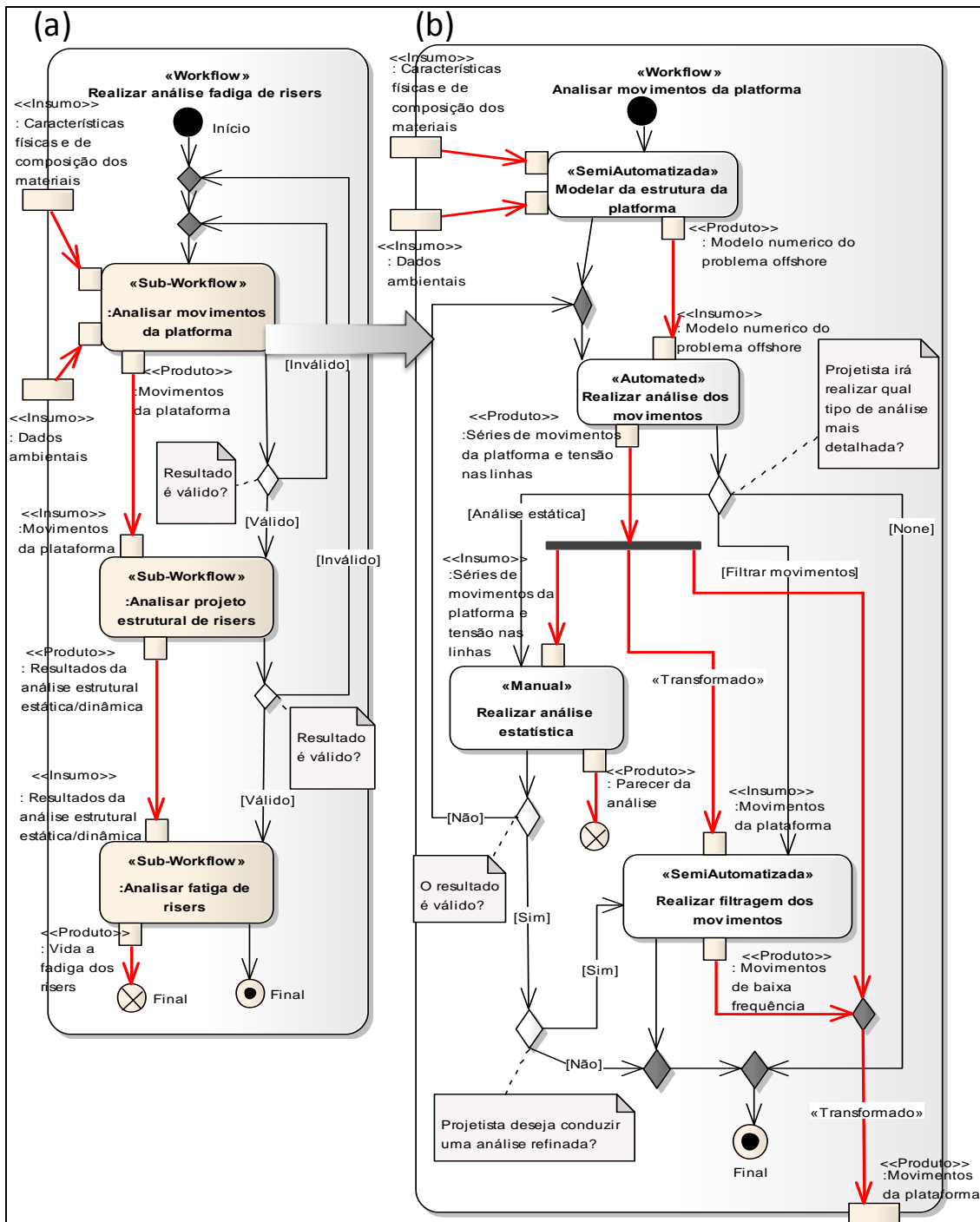


Figura 5.6. Modelos adaptados da especificação. (a) workflow abstrato de “Analisar de fadiga de risers”. (b) modelo da atividade composta “Analisar dos movimentos da plataforma”.

Outro fato era que o especialista do domínio não possuía conhecimento em especificação e notação UML 2. Então, durante as reuniões periódicas, houve a necessidade de reservar tempo para explicar a notação e a própria abordagem. Em geral, as reuniões duravam duas horas, pois era o tempo disponível do especialista do domínio para participar das reuniões e responder as perguntas. Ao final, o especialista

no domínio já entendia a notação UML 2, podendo discutir com o analista os detalhes do modelo.

A introdução das modificações na notação e o uso de *tagged values* para armazenar as informações textuais juntamente com o modelo facilitaram a manipulação da especificação. Contudo, como houve um aumento no número de diagramas e, conseqüentemente, de informações textuais sobre atividades, ferramentas e artefatos, percebeu-se que uma possível melhoria estaria relacionada à automatização de tarefas da abordagem, que incluiriam melhorias na gestão de conteúdos e a inserção automática dessas informações, como insumos, pré-atividades, dentre outras.

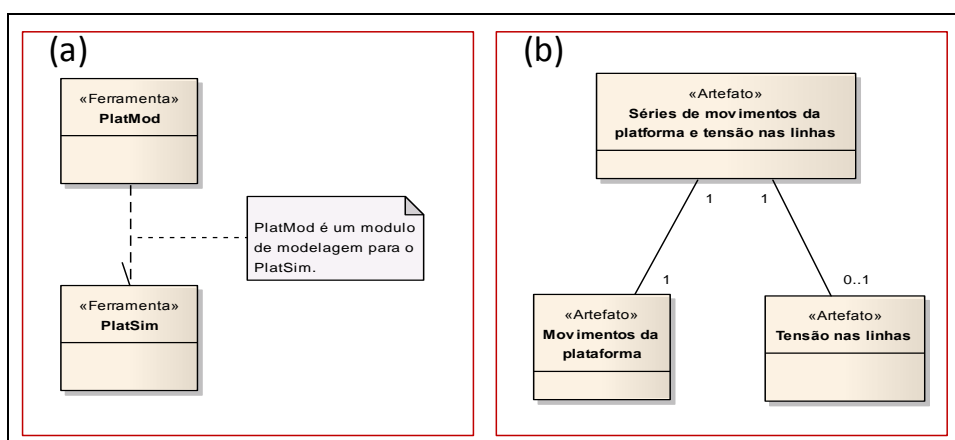


Figura 5.7. (a) Extrato do diagrama de Ferramentas. (b) Extrato do diagrama de Artefatos.

Etapa de verificação e validação

Após a execução da especificação, uma inspeção foi realizada na especificação de *workflow* científico abstrato. É importante ressaltar que na verificação os indivíduos envolvidos com a tarefa de especificação *workflow* não participaram. Apesar de recomendar-se que a inspeção seja realizada por especialistas do domínio, no caso da aplicação em campo, a inspeção foi realizada por cinco (5) engenheiros de *software* com pouco conhecimento do domínio do experimento, que estavam envolvidos no projeto, mas não nas atividades de especificação. O documento foi apresentado aos inspetores durante uma reunião e depois distribuído para cada um, sendo dado o prazo de 1 semana para que a tarefa fosse realizada.

Após este prazo, como resultado da inspeção, 106 discrepâncias foram relatadas na especificação de *workflow*. Após uma primeira avaliação da lista de discrepâncias, 15 falsos positivos foram encontrados, que foram eliminados do conjunto, assim chegando-se a 91 discrepâncias (ainda duplicadas). Depois de uma segunda análise, foram encontradas 10 duplicatas, que foram agrupadas aos seus

respectivos pares, formando uma única descrição de defeito. Com isso, ao final, identificou-se a 81 defeitos no documento. A Tabela 5.4 apresenta a distribuição do número de discrepâncias relatadas por inspetor e o tempo gasto por cada um nesta tarefa. A média de tempo gasto por cada inspetor foi de aproximadamente 107 min. para realizar a tarefa de inspeção. Já na Figura 5.8 apresenta-se a distribuição de defeitos por cada tipo presente na taxonomia adotada nesta abordagem, destaca-se que a maioria defeitos era do tipo omissão (22 do total).

Tabela 5.4. Dados de cada inspetor relacionados ao número de defeitos e tempo gasto na inspeção.

Nº do Inspetor	Tempo gasto (min.)	Nº discrepâncias relatadas	Nº Falso positivos	Desempenho (Nº por mim)
1	120	33	4	0.24
2	90	12	1	0.12
3	170	45	7	0.22
4	51	11	1	0.20
5	105	5	2	0.03
Total:	536	106	15	

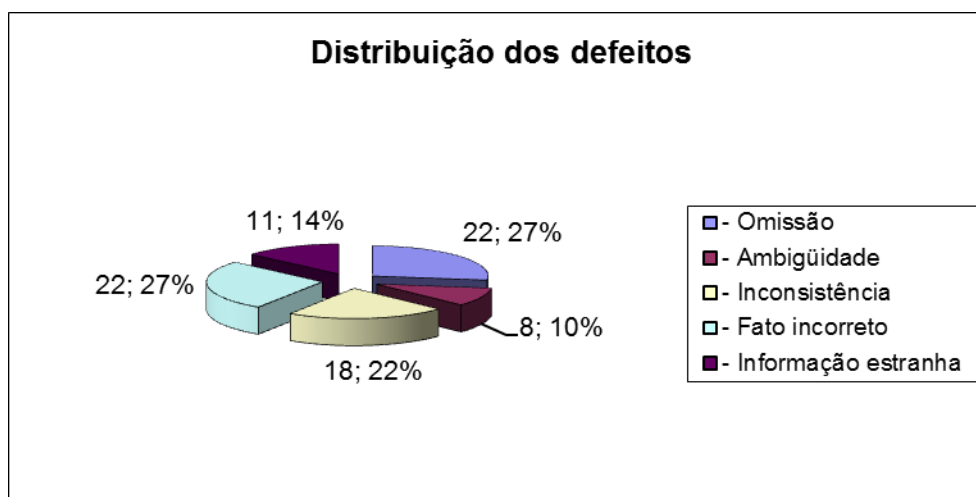


Figura 5.8. Distribuição dos defeitos pelas categorias.

Um fato de destaque foi que três inspetores relataram, em sua maioria, defeitos do tipo omissão que, em geral, relatavam omissões de definições sobre termos muito utilizados na especificação. Estes termos são específicos do domínio do experimento e esse alto número de relatos pode ser reflexo da falta de conhecimento profundo do domínio por parte dos inspetores. Porém isto não desqualifica o relato como defeito. Na verdade, esse fato motivou a adoção de uma taxonomia de termos do domínio e suas definições como parte do *template* de especificação de *workflow* científico abstrato (vide Anexo D). Após isso, a especificação do *workflow* foi corrigida no decorrer de mais cinco reuniões (duas horas cada).

A validação da especificação foi realizada em uma reunião, realizando a leitura dos modelos com a participação de três especialistas do domínio. E por consequência indireta da validação, foram detectados mais quatro defeitos, relacionados a omissão de insumos e fatos incorretos sobre atividades, que foram posteriormente corrigidos e assim a especificação do *workflow* para a análise de fadiga de *riser* foi aprovada. No entanto, durante a validação foi detectado que existiam mais *sub-workflows* que não foram identificados, pois contemplavam outros métodos de análise de fadiga que não faziam parte da metodologia descrita pelo especialista. Isso não caracteriza defeito, pois desde o início era sabido que existiam outras formas de realizar a análise de fadiga, contudo o projeto Galileu, a princípio, focava neste método. Até o momento, esses novos *workflows* científicos abstratos detectados ainda se encontravam em identificação, contudo, o *workflow* concreto para análise de fadiga foi concebido a partir da especificação.

Os defeitos encontrados durante a validação passaram pela inspeção realizada anteriormente, entretanto como existe ainda uma tarefa de garantia da qualidade prevista no procedimento para concepção, mesmo não sendo o objetivo, permitiram a identificação destes defeitos. Contudo, estes poderiam ser pegos na verificação se fosse utilizado, ao invés de inspeção *ad hoc*, técnicas de inspeção mais formal, como, por exemplo, *checklists*. Estes instrumentos seriam adaptados para guiar o inspetor na verificação de questões importantes relacionadas com o domínio do experimento *in silico* ou com a completude da informação. Isso poderia diminuir a influência que a experiência do inspetor exerce na tarefa de verificação proposta nesta abordagem, o que por consequência poderia melhorar seu desempenho, isto é, número de defeitos encontrados.

Etapa de implementação do *workflow* concreto

Uma vez com a especificação validada, foi criado o *workflow* concreto para a análise de fadiga. O projeto do *workflow* concreto foi realizado por outro indivíduo (*implementador*) que não participou da especificação e durou cerca de 20 horas. Esse projeto consistiu na escolha das ferramentas e conversores, além da avaliação das restrições do ambiente. A concepção (*implementação*) do *workflow* concreto foi realizada no Vistrails (CALLAHAN ET AL. 2006) e esta durou 50 horas. A avaliação do *workflow* concreto foi realizada em conjunto pelo *implementador* e o especialista do domínio, executando uma rodada com dados reais, durando cerca de 10 horas.

A Figura 5.9 apresenta três *sub-workflows* gerados a partir da especificação. A Figura 5.9 (a) representa o *workflow* completo, incluindo todas as atividades do experimento de análise de fadiga. A Figura 5.9 (b) representa o *sub-workflow* “Analisar

movimento da plataforma”, que em alguns casos é executado pelo projetista de *risers* separadamente. A Figura 5.9 (c) representa os *sub-workflows* abstratos “Analisar projeto estrutural de *risers*” e “Analisar fadiga de *risers*” que, em geral, são realizados juntos pelo projetista de *risers*.

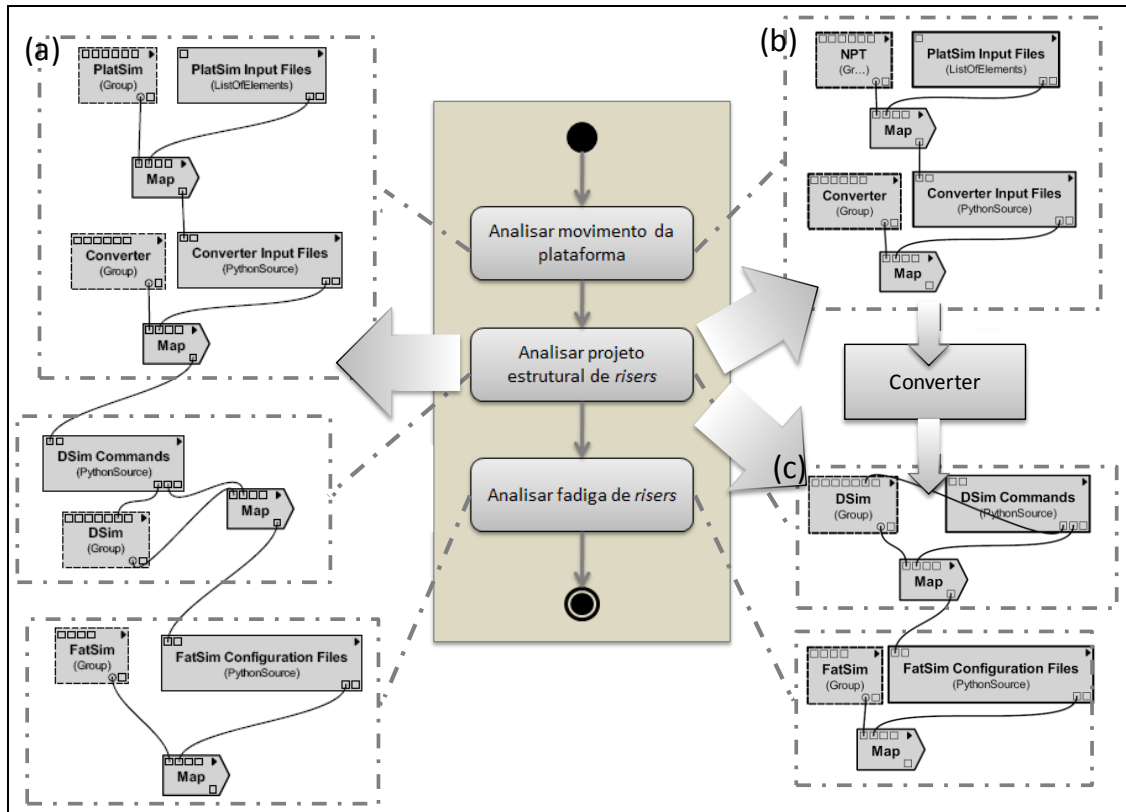


Figura 5.9. Exemplo de *workflows* concretos definidos a partir da especificação de *workflow* abstrato.

Uma oportunidade de pesquisa, identificada nesta aplicação em campo, diz respeito à concepção do *workflow* concreto a partir da especificação de *workflow* abstrato. Atualmente, isso é feito de maneira *ad hoc*, através da transformação das informações capturadas na especificação de *workflow* abstrato para a representação no nível concreto em um SGWfC. Contudo, através do uso de abordagens como geração semi-automatizada de modelos, poder-se-ia a partir das informações já identificadas, criar modelos para diversos SGWfC existentes, somente deixando algumas informações mais específicas da definição do *workflow* concreto em branco para o preenchimento do *implementador* (e.g. definição do IP de um servidor). Acredita-se que com isso, poder-se-ia reduzir o número de defeitos introduzidos na mudança de nível de abstração e também poder-se-ia reduzir o esforço aplicado nesta tarefa. Contudo, existem limitações para essa pesquisa, como as diferentes

linguagens de execução existentes e a quantidade de informações específicas para cada SGWfC, que pode variar de um para outro totalmente.

5.4.3 Análise da aplicação em campo

Novamente, os relatos de uso da abordagem para concepção descreveram-na como uma forma viável de capturar a especificação de *workflow* abstrato, pois os analistas e os especialistas de domínio foram capazes de identificar as atividades e seus objetivos, os artefatos produzidos e consumidos, e as ferramentas e, ao final, concebeu-se uma especificação *workflow* científico. Nesta nova aplicação, além dos formulários e modelos de *workflow*, foram criados diagramas auxiliares contendo os artefatos e ferramentas. Além disso, a notação estendida criada para essa abordagem foi bem recebida pelo especialista do domínio, pois permitia a identificação dos elementos do *workflow* no próprio modelo. A abordagem permitiu identificar e representar os detalhes não habitualmente explicitados quando abordagens *ad hoc* são utilizadas.

Observou-se novamente que é necessário conhecimento sobre modelagem de diagrama de atividades UML para que a abordagem funcione. A aprendizagem sobre a notação requer tempo de alguns participantes, especialmente os especialistas do domínio não familiarizados com conceitos de Engenharia de Software. Portanto, a tarefa "Definir o contexto do *workflow* científico" pode ser utilizada para aprendizagem, porque permite que os pesquisadores não familiarizados com a UML assimilem os conceitos e as técnicas da modelagem. Ainda assim, de acordo com a observação feita por um dos especialistas do domínio: "a especificação é valiosa porque pode apoiar a divulgação do conhecimento a outros pesquisadores (que são novatos no domínio)".

Diferentemente da prova de conceito, esta especificação foi utilizada para concepção do *workflow* científico concreto e a especificação do *workflow* concreto foi utilizada como base e proveram informações importantes acerca do domínio do experimento *in silico*, segundo relato do *implementador*. Assim, indícios da validade da abordagem para concepção da especificação foram identificados, reforçado pelo fato que o *implementador* não participou das tarefas de concepção do *workflow* abstrato.

Apesar dos indícios positivos novamente encontrados sobre a viabilidade e as vantagens da abordagem, devem-se apresentar alguns comentários acerca da aplicação em campo realizada. O contexto no qual a aplicação em campo foi realizada não foi caracterizado. Desta maneira, não se pode afirmar totalmente que não existiu influência de fatores desconhecidos para o sucesso da aplicação da abordagem. Outro

ponto é que os indícios encontrados não podem ser considerados verdadeiros em todos os possíveis contextos, pois a aplicação foi em contexto específico e a generalização desses resultados não é possível. Desta forma, destaca-se que se deve ainda aplicar outros estudos experimentais, incluindo estudos de caso e experimentos controlados, em diferentes domínios, para por a prova esta hipótese.

5.5 Considerações Finais do Capítulo

O uso de experimentos *in silico* tem aumentado em diferentes campos da ciência, incluindo a Engenharia de *Software*. Conforme esses experimentos tornam-se grandes e mais complexos, mais abordagens formais, tais como *workflow* científico, são necessárias. Engenharia de *Software* está aprendendo sobre como usar esse tipo de experimento para obter vantagens, como automação, controle, execução e proveniência dos dados gerados em sua execução e o seu ciclo de aprendizagem. Portanto, foi introduzida uma abordagem destinada a apoiar a concepção de experimentos *in silico* de Engenharia de *Software* e, posteriormente, modificada para atender melhor a esse objetivo. Acredita-se que a concepção dos *workflows* científicos em nível abstrato deve ser independente da SGWfC. Isto pode simplificar a sua replicação e adaptação quando precisar ser executada em diferentes infraestruturas ou *engines* de *workflow* científico.

A abordagem, durante sua aplicação em campo, novamente apresentou indícios positivos sobre sua viabilidade e sobre as vantagens em se formalizar o experimento *in silico* como um *workflow* científico abstrato. De acordo com a observação feita por um dos especialistas do domínio: "a especificação é valiosa porque pode apoiar a divulgação do conhecimento a outros pesquisadores (que são novatos no domínio)". No entanto, observou-se que ainda é necessário conhecimento sobre modelagem de diagramas de atividade UML para que a abordagem funcione. Entretanto com o tempo, o especialista do domínio foi capaz de absorver a notação e passou a debater com o analista sobre a representação do *workflow*. Deve-se destacar que os indícios identificados não podem ser considerados verdadeiros para todos os contextos possíveis e, desta maneira, futuramente deve-se realizar estudos experimentais mais específicos e em ambientes variados a fim de por a prova a hipótese que norteou a criação desta abordagem.

Mesmo com as melhorias efetuadas na abordagem após a aplicação em campo, identificou-se que ainda existiam novas oportunidades de melhorias na sua organização. Uma delas é o uso de técnicas de inspeção mais formais, como, por exemplo, a introdução de *checklists* adaptados para guiar o inspetor na verificação de

questões importantes relacionadas com o domínio de experimento ou completude informação. Outro ponto é a realização de todas as tarefas da abordagem revelou que algumas merecem mais atenção quanto à sua automatização, incluindo melhorias na gestão de conteúdos e inserção automática de informações (por exemplo, insumos e produtos, pré-atividades, entre outros). Com isso espera-se reduzir algumas questões relacionadas com as variações observadas no esforço de modelagem, manuseio e preenchimento dos formulários. Por último, a oportunidade em apoiar a concepção do *workflow* concreto por meio de abordagens semi-automatizadas, como a geração de modelos pré-preenchidos a partir das informações identificadas na concepção do *workflow* abstrato.

CAPÍTULO 6 - CONCLUSÕES

*Neste capítulo é apresentada a conclusão da pesquisa. São descritas as considerações finais a respeito do atendimento a hipótese de pesquisa que norteou este trabalho. Também são apresentados os resultados obtidos que tratam da criação da abordagem, sua aplicação em campo e as publicações resultantes deste esforço. Abordam-se as principais contribuições de pesquisa, que são o procedimento e os instrumentos para concepção de workflow científico para experimentos *in silico*, além da criação de corpo de conhecimento. Por último, apresentam-se as limitações desta pesquisa e as futuras linhas de pesquisa oriundas da criação desta abordagem.*

6.1 Considerações finais

A Engenharia de Software (ES) vem utilizando a Experimentação como instrumento para a criação de um corpo de conhecimento composto por diversos itens, e é através de estudos experimentais que estes itens são avaliados quando relacionados a seus possíveis comportamentos em processos de *software* e como estes impactam no produto gerado. Diferentes tipos de estudos experimentais podem ser utilizados para essa verificação, e, em alguns casos, contam com a participação de profissionais. Contudo, estes estudos apresentam limitações em situações nas quais o tempo para observação do comportamento é demasiadamente longo, pois isto pode inviabilizar a participação de profissionais, tornando-a assim mais difícil e acarretando possivelmente riscos de continuidade da pesquisa. Condições como essa têm motivado o uso cada vez mais freqüente de estudos baseados em simulação na Engenharia de Software Experimental: os experimentos *in virtuo* e, em especial, *in silico*.

Para apoiar os estudos baseados em simulação, esta e outras áreas da ciência vêm utilizando tecnologias como *workflow* científico e Sistemas Gerenciadores de *Workflow* Científico (SGWfC). O *workflow* científico é um modelo que representa a ordem das atividades a fim de realizar um determinado objetivo, sendo este interpretado e executado pelos SGWfCs. Neste contexto, as atividades do *workflow* científico fazem uso, em geral, de ferramentas (sistema ou softwares) durante a sua execução. Os principais benefícios obtidos com o uso desta tecnologia são: registro da

proveniência dos dados; automação da execução do fluxo de atividades; controle e invocação das ferramentas; e manipulação dos dados consumidos e gerados.

Apesar desses benefícios, o uso dessa tecnologia gera novas questões associadas à sua especificação e modelagem. Adicionalmente, experimentos *in silico*, naturalmente, adicionam complexidade a sua realização, pois requerem a definição de infraestrutura mais complexa, bem como maior conhecimento do domínio onde a pesquisa será realizada. Considerando esse contexto, a composição, isto é, a definição de um *workflow* para o experimento *in silico* (e *in virtuo*) se torna não trivial para o pesquisador. De fato, Modelagem Computacional já foi identificada como um dos desafios para computação por diversos grupos de pesquisa, como, por exemplo, a SBC.

Assim, foi realizada uma revisão *ad hoc* da literatura técnica e não foi possível identificar uma abordagem de composição para experimentos *in silico* (e *in virtuo*) que detalhasse as tarefas e definisse meios de garantir a qualidade dos *workflows* gerados, partindo do *workflow* abstrato até o concreto. Desta maneira, buscou-se meios mais formais para a revisão da literatura através da realização de uma *quasi* revisão sistemática, esta indicou que existe a necessidade de: documentação e representação em níveis mais altos de abstração; métodos que auxiliem na concepção do *workflow* científico.

De forma geral, a composição, mais especificamente a concepção, é realizada diretamente no nível concreto e de maneira *ad hoc*, o que pode acarretar riscos para a pesquisa, como perda de informação, dificuldade para reaplicação e falha no registro da proveniência. Adicionalmente, existe a falta de apoio dos SGWfCs para documentação mais detalhada do estudo, não permitindo a especificação de atividades manuais ou semi-automatizadas, e a representação de diferentes fluxos de execução ligados ao mesmo estudo. Por isso, foi proposta uma abordagem que represente o experimento *in silico* em documento mais formal (especificação), garantindo a sua qualidade e a padronização de produtos gerados, e, ao mesmo tempo, sirva como um insumo para a futura derivação do *workflow* abstrato, contido nesse documento, para o concreto.

Como resultado, a formalização do experimento por meio de modelos e formulários permite uma maior exploração da especificação como um importante recurso para implementar experimentos *in silico* (e *in virtuo*) utilizando SGWfC ou infraestruturas computacionais diferentes. Tendo sido a abordagem projetada para ser aplicada pelos engenheiros de *software* ou especialistas de outros domínios (cientistas ou pesquisadores), somente há a restrição de compreender a notação de diagrama de atividades da UML 2. O modelo descrito como um diagrama de atividades permite que

a seqüência de atividades e os dados transmitidos sejam observados, enquanto os formulários são capazes de sintetizar a informação, o que possibilita o acesso rápido aos detalhes do experimento. No entanto, observou-se que ainda é necessário conhecimento sobre modelagem de diagramas de atividade UML para que a abordagem funcione.

Destaca-se que em ambas as aplicações foi relatado que a abordagem para concepção é uma forma viável de capturar a especificação de *workflow* abstrato. A principal justificativa reside no fato de que os analistas e os especialistas de domínio foram capazes de identificar as atividades e seus objetivos, os artefatos produzidos e consumidos, e as ferramentas, criando-se ao final uma especificação *workflow* científico. Mesmo com esses indícios, destaca-se que ainda devem-se aplicar mais estudos experimentais, incluindo estudos de caso e experimentos controlados para por a prova hipótese da pesquisa.

6.2 Resultados obtidos

Considerando os objetivos de pesquisa traçados na seção 1.3, o principal objetivo atingido foi a definição de uma abordagem que auxilie o engenheiro de *software* (ou pesquisador) nas tarefas de concepção e verificação dos requisitos de *workflow* científico para uso em experimentos *in virtuo* e *in silico*. O objetivo foi alcançado pelo atendimento a dois pontos:

1. **Identificação do cenário sobre composição de *workflow* científico:** foi identificado o cenário sobre composição de *workflow* científico para experimentos *in virtuo* e *in silico* através do uso de *quasi*-revisões sistemáticas, formando um corpo de conhecimento atualizável e um protocolo de pesquisa que pode ser reaplicado. Também houve a identificação, através da revisão da literatura, das características que definem os elementos do *workflow* científico;
2. **Definição do procedimento para concepção de *workflow* científico:** foram estabelecidas as tarefas do procedimento para identificação e verificação de requisitos do *workflow* científico e a notação a ser utilizada na representação do modelo de *workflow* científico, através de diagramas de atividades e de classes, e formulários de captura textual das informações sobre os requisitos do *workflow* científico.

Como resultado secundário, obtiveram-se publicações nacionais e internacionais, que divulgaram as descobertas e propostas da pesquisa. A seguir um breve resumo do conteúdo e sua referência bibliográfica:

1. PEREIRA & TRAVASSOS (2009a): Artigo publicado no *workshop* de teses do SBES do ano de 2009. Artigo descreve o escopo inicial da pesquisa, a sua motivação e a proposta de qualificação de mestrado. O escopo inicial descrito é composto, não só pela proposição de procedimento para concepção de *workflow* científico abstrato, mas também para concepção de *workflow* científico em nível concreto. Também foi exposta as ideias iniciais sobre a possibilidade do desenvolvimento de ambiente para apoiar toda a sub-etapa de concepção do ciclo de vida do *workflow*, consistindo em derivação semi-automatizada de modelos para diferentes SGWfCs.
2. PEREIRA & TRAVASSOS (2009b): Artigo publicado no *workshop* e-Science. Artigo descreve a primeira versão da abordagem, contendo os primeiros modelos de formulários de captura textual e a versão inicial da notação estendida de diagrama de atividades. Neste artigo, foram utilizados, como exemplos de aplicação, o experimento *in virtuo* sobre Evolução de Software, apresentando-se formulários preenchidos e modelos de *workflow* abstrato.
3. PEREIRA ET AL. (2009a): Artigo publicado no ESELAW. Artigo descreve a primeira versão da abordagem, contudo o foco principal foi na aplicação da prova de conceito no domínio de Evolução de *software*. Neste artigo, foi relatado os indícios positivos da viabilidade do uso da abordagem, e alguns pontos de melhoria identificados através da aplicação, em especial, no que diz respeito a evolução da notação e dos formulários.
4. PEREIRA ET AL. (2009b): Resumo de uma página e *poster* publicados na conferência e-Science IEEE. Este resumo descrevia a primeira versão da abordagem e exemplos de artefatos gerados com sua aplicação. O exemplo utilizado foi o experimento de modelagem *offshore* para exploração de petróleo (Engenharia Civil), sendo apresentados modelos de *workflow* abstrato e formulários preenchidos.
5. PEREIRA & TRAVASSOS (2010): Artigo publicado na conferência ESEM. Este descreve a segunda versão da abordagem, descrevendo todos os formulários e a notação estendida da para modelagem. Neste artigo, também foi apresentado o estudo em campo no domínio de

exploração de petróleo em alto mar, mais precisamente, foi apresentado o experimento de modelagem *offshore* para exploração de petróleo. O estudo em campo indicou, novamente, a viabilidade do uso da abordagem para concepção de *workflow* abstrato, e permitiu a identificação de oportunidades de pesquisas futuras, como a derivação semi-automatizada de *workflow* concretos e o desenvolvimento de ambiente para apoiar a sub-etapa de concepção de *workflow*.

Outras publicações que envolvem *workflow* científico e tiveram participação do autor, porém não dizem respeito respectivamente à abordagem criada. Sendo estes:

1. MATTOSO ET AL. (2009): Artigo aborda os desafios da computação relacionados à modelagem computacional e ao uso de simulação, em especial a necessidade de novas tecnologias para apoiar essa categoria de experimentos. Neste artigo é rereferenciada a falta de métodos para concepção, em especial, no nível abstrato.
2. MATTOSO ET AL. (2010): Artigo para *Journal* descreve o ciclo de vida do *workflow* científico, definindo e descrevendo as quatro principais etapas: Composição; Execução; Análise; e, Proveniência. Além disso, o artigo apresenta uma revisão dos ciclos de vida já existentes, descrevendo os motivos pelos quais não representam todas as etapas existentes.
3. MELLO ET AL. (2010): Artigo descreve *checklist* para diagrama de atividades, pois argumenta que não existe tecnologias para a garantia da qualidade de modelos deste tipo. Como prova de conceito foi utilizado os modelos concebidos pela abordagem desta dissertação, mais especificamente, a especificação do experimento de modelagem *offshore* para exploração de petróleo.

6.3 Contribuições da pesquisa

A principal contribuição da pesquisa é a abordagem para concepção de *workflow* científico em nível abstrato, composta por procedimento e instrumentos para concepção. Este procedimento compõe o conjunto de tecnologias desenvolvidas pelo Grupo ESE para apoiar a Experimentação na Engenharia de *Software* e, futuramente, essa será parte integrante do ambiente de experimentação eSEE. O ambiente eSEE (TRAVASSOS ET AL., 2008) permite a modelagem do processo de experimentação, a instanciação de um ambiente e organiza conhecimento sobre experimentação,

incluindo a simulação. Por isso, o eSEE pode ser adaptado para incluir o conhecimento sobre *workflow* científicos e permitir a integração dos SGWfCs para apoiar aos experimentos *in silico*. A proposta é que futuramente, o ambiente seja capaz de permitir a concepção do *workflow* em ambos os níveis, e este permita sua execução (utilizando SGWfC) e seu gerenciamento. Além disso, a abordagem estará disponível para o uso por outros grupos, que desejem especificar seus *workflows* de maneira sistemática e obter as vantagens tais como formalização do experimento, sua documentação formal como *workflow* abstrato e a identificação de defeitos.

Outra contribuição é a definição de um protocolo de revisão sistemática que permite a busca de artigos relacionados à composição de *workflow* científico, além do corpo de conhecimento que reflete o cenário da composição de *workflow* científico. Com esse protocolo, qualquer outro pesquisador pode replicar o estudo e verificar se os resultados obtidos são estes e não somente isso, ele também pode atualizar o corpo de conhecimento com novas informações desenvolvidas ao longo do tempo. O corpo de conhecimento apresenta o cenário de composição de *workflow* científico atual (2010), demonstrando em quais pontos estão sendo aplicados os esforços acadêmicos e em quais ainda há oportunidades em aberto para pesquisa.

6.4 Limitações

Mesmo, com as melhorias efetuadas, após a aplicação em campo, identificou-se que a abordagem ainda permite melhorias na sua organização, tais como:

- O uso de técnicas de inspeção mais formais, como *checklist* ou técnicas de leitura, pois a inspeção *ad hoc*, mesmo com a adoção de heurísticas, ainda é suscetível a experiência do inspetor no domínio do experimento *in silico*;
- A automatização de tarefas da abordagem, incluindo melhorias na gestão de conteúdos e inserção automática de informações (por exemplo, insumos e produtos, pré-atividades, entre outros). Mesmo com a adoção de estereótipos e o uso de *tagged values*, ainda há redundância nas informações inseridas nos modelos, que poderiam ser evitadas e assim diminuir a probabilidade de inclusão de defeitos. Com isso espera-se reduzir algumas questões relacionadas com as variações observadas no esforço de modelagem, manuseio e preenchimento dos formulários;
- A abordagem está restrita a concepção de *workflow* científico abstrato, pois apoia somente a especificação através de diagramas e formulários dos requisitos de experimentos *in virtuo* e *in silico*. Desta maneira, a concepção de *workflows* concretos ainda está em aberto, pois não há métodos ou

procedimentos que auxiliem na sua derivação a partir da especificação em alto nível.

A hipótese desta pesquisa considera *ser possível minimizar os riscos envolvidos com a concepção de experimentos in silico (e in virtuo) que utilizam tecnologia de workflow científico, reduzindo problemas de omissão, fato incorreto, ambiguidade, informação estranha e inconsistência nos seus requisitos identificados*. Neste trabalho, a avaliação se deu através de prova de conceito e aplicação em campo que, apesar dos indícios favoráveis, não permitem uma generalização dos resultados para todos os contextos existentes e nem para todos os possíveis casos de uso.

Assim, ainda existe a necessidade de avaliações da abordagem através de estudos experimentais em outros contextos, a fim de por a prova a sua viabilidade e a sua eficácia, conforme proposto pela pesquisa. A partir da hipótese, uma possível questão de pesquisa para nortear estudos experimentais futuros poderia ser definida desta forma:

Q1: O uso de uma abordagem de apoio à definição dos requisitos de *workflow* científico para experimento *in silico* reduz o número de defeitos (omissão, fato incorreto, ambiguidade, informação estranha e inconsistência) inseridos?

Possíveis futuros estudos experimentais para avaliar a questão Q1 seriam:

- *Survey* com especialistas em concepção de *workflow* científico tendo como objetivo avaliar os estereótipos definidos para notação estendida do diagrama, se estes são suficientes ou se é necessário alterar ou acrescentar mais algum. Outra oportunidade seria a avaliação das características listadas (campos) nos três formulários se são suficientes ou se precisam ser alteradas e ou acrescentadas.
- Estudos experimentais controlados, aplicando a abordagem em ambientes controlados, com os fatores de influência devidamente identificados. Dever-se-ia avaliar a eficiência da abordagem, isto é, a redução de defeitos e a aderência das informações identificadas como base para derivação do *workflow* concreto.
- Estudo de caso em projeto com o objetivo de verificar como a abordagem influencia na concepção em ambiente real, tendo como objetivo a geração de *workflow* concreto.

6.5 Futuras linhas de pesquisa

Esta pesquisa vislumbrou três futuras linhas de pesquisa possíveis. Estas estão relacionadas à **derivação semi-automática** de modelos de *workflow* concreto, à **extensão do procedimento para concepção** para que englobe também a concepção do nível concreto; à construção de **ambiente para concepção**.

A derivação semi-automática permitiria a criação de modelos de *workflow* concreto a partir da especificação de *workflow* abstrato, utilizando-se de regras de composição oriundas de práticas e heurísticas aplicadas no desenvolvimento deste tipo de aplicação. Um objetivo seria a diminuição do esforço para composição, pois os modelos seriam populados de maneira prévia e posteriormente completados pelo implementador ou pesquisador. Outro objetivo seria a redução de defeitos oriundos da concepção *ad hoc* atualmente utilizada, pois as regras permitiriam a criação de modelos estruturalmente avaliados. A principal limitação diz respeito a derivação de modelos de *workflow* concretos para todos os sistemas gerenciadores existentes, pois esses possuem linguagens de execução próprias que variam de um para outro e não existe grande correlação entre estas. Abordagens como a GexpLine (OGASAWARA ET AL., 2009), já permitem a derivação semi-automática de *workflows* científicos concretos. Portanto, poder-se-ia adaptar a abordagem desta dissertação para que os modelos abstratos fossem insumo para esta ferramenta, gerando modelos pré-preenchidos numa representação intermediária e posteriormente em nível concreto.

A extensão do procedimento diz respeito ao uso de procedimento para concepção também na criação do *workflow* concreto, pois assim como no nível abstrato, ainda é feita de maneira *ad hoc*. Seria interessante também investir na pesquisa de técnicas de teste para os *workflows* concretos, pois estes já são capazes de serem executados, ou seja, possuem comportamento dinâmico, diferente dos *workflows* abstratos que são essencialmente estáticos.

A construção de um ambiente para concepção do *workflow* científico, que incluía a gestão de conteúdo e a inserção automática de informações. Esse ambiente também poderia realizar a derivação semi-automática do *workflow* concreto, a invocação de SGWfC e o gerenciamento da execução. Novamente, poderia-se utilizar ferramentas como GexpLine na derivação. A integração com o ambiente eSEE para gerenciamento de estudos experimentais, desenvolvido e mantido pelo grupo ESE, levando o apoio aos experimentos *in silico* e *workflow* científico.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRANTES, J. F. E TRAVASSOS, G. H., 2007, *Revisão quasi-Sistemática da Literatura: Caracterização de Métodos Ágeis de Desenvolvimento de Software*. In: Relatório Técnico RT-ES 714/07, Programa de Engenharia de Sistemas de Computação - PESC, COPPE/UFRJ, Rio de Janeiro, Rio de Janeiro, Brasil.
- ALTINTAS, I., BERKLEY, C., JAEGER, E., JONES, M., LUDASCHER, B., E MOCK, S., 2004, "Kepler: an extensible system for design and execution of scientific workflows". In: *Proceedings of the 16th International Conference on Scientific and Statistical Database Management*, pp. 423-424, Santorini Island, Greece, 19 July.
- AMARAL, E. A. G. G., 2003, *Empacotamento de Experimentos em Engenharia de Software*. Dissertação de M.Sc., Programa de Engenharia de Sistemas e Computação – PESC, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
- AMIGONI, F.; FUGINI, M., LIBERATI, D., 2007, "A virtual laboratory for web and grid enabled scientific experiments". In: *Proceedings of 9th International Conference on Enterprise Information Systems*, pp. 227-230, Madeira, June.
- ARAÚJO, M. A., 2009, *Um Modelo para Observação de Evolução de Software*. Tese D.Sc., Programa de Engenharia de Sistemas e Computação – PESC, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
- Baker, M., Buyya, R., and Hyde, D., 1999, "Cluster Computing: A High-Performance Contender", *IEEE Computer*, v. 32, n.7, (July), pp. 79-80, 83.
- BARBARA, D., MEHROTRA, S., RUSINKIEWICZ, M., 1996, "INCAs: Managing dynamic workflows in distributed environments", *Journal of Database Management*, v. 7, n. 1, (Dec.), pp. 5-5.
- BARROS, M.O., 2001, *Gerenciamento de Projetos baseado em Cenários. Uma abordagem de modelagem dinâmica e simulação*. Tese D.Sc., Programa de Engenharia de Sistemas e Computação – PESC, COPPE, Universidade Federal do Rio de Janeiro, RJ, Brasil.
- BASILI, V., CARVER, J., CRUZES, D., HOCHSTEIN, L., HOLLINGSWORTH, J., SHULL, F., ZELKOWITZ, M., 2008, "Understanding the High-Performance-Computing Community: A Software Engineer's Perspective", *IEEE Software*, v. 25, n. 4, pp. 29-36, July/August.

- BIOLCHINI, J., MIAN, P. G., NATALI, A. C. C., TRAVASSOS, G. H., 2005, *Systematic review in software engineering: relevance and utility*. In: Relatório Técnico RT - ES 679 / 05, PESC/COPPE, Universidade Federal do Rio de Janeiro, Brasil.
- BOCK, C., 2003a, "UML 2 Activity and Action Models", *Journal of Object Technology*, v. 2, n. 4, (July-August), pp. 43-53.
- BOCK, C., 2003b, "UML 2 Activity and Action Models Part 2: Actions", *Journal of Object Technology*, v. 2, n. 5 (September-October), pp. 41-56.
- BOCK, C. 2003c, "UML 2 Activity and Action Models Part 3: Control Nodes", *Journal of Object Technology*, v. 2, n. 6 (November-December), pp. 7-23.
- BOCK, C., 2004, "UML 2 Activity and Action Models Part 4: Object Nodes", *Journal of Object Technology*, v. 3, n. 1 (January-February), pp. 27-41.
- BOOSE, E.R. AND ELLISON, A.M. AND OSTERWEIL, L.J. AND CLARKE, L.A. AND PODOROZHNY, R. AND HADLEY, J.L. AND WISE, A. AND FOSTER, D.R., 2007, "Ensuring reliable datasets for environmental models and forecasts", *Ecological Informatics*, v. 2, n. 3, Oct., pp. 237-247.
- BOSE, R., FREW, J., 2005, "Lineage retrieval for scientific data processing: a survey", *ACM Computing Surveys*, v. 37, n. 1, p. 1-28.
- BOSIN, A., DESSÌ, N., FUGINI, M.G., LIBERATI, D., PES, B., 2006, "Applying enterprise models to design cooperative scientific environments", In: *Business Process Management Workshops*, v. 3812/2006, *Lecture Notes in Computer Science*, Springer, pp. 281-292, 2006.
- CALLAHAN, S. P., FREIRE, J., SANTOS, E., SCHEIDEGGER, C. E., SILVA, C. T. E VO, H. T., 2006, "VisTrails: visualization meets data management" In: *ACM SIGMOD International Conference on Management of Data*, pp. 745-747, Chicago, Illinois, USA, June 26-29.
- CHARD, K., ONYUKSEL, C., WEI TAN, SULAKHE, D., MADDURI, R., FOSTER, I., 2008, "Build Grid Enabled Scientific Workflows Using gRAVI and Taverna", Fourth IEEE International Conference on eScience, Indianapolis, Indiana, USA, pp. 614 – 619, 7-12 Dec. 2008.
- CHAPETTA, W. A., 2006, *Uma Infra-estrutura para Planejamento, Execução e Empacotamento de Estudos Experimentais em Engenharia de Software*. Dissertação de M.Sc., Programa de Engenharia de Sistemas e Computação – PESC, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
- CORREA, F.N., 2003. *Application of Hybrid Methodologies in Parametric Studies about the behavior of Offshore Systems*. Dissertação de M. Sc., Programa de Engenharia Civil, COPPE/UFRJ, Rio de Janeiro, Rio de Janeiro, Brasil.

- CRA, 2005, *Conference on Grand Research Challenges in Computer Science and Engineering*. Disponível em: <http://archive.cra.org/Activities/grand.challenges/architecture/home.html/>. Acesso em: 08 de Nov. de 2010.
- DAVIS, A., DIESTE, O., HICKEY, A., JURISTO, N., MORENO, A. M., 2006, “Effectiveness of Requirements Elicitation Techniques: Empirical Results Derived from a Systematic Review”. *Requirements Engineering, IEEE International Conference on*, 14th IEEE International Requirements Engineering Conference (RE'06), pp. 179-188, Minneapolis/St. Paul, Minnesota, USA, 11-15 Sep.
- DEELMAN, E., GANNON, D., SHIELDS, M. E TAYLOR, I., 2009, “Workflows and e-Science: An overview of workflow system features and capabilities”, *Future Generation Computer Systems*, v. 25, n. 5 (May 2009), pp. 528-540.
- DE ROURE, D., GOBLE, C. AND STEVENS, R., 2009, “The Design and Realisation of the myExperiment Virtual Research Environment for Social Sharing of Workflows”, *Future Generation Computer Systems*, v. 25, pp. 561-567. doi:10.1016/j.future.2008.06.010
- ELLISON, A. M., OSTERWEIL, L. J., CLARKE, L., HADLEY, J. L., WISE, A., BOOSE, E., FOSTER, D. R., HANSON, A., JENSEN, D., KUZEJA, P., RISEMAN, E., SCHULTZ, H., 2006, “Analytic webs support the synthesis of ecological data sets”, *Ecology*, v. 87 n.6 (Jun), pp. 1345–1358.
- FAGAN, M. E., 1976, “Design and Code Inspection to reduce erros in program development”, *IBM Systems Journal*, v. 15, n. 3, pp. 182-211.
- FAHRINGER, T., PLLANA, S., VILLAZON, A., 2004, “AGWL: Abstract Grid Workflow Language”. In: *International Conference on Computational Science, Programming Paradigms for Grids and Meta-computing Systems*, Krakow, Poland, Springer, Berlin, Heidelberg, New York, June.
- FOSTER, I., KESSELMAN, C., 1999, *The Grid: Blueprint for a future computing infrastructure*. 1ed., Morgan Kaufmann, USA.
- FOWLER, M., 2003, *UML Distilled: A Brief Guide to the Standard Object Modeling Language*. 3 ed. Boston, Addison-Wesley Professional.
- FREIRE, J., KOOP, D., SANTOS, E., SILVA, C. T., 2008. “Provenance for Computational Tasks: A Survey”, *Computing in Science and Engineering*, v. 10, n. 3, pp. 11-21.
- GANNON, D., PLALE, B., MARRU, S., KANDASWAMY, G., SIMMHAN, Y., SHIRASUNA, S., 2007. “Dynamic, Adaptive Workflows for Mesoscale Meteorology”. In: *Workflows for e-Science*, Springer, pp. 126-142.

- GIL, Y. 2007. "Workflow Composition: Semantic Representations for Flexible Automation". In: *Workflows for e-Science*, 2007, Part II, pp. 244-257.
- GIL, Y., RATNAKAR, V., DEELMAN, E., MEHTA, G., KIM, J., 2007a, "Wings for Pegasus: Creating large-scale scientific applications using semantic representations of computational workflows". In: *Proceedings of the Nineteenth Conference on Innovative Applications of Artificial Intelligence (IAAI-07)*, pp. 22–26, Vancouver, British Columbia, Canada, July.
- GIL, Y., GONZÁLEZ-CALERO, P. A., AND DEELMAN, E., 2007b. On the black art of designing computational workflows. In: *Proceedings of the 2nd Workshop on Workflows in Support of Large-Scale Science*, pp. 53-62, Monterey, California, USA, June 25 - 25. ACM, New York, NY, 53-62.
- GODERIS, A., DE ROURE, D., GOBLE, C., BHAGAT, J., CRUICKSHANK, D., FISHER, P., MICHAELIDES, D., TANO, F., 2008. Discovering Scientific Workflows: The myExperiment Benchmarks. In: *IEEE Transactions on Automation Science and Engineering*, pp. 12, April.
- HOLLINGSWORTH, D., 1994, *Workflow management coalition: The workflow reference model*. Document TC00-1003, Workflow Management Coalition, Dec. 1994. Disponível em: < <http://www.wfmc.org/standards/docs/tc003v11.pdf>>. Acesso em: 20 de Nov. 2010.
- HULL, D., WOLSTENCROFT, K., STEVENS, R., GOBLE, C., POCOCK, M.R., LI, P., OINN, T., 2006. "Taverna: a tool for building and running workflows of services", *Nucleic Acids Research*, vol. 34, n. 2, pp. 729-732 (1 July 2006).
- HUMPHREY, W. S., 1989, *Managing the Software Process*. 1 ed., Pittsburgh, PA, Reading. Addison-Wesley Publishing Co.
- IENDRIKE, H.D.S., 2003, *Métodos para Projeto de Workflow a partir do Modelo de Negócio de Organização*. Dissertação Msc, NCE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
- IEEE 610.12, 1990, "*IEEE Standard Glossary of Software Engineering Terminology*". IEEE Std 610.12-1990, vol., no., pp.1, 1990. doi: 10.1109/IEEESTD.1990.101064
- IEEE, 1998, "*IEEE standard for software reviews*". IEEE Std 1028-1997 , vol., no., pp.360 (viii+37), 4 Mar.
- INGALLS, R. G., 2002, "Introduction to Simulation", *Winter Simulation Conference Proceedings*, pp. 7-16, Arlington, VA, USA, December 09-December 12.
- JABLONSKI, S.; VOLZ, B., REHMAN, M., 2007, "A conceptual modeling and execution framework for process based scientific applications". In: *ACM 1st Workshop on CyberInfrastructure: Information Management in eScience, CIMS '07, Co-*

- located with the 16th ACM Conference on Information and Knowledge Management, CIKM 2007, November 6-9, 2007, Lisboa, Portugal, pp. 23-30
- JACOBSON I, ERICSSON M, JACOBSON A., 1994, *The Object Advantage: Business Process Reengineering With Object Technology*. Addison-Wesley Professional.
- JURISTO, N., MORENO, A. M., 2001, *Basics of software engineering experimentation*. 1 ed., AA Dordrecht, Netherlands, Kluwer Academic Publishers.
- KAESTLE, G., SHEK, E. C., DAO, S. K., 1999, "Sharing experiences from scientific experiments". *Eleventh International Conference on Scientific and Statistical Database Management*, pp. 168-177, Cleveland, OH, USA, 28-30 Jul.
- KALETAS, E.; AFSARMANESH, H., HERTZBERGER, L., 2004, "A methodology for integrating new scientific domains and applications in a virtual laboratory environment". In: *ICEIS 2004 - Proceedings of the Sixth International Conference on Enterprise Information Systems*, Porto, Portugal, April 14-17, 2004, pp. 265-272.
- KALETAS, E.; AFSARMANESH, H. & HERTZBERGER, L., 2003, "Modelling multi-disciplinary scientific experiments and information". In: *Computer and Information Sciences, - ISCIS 2003 Lecture Notes in Computer Science*, 2003, Volume 2869/2003, 75-82, Springer.
- KALINOWSKI, M., 2004, *Infra-estrutura computacional de apoio ao processo de inspeção de software*. Dissertação M.Sc. em Engenharia de Sistemas e Computação, PESC/COPPE, Universidade Federal do Rio de Janeiro, RJ, Brasil, p. 109.
- KALYANAM, R., ZHAO, L., PARK, T., GOASGUEN, S., 2007, "A Web Service-Enabled Distributed Workflow System for Scientific Data Processing". In: *Future Trends of Distributed Computing Systems, 2007. FTDCS '07. 11th IEEE International Workshop on DOI - 10.1109/FTDCS.2007.9*, 2007, pp. 7-14.
- KAVANAGH, J, HALL, W., 2008, "Grand Challenges in Computing Research 2008", Disponível em: <http://www.ukcrc.org.uk/grand_challenges/news/gccr08final.pdf>. Acesso em: 11 de Nov. 2010.
- KIM, J., GIL, Y., RATNAKAR, V., 2006, "Semantic Metadata Generation for Large Scientific Workflows". In: *The Semantic Web - ISWC 2006, Lecture Notes in Computer Science*, 2006, v. 4273/2006, pp. 357-370. doi: http://dx.doi.org/10.1007/11926078_26.
- KITCHENHAM, B., 2004, *Procedures for performing systematic reviews*. In: TR/SE-0401/NICTA Technical Report 0400011T, v. 1, p. 33, Keele University.

- KITCHENHAM, B., PEARLBRETON, O., BUDGEN, D., TURNER, M., BAILEY, J., LINKMAN, S., 2009, "Systematic literature reviews in software engineering – A systematic literature review", *Information and Software Technology*, v. 51, n. 1, pp. 7-15. Elsevier B.V. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0950584908001390>>.
- LACROIX, Z.; LEGENDRE, C. & TUZMEN, S., 2009, "Reasoning on scientific workflows". In: *SERVICES 2009 - 5th 2009 World Congress on Services*, 2009, pp. 306-313.
- LANUBILE, F., MALLARDO, T., 2003, "An Empirical Study of Web-Based Inspection Meetings". *Proceedings of the 2nd International Symposium on Empirical Software Engineering*, Roman Castles (Rome), Italy, IEEE Computer Society Press, September/October.
- LAUESEN, S. 2002. *Software Requirements: Styles and Techniques*. 1ed, USA, Addison-Wesley.
- LI, Q., SHAN, Z., HUNG, P. C. K., CHIU, D. K. W., CHEUNG, S. C., 2006, "Flows and views for scalable scientific process integration". In: *INFOSCALE '06. Proceedings of the First International Conference on Scalable Information Systems*, May 29-June 1 2006, Hong Kong.
- LIU, D., WANG, Q., XIAO, J., 2009, "The Role of Software Process Simulation Modeling in Software Risk Management: A Systematic Review". In: *Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement*, pp. 302 – 311, Florida, 15-16 Oct. 2009.
- LIU, X., DOU, W., CHEN, J., FAN, S., CHEUNG, S. C., CAI, S., 2007, "On design, verification, and dynamic modification of the problem-based scientific workflow model", *Simulation Modelling Practice and Theory*, v. 15, n. 9, (October) 2007, pp. 1068-1088.
- LIU, X., DOU, W., FAN, S., CAI, S., 2006, "The problem-based scientific workflow design and performance in Grid environments". In: *Proceedings of the Fifth International Conference on Grid and Cooperative Computing Workshops (GCCW'06)*, pp, 267 – 274, Hunan, China, October 21-October 23.
- LIVNY, I., IOANNIDIS, Y., LIVNY, M., HABER, E., MILLER, R., TSATALOS, O., WIENER, J., 1994, "Desktop Experiment Management". *IEEE Data Engineering Bulletin*, v. 16, n.1, March 1993.
- LUDASCHER, B., ALTINTAS, I. E GUPTA, A., 2003, "Compiling abstract scientific workflows into Web service workflows" In: *Conference on Scientific and Statistical Database Management*, 2003. 15th International DOI -

- 10.1109/SSDM.2003.1214990, 2003, pp. 251-254, Cambridge, MA, USA, 9-11 July 2003.
- MAFRA, S. N., TRAVASSOS, G. H., 2005, “Técnicas de Leitura de Software: Uma Revisão Sistemática”. In: *Proceedings of XIX SBES*, Uberlândia, MG, Brasil, p. 16.
- MARTINHO, W., TRAVASSOS, G. H., 2010, “Towards the conception of scientific workflows for in silico experiments in software engineering”. In: *International Symposium on Empirical Software Engineering and Measurement, Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*, 2010, Bolzano-Bozen. p. 1-4.
- MATTOS, A., SILVA, F.J.C., RUBERG, N., DA CRUZ, S.M.S., MATTOSO, M., 2008, *Gerência de Workflows Científicos: uma análise crítica no contexto da bioinformática*. In: Relatório Técnico, PESC/COPPE, Universidade Federal do Rio de Janeiro. Disponível em:<<http://www.cos.ufrj.br/uploadfiles/1204740464.pdf>>. Acesso em: 10 Out. 2010.
- MATTOSO, M. E DA CRUZ, S.M.S. 2008. “Gerência de workflows científicos: Oportunidades de pesquisa em bancos de dados”. In: *Proceedings of the 23rd Brazilian symposium on Databases*, Application track: Tutorials table of contents, Campinas, Sao Paulo, Brazil, pp. 313-314.
- MATTOSO, M. L. Q., WERNER, C. M. L., TRAVASSOS, G. H., BRAGANHOLO, V., MURTA, L., OGASAWARA, E., OLIVEIRA, D; OLIVEIRA, F.T., PEREIRA, W. M., “Desafios no apoio à composição de experimentos científicos em larga escala”. In: *CSBC- SEMISH 2009*, CSBC- SEMISH Seminário Integrado de Software e Hardware, 2009, Bento Gonçalves, Brasil, pp. 307-321.
- MATTOSO, M., WERNER, C., TRAVASSOS, G., BRAGANHOLO, V., MURTA, L., OGASAWARA, E., OLIVEIRA, D., CRUZ, S., MARTINHO, W., 2010, “Towards Supporting the Life Cycle of Large Scale Scientific Experiments”, *International Journal of Business Process Integration and Management*, v. 5, n. 1, pp. 79–92.
- MATTOSO, M., WERNER, C., TRAVASSOS, G.H., BRAGANHOLO, V. MURTA, L., 2008. “Gerenciando Experimentos Científicos em Larga Escala”. In: *SEMISH – CSBC 2008*, Belém, Pará - Brasil, 2008, pp. 121-135.
- MEDEIROS, C. B., PEREZ-ALCAZAR, J., DIGIAMPIETRI, L., PASTORELLO JR., G. Z., SANTANCHE, A., TORRES, R. S., MADEIRA, E., BACARIN, E., 2005, “WOODSS and the Web: Annotating and reusing scientific workflows”, *Sigmod Record*, v. 34, n. 3, Sept. 2005 (on line).

- MELLO, R. M., MARTINHO, W., TRAVASSOS, G. H., 2010, "Inspection of the Activity Diagrams of the Requirements Specification". In: *Brazilian Symposium on Software Engineering (SBES)*, 2010, pp. 168-177, September 27-October 01, Salvador - Brazil.
- MELO, W., SHULL, F., TRAVASSOS, G. H., 2001, *Software Review Guidelines*. In: Technical Report ES-556/01, Systems Engineering and Computer Science Program – PESC, COPPE/UFRJ, Rio de Janeiro, Brasil.
- NITRD, 2005, *Computational Science - ensuring America's Competitiveness*. Disponível em: http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf Acesso em: 03 de Dez. de 2010.
- OASIS, 2007, *Business Process Execution Language. Versão: 2.0*. Disponível em: <http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.pdf>. Acesso em: 08 de Nov. de 2010.
- OBJECT MANAGEMENT GROUP, 2009, *OMG Unified Modeling Language Specification*. Disponível em: <http://www.omg.org/spec/UML/2.2/>. Acesso em: 10 de Out. de 2010.
- OBJECT MANAGEMENT GROUP, 2010a, *Object Constraint Language Specification*. Disponível em: <http://www.omg.org/spec/OCL/2.2/>. Acesso em: 03 de Out. de 2010.
- OBJECT MANAGEMENT GROUP, 2010b, *OMG Business Process Model and Notation. Versão: 2.0*. Disponível em: <http://www.omg.org/spec/BPMN/2.0/Beta2/>. Acesso em: 08 de Nov. de 2010.
- OGASAWARA, E., PAULINO, C., MURTA, L., WERNER, C., MATTOSO, M., 2009, "Experiment Line: Software Reuse in Scientific Workflows". In: *Lecture Notes in Computer Science - Proceedings of the 21th International Conference of Scientific and Statistical Database Management*, pp. 264–272, 5566 LNCS, New Orleans, LA, USA.
- OINN, T., ADDIS, M., FERRIS, J., MARVIN, D., SENGER, M., GREENWOOD, M., CARVER, T., GLOVER, K., POCOCK, M. R., WIPAT, A., LI, P., 2004, "Taverna: A tool for the composition and enactment of bioinformatics workflows". *Bioinformatics*, v. 20, n. 17 (2004) pp. 3045–3054.
- OLIVEIRA, D., CUNHA, L., TOMAZ, L., PEREIRA, V., MATTOSO, M., 2009, "Using Ontologies to Support Deep Water Oil Exploration Scientific Workflows". In: *IEEE International Workshop on Scientific Workflows*, pp. 364-367, LA, California, USA, July 06-July 10.

- OSTERWEIL, L.; CLARKE, L.; ELLISON, A.; BOOSE, E.; PODOROZHNY, R. & WISE, A., 2010, "Clear and precise specification of ecological data management processes and dataset provenance". *IEEE Transaction on Automation Science and Engineering*, 2010, v. 7, n.1, pp. 189-195
- OSTERWEIL, L.; CLARKE, L.; ELLISON, A.; PODOROZHNY, R.; WISE, A.; BOOSE, E. & HADLEY, J., 2008, "Experience in using a process language to define scientific workflow and generate dataset provenance". In: *16th ACM SIGSOFT International Symposium on the Foundations of Software Engineering, SIGSOFT 2008/FSE-16*, 2008, PP. 319-329, ATLANTA, GEORGIA.
- PAI, M., MCCULLOCH, M., GORMAN, J.D., PAI, N., ENANORIA, W., KENNEDY, G., THARYAN, P., COLFORD, J.M. JR., 2004, "Systematic reviews and meta-analyses: an illustrated, step-by-step guide". *The National Medical Journal of India*, v. 17, n. 2, Mar-Apr, pp. 86-95.
- PENNINGTON, D. D., HIGGINS, D., PETERSON, A. T., JONES, M. B., LUDÄSCHER, B., BOWERS, S., 2007, "Ecological Niche Modeling Using the Kepler Workflow System". In: *Workflows for e-Science 2007, Part I*, Springer, pp. 91-108.
- PEREIRA, W. M., TRAVASSOS, G. H., 2009a, "Abordagem para concepção de workflows científicos em experimentos in virtuo e in silico em Engenharia de Software". In: *14º Workshop de Teses e Dissertações em Engenharia de Software (WTES 2009)*, pp. 36-41, 2009, Fortaleza, Brasil.
- PEREIRA, W. M., TRAVASSOS, G. H., 2009b, "Abordagem para concepção de experimentos científicos em larga escala suportados por workflows científicos". In: *III e-Science Workshop*, pp. 25-32, 2009, Fortaleza, Brasil.
- PEREIRA, W. M., ARAUJO, M. A. P., TRAVASSOS, G. H., 2009a, "Apoio na Concepção de Workflow Científico Abstrato para Estudos in virtuo e in silico em Engenharia de Software". In: *VI Experimental Software Engineering Latin American Workshop*, pp. 22-31, 2009, São Carlos - SP.
- PEREIRA, W. M.; OGOSAWARA, E. ; OLIVEIRA, D. ; CHIRIGATI, F. ; CORREA, F. ; JACOB, B. ; SANTOS, I., TRAVASSOS, G. H., MATTOSO, M. L. Q., "A Conception Process for Abstract Workflows: An Example on Deep Water Oil Exploitation Domain". In: *5th IEEE International Conference on e-Science*, 2009b, Oxford, London. In Press, Poster, 2009.
- PLLANA, S., QIN, J., FAHRINGER, T., 2005, "Teuta: A tool for uml based composition of scientific grid workflows." In: *1st Austrian Grid Symposium*, p. 12, Schloss Hagenberg, Austria. Springer Verlag.
- PRESSMAN, R. S., 2010, *Software Engineering: A Practitioner's Approach*. 7 ed., New York: McGraw-Hill.

- RUS, I., HALLING, M., BIFFL, S., 2003, "Supporting Decision- Making in Software Engineering with Process Simulation and Empirical Studies". *International Journal of Software Engineering and Knowledge Engineering (IJSEKE)*, v. 13, n. 5, pp. 531-545.
- SAUER, C., JEFFERY, D.R., LAND, L., YETTON, P., 2000, "The Effectiveness of Software Development Technical Review: A Behaviorally Motivated Program of Research", *IEEE Transactions on Software Engineering*, v. 26, n.1, pp. 1-14, Janeiro.
- SHULL, F., 1998, *Developing Techniques for Using Software Documents: A Series of Empirical Studies*. Ph.D. dissertation, University of Maryland College Park, College Park, Maryland, United States.
- SI ALHIR, S., 2002, *Guide to applying the UML*. 1 ed. New York, Springer-Verlag, ISBN 0387952098.
- SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 2006, *Grandes Desafios da Computação no Brasil: 2006-2016*. Disponível em: <http://www.sbc.org.br/index.php?language=1&content=downloads&id=272>. Acesso em: 03 de Dez. de 2010.
- SOLINGEN, R.V., BERGHOUT, E., 1999. *Goal/Question/Metric Method*. 1ed., USA, McGraw-Hill Education.
- SORDE, S. W., AGGARWAL, S. K., SONG, J., KOH, M., SEE, S., 2007, "Modeling and verifying non-DAG workflows for computational grids". In: *2007 IEEE Congress on Services, SERVICES 2007*, pp. 237-243, Salt Lake City, UT, 9-13 July 2007.
- STEVENS, R. D., TIPNEY, H. J., WROE, C. J., OINN, T. M., SENGER, M., LORD, P. W., GOBLE, C. A., BRASS, A., TASSABEHJI, M. 2004, "Exploring Williams-Beuren syndrome using myGrid". *Bioinformatics*, v. 20, n. 1, Aug., pp. i303-310.
- SUN, Z., YUE, P., 2010, "The use of Web 2.0 and geoprocessing services to support geoscientific workflows". In: *18th International Conference on Geoinformatics, 2010*, pp. 1-5, 18-20 June 2010. DOI - 10.1109/GEOINFORMATICS.2010.5567702,
- TAYLOR, I. J., DEELMAN, E., GANNON, D. B., SHIELDS, M., 2007, *Workflows for e-Science: Scientific Workflows for Grids*. 1 ed., USA, Springer.
- TRAVASSOS, G. H. ; SANTOS, P. S. M. ; MIAN, P. ; DIAS NETO, A. C. ; BIOLCHINI, J., 2008, "An Environment to Support Large Scale Experimentation in Software Engineering". In: *IEEE International Conference on Engineering of Complex Computer Systems, Proceedings of ICECCS 2008, 2008*, pp. 193-202, Belfast.
- TRAVASSOS, G. H., BARROS, M. O., 2003, "Contributions of In Virtuo and In Silico Experiments for the Future of Empirical Studies in Software Engineering". In:

- Proceedings of 2nd Workshop in Workshop Series on Empirical Software Engineering the Future of Empirical Studies in Software Engineering*, Roma. Fraunhofer IRB Verlag, Roman Castles, Italy, pp. 117–130.
- TRAVASSOS, G. H., SHULL, F., CARVER, J., BASILI, V. R., 2002, *Reading Techniques for OO Design Inspections*. In: Technical Report CS-TR-4353, University of Maryland, USA. Disponível em: <<http://www.cs.umd.edu/Library/TRs/CS-TR-4353/CS-TR-4353.pdf>>.
- TROGER, A., FERNANDES, A. A. A., 2004, “A language for comprehensively supporting the in vitro experimental process in silico”. In: *Proceedings of the Fourth IEEE Symposium on Bioinformatics and Bioengineering (BIBE'04)*, pp. 47 – 56, 19 - 21 May 2004, Taichung, Taiwan, ROC.
- UKCRC, 2004, *Grand Challenges for Computer Research*. Disponível em: <http://www.ukcrc.org.uk/grand-challenge/>. Acesso em: 08 de Nov. de 2010.
- VERDI, K. K., ELLIS, H. J. C., GRYK, M. R., 2007, “Conceptual-level workflow modeling of scientific experiments using NMR as a case study”. *BMC Bioinformatics*, 2007, v. 8, n. 31, Jan.
- WOHLIN, C., HÖST, M., RUNESON, P., OHLSSON, M.C., REGNELL, B., WESSLÉN, A., 2000, *Experimentation in Software Engineering*. 1st ed. Massachusetts, Kluwer Academic Publishers.
- WORKFLOW MANAGEMENT COALITION (WFMC), 2008. *XML Process Definition Language*. Versão: 2.1. Disponível em: <<http://www.wfmc.org/xpdl.html>>. Acessado em: 08 Nov. 2010.
- WORLD WIDE WEB CONSORTIUM. 2010. *Web Services Description Language 1.1*. Disponível em: <<http://www.w3.org/TR/wsdl>>. Acessado em: 08 Nov. 2010.
- YANG, P., YANG, Z., LU, S., 2007, “Formal modeling and analysis of scientific workflows using hierarchical state machines”. In: *Third IEEE International Conference on e-Science and Grid Computing*, pp. 619—626, Bangalore, India, December 10-December 13 2007.
- YU, J., BUYYA, R., 2005, “A Taxonomy of Workflow Management Systems for Grid Computing”. *Journal of Grid Computing*, v. 3, n 3, pp. 171-200, doi: 10.1007/s10723-005-9010-8.
- YOURDON, E., 1989, *Modern Structured Analysis*. 1 ed., USA, Yourdon Press.
- ZHANG, H., KITCHENHAM, B., PFAHL, D., 2008, “Software process simulation over the past decade: trends discovery from a systematic review”. In: *ESEM '08*. ACM, NY, pp. 345-347.

ANEXO A – STRINGS DE BUSCA DA QUASI REVISÃO SISTEMÁTICA

A seguir são apresentadas as strings de busca, tanto para questão primária quanto para secundária, adaptadas para cada uma das bases de busca definidas anteriormente (Scopus, Science Direct, IEEEXplore e Compendex EI).

String para questão primária:

(scientific workflow OR scientific workflows OR scientific experiment OR scientific experiments OR science process OR science processes OR scientific process OR scientific processes OR workflow in e-science OR e-science workflow OR e-science workflows) AND ((development approach) OR (development method) OR (development technique) OR (development model) OR (development process) OR (composition approach) OR (composition method) OR (composition technique) OR (composition model) OR (composition process) OR (modeling approach) OR (modeling method) OR (modeling technique) OR (modeling model) OR (modeling process) OR (modelling approach) OR (modelling method) OR (modelling technique) OR (modelling model) OR (modelling process) OR (specification approach) OR (specification method) OR (specification technique) OR (specification model) OR (specification process) OR (design approach) OR (design method) OR (design technique) OR (design model) OR (design process) OR (definition approach) OR (definition method) OR (definition technique) OR (definition model) OR (definition process) OR (orchestration approach) OR (orchestration method) OR (orchestration technique) OR (orchestration model) OR (orchestration process) OR (conception approach) OR (conception method) OR (conception technique) OR (conception model) OR (conception process) OR (description approach) OR (description method) OR (description technique) OR (description model) OR (description process) OR (analysis approach) OR (analysis method) OR (analysis technique) OR (analysis model) OR (analysis process) OR (representation approach) OR (representation method) OR (representation technique) OR (representation model) OR (representation process))

String da questão primária para a base Scopus:

TITLE-ABS-KEY(((scientific workflow) OR (scientific workflows) OR (scientific experiment)) OR (scientific experiments)) OR (science process) OR (science processes) OR (scientific process) OR (scientific processes) OR (workflow in e-science) OR (workflows in e-science) OR (e-science workflow) OR (e-science workflows)) AND ((development W/5 approach) OR (development W/5 method) OR (development W/5 technique) OR (development W/5 model) OR (development W/5 process) OR (composition W/5 approach) OR (composition

W/5 method) OR (composition W/5 technique) OR (composition W/5 model) OR (composition W/5 process) OR (modeling W/5 approach) OR (modeling W/5 method) OR (modeling W/5 technique) OR (modeling W/5 model) OR (modeling W/5 process) OR (modelling W/5 approach) OR (modelling W/5 method) OR (modelling W/5 technique) OR (modelling W/5 model) OR (modelling W/5 process) OR (specification W/5 approach) OR (specification W/5 method) OR (specification W/5 technique) OR (specification W/5 model) OR (specification W/5 process) OR (design W/5 approach) OR (design W/5 method) OR (design W/5 technique) OR (design W/5 model) OR (design W/5 process) OR (definition W/5 approach) OR (definition W/5 method) OR (definition W/5 technique) OR (definition W/5 model) OR (definition W/5 process) OR (orchestration W/5 approach) OR (orchestration W/5 method) OR (orchestration W/5 technique) OR (orchestration W/5 model) OR (orchestration W/5 process) OR (conception W/5 approach) OR (conception W/5 method) OR (conception W/5 technique) OR (conception W/5 model) OR (conception W/5 process) OR (description W/5 approach) OR (description W/5 method) OR (description W/5 technique) OR (description W/5 model) OR (description W/5 process) OR (analysis W/5 approach) OR (analysis W/5 method) OR (analysis W/5 technique) OR (analysis W/5 model) OR (analysis W/5 process) OR (representation W/5 approach) OR (representation W/5 method) OR (representation W/5 technique) OR (representation W/5 model) OR (representation W/5 process)))

String da questão primária para a base IEEE Xplore:

((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((development <near/5> approach) OR (development <near/5> method) OR (development <near/5> technique) OR (development <near/5> model) OR (development <near/5> process) OR (composition <near/5> approach) OR (composition <near/5> method) OR (composition <near/5> technique) OR (composition <near/5> model) OR (composition <near/5> process) OR (modeling <near/5> approach) OR (modeling <near/5> method) OR (modeling <near/5> technique) OR (modeling <near/5> model) OR (modeling <near/5> process) OR (modelling <near/5> approach) OR (modelling <near/5> method) OR (modelling <near/5> technique) OR (modelling <near/5> model) OR (modelling <near/5> process) OR (specification <near/5> approach) OR (specification <near/5> method) OR (specification <near/5> technique) OR (specification <near/5> model) OR (specification <near/5> process) OR (design <near/5> approach) OR (design <near/5> method) OR (design <near/5> technique) OR (design <near/5> model) OR (design <near/5> process) OR (definition <near/5> approach) OR (definition <near/5> method) OR (definition <near/5> technique) OR (definition <near/5> model) OR (definition <near/5> process) OR (orchestration <near/5> approach) OR (orchestration <near/5> method) OR (orchestration <near/5> technique) OR (orchestration <near/5> model) OR (orchestration <near/5> process) OR (conception <near/5> approach) OR (conception <near/5> method) OR (conception <near/5> technique) OR

(conception <near/5> model) OR (conception <near/5> process) OR (description <near/5> approach) OR (description <near/5> method) OR (description <near/5> technique) OR (description <near/5> model) OR (description <near/5> process) OR (analysis <near/5> approach) OR (analysis <near/5> method) OR (analysis <near/5> technique) OR (analysis <near/5> model) OR (analysis <near/5> process) OR (representation <near/5> approach) OR (representation <near/5> method) OR (representation <near/5> technique) OR (representation <near/5> model) OR (representation <near/5> process))) <in> metadata

Devido a limitações na base de busca, é necessário dividir a string em três sub-strings de forma que a base seja capaz de interpretá-la corretamente. A String de busca original é dividida em 3 partes, descritas abaixo:

Parte 1: (((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((development <near/5> approach) OR (development <near/5> method) OR (development <near/5> technique) OR (development <near/5> model) OR (development <near/5> process) OR (composition <near/5> approach) OR (composition <near/5> method) OR (composition <near/5> technique) OR (composition <near/5> model) OR (composition <near/5> process) OR (modeling <near/5> approach) OR (modeling <near/5> method) OR (modeling <near/5> technique) OR (modeling <near/5> model) OR (modeling <near/5> process) OR (modelling <near/5> approach) OR (modelling <near/5> method) OR (modelling <near/5> technique) OR (modelling <near/5> model) OR (modelling <near/5> process) OR (specification <near/5> approach) OR (specification <near/5> method) OR (specification <near/5> technique) OR (specification <near/5> model) OR (specification <near/5> process))) <in> (metadata)

Parte 2: (((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((design <near/5> approach) OR (design <near/5> method) OR (design <near/5> technique) OR (design <near/5> model) OR (design <near/5> process) OR (definition <near/5> approach) OR (definition <near/5> method) OR (definition <near/5> technique) OR (definition <near/5> model) OR (definition <near/5> process) OR (orchestration <near/5> approach) OR (orchestration <near/5> method) OR (orchestration <near/5> technique) OR (orchestration <near/5> model) OR (orchestration <near/5> process) OR (conception <near/5> approach) OR (conception <near/5> method) OR (conception <near/5> technique) OR (conception <near/5> model) OR (conception <near/5> process))) <in> (metadata)

Parte 3: (((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((description <near/5> approach)

OR (description <near/5> method) OR (description <near/5> technique) OR (description <near/5> model) OR (description <near/5> process) OR (analysis <near/5> approach) OR (analysis <near/5> method) OR (analysis <near/5> technique) OR (analysis <near/5> model) OR (analysis <near/5> process) OR (representation <near/5> approach) OR (representation <near/5> method) OR (representation <near/5> technique) OR (representation <near/5> model) OR (representation <near/5> process))) <in> (metadata)

Obs.: Na base IEEE Xplore não foi possível efetuar filtro na string de busca para procurar somente no *abstract*, *title* e *keywords*. Desta maneira foi executada uma busca mais ampla e foi incorporado o ônus da separação manual dos trabalhos retornados.

String da questão primária para a base Compendex EI:

((({scientific workflow}) OR ({scientific workflows}) OR ({scientific experiment}) OR ({scientific experiments}) OR ({science process}) OR ({science processes}) OR ({scientific process}) OR ({scientific processes}) OR ({workflow in e-science}) OR ({workflows in e-science}) OR ({e-science workflow}) OR ({e-science workflows})) AND ((development ONEAR/5 approach) OR (development ONEAR/5 method) OR (development ONEAR/5 technique) OR (development ONEAR/5 model) OR (development ONEAR/5 process) OR (composition ONEAR/5 approach) OR (composition ONEAR/5 method) OR (composition ONEAR/5 technique) OR (composition ONEAR/5 model) OR (composition ONEAR/5 process) OR (modeling ONEAR/5 approach) OR (modeling ONEAR/5 method) OR (modeling ONEAR/5 technique) OR (modeling ONEAR/5 model) OR (modeling ONEAR/5 process) OR (modelling ONEAR/5 approach) OR (modelling ONEAR/5 method) OR (modelling ONEAR/5 technique) OR (modelling ONEAR/5 model) OR (modelling ONEAR/5 process) OR (specification ONEAR/5 approach) OR (specification ONEAR/5 method) OR (specification ONEAR/5 technique) OR (specification ONEAR/5 model) OR (specification ONEAR/5 process) OR (design ONEAR/5 approach) OR (design ONEAR/5 method) OR (design ONEAR/5 technique) OR (design ONEAR/5 model) OR (design ONEAR/5 process) OR (definition ONEAR/5 approach) OR (definition ONEAR/5 method) OR (definition ONEAR/5 technique) OR (definition ONEAR/5 model) OR (definition ONEAR/5 process) OR (orchestration ONEAR/5 approach) OR (orchestration ONEAR/5 method) OR (orchestration ONEAR/5 technique) OR (orchestration ONEAR/5 model) OR (orchestration ONEAR/5 process) OR (conception ONEAR/5 approach) OR (conception ONEAR/5 method) OR (conception ONEAR/5 technique) OR (conception ONEAR/5 model) OR (conception ONEAR/5 process) OR (description ONEAR/5 approach) OR (description ONEAR/5 method) OR (description ONEAR/5 technique) OR (description ONEAR/5 model) OR (description ONEAR/5 process) OR (analysis ONEAR/5 approach) OR (analysis ONEAR/5 method) OR (analysis ONEAR/5 technique) OR (analysis ONEAR/5 model) OR (analysis ONEAR/5 process) OR (representation ONEAR/5 approach) OR (representation ONEAR/5 method) OR (representation ONEAR/5 technique) OR (representation ONEAR/5 model) OR (representation ONEAR/5 process))) wn KY

Devido a limitações na base de busca, é necessário dividir a string em três sub-strings de forma que a base seja capaz de interpretá-la corretamente. A String de busca original é dividida em 2 partes, descritas abaixo:

Parte 1: ((({scientific workflow}) OR ({scientific workflows}) OR ({scientific experiment}) OR ({scientific experiments}) OR ({science process}) OR ({science processes}) OR ({scientific process}) OR ({scientific processes}) OR ({workflow in e-science}) OR ({workflows in e-science}) OR ({e-science workflow}) OR ({e-science workflows})) AND ((development ONEAR/5 approach) OR (development ONEAR/5 method) OR (development ONEAR/5 technique) OR (development ONEAR/5 model) OR (development ONEAR/5 process) OR (composition ONEAR/5 approach) OR (composition ONEAR/5 method) OR (composition ONEAR/5 technique) OR (composition ONEAR/5 model) OR (composition ONEAR/5 process) OR (modeling ONEAR/5 approach) OR (modeling ONEAR/5 method) OR (modeling ONEAR/5 technique) OR (modeling ONEAR/5 model) OR (modeling ONEAR/5 process) OR (modelling ONEAR/5 approach) OR (modelling ONEAR/5 method) OR (modelling ONEAR/5 technique) OR (modelling ONEAR/5 model) OR (modelling ONEAR/5 process) OR (specification ONEAR/5 approach) OR (specification ONEAR/5 method) OR (specification ONEAR/5 technique) OR (specification ONEAR/5 model) OR (specification ONEAR/5 process) OR (design ONEAR/5 approach) OR (design ONEAR/5 method) OR (design ONEAR/5 technique) OR (design ONEAR/5 model) OR (design ONEAR/5 process) OR (definition ONEAR/5 approach) OR (definition ONEAR/5 method) OR (definition ONEAR/5 technique) OR (definition ONEAR/5 model) OR (definition ONEAR/5 process) OR (orchestration ONEAR/5 approach) OR (orchestration ONEAR/5 method) OR (orchestration ONEAR/5 technique) OR (orchestration ONEAR/5 model) OR (orchestration ONEAR/5 process))) wn KY

Parte 2: ((({scientific workflow}) OR ({scientific workflows}) OR ({scientific experiment}) OR ({scientific experiments}) OR ({science process}) OR ({science processes}) OR ({scientific process}) OR ({scientific processes}) OR ({workflow in e-science}) OR ({workflows in e-science}) OR ({e-science workflow}) OR ({e-science workflows})) AND ((conception ONEAR/5 approach) OR (conception ONEAR/5 method) OR (conception ONEAR/5 technique) OR (conception ONEAR/5 model) OR (conception ONEAR/5 process) OR (description ONEAR/5 approach) OR (description ONEAR/5 method) OR (description ONEAR/5 technique) OR (description ONEAR/5 model) OR (description ONEAR/5 process) OR (analysis ONEAR/5 approach) OR (analysis ONEAR/5 method) OR (analysis ONEAR/5 technique) OR (analysis ONEAR/5 model) OR (analysis ONEAR/5 process) OR (representation ONEAR/5 approach) OR (representation ONEAR/5 method) OR (representation ONEAR/5 technique) OR (representation ONEAR/5 model) OR (representation ONEAR/5 process))) wn KY

String da questão primária para a base Science Direct:

tak((({scientific workflow}) OR ({scientific workflows}) OR ({scientific experiment}) OR ({scientific experiments}) OR ({science process}) OR ({science processes}) OR ({scientific process}) OR ({scientific processes}) OR ({workflow in e-science}) OR ({workflows in e-science}))

OR ({e-science workflow}) OR ({e-science workflows})) AND ((development W/5 approach) OR (development W/5 method) OR (development W/5 technique) OR (development W/5 model) OR (development W/5 process) OR (composition W/5 approach) OR (composition W/5 method) OR (composition W/5 technique) OR (composition W/5 model) OR (composition W/5 process) OR (modeling W/5 approach) OR (modeling W/5 method) OR (modeling W/5 technique) OR (modeling W/5 model) OR (modeling W/5 process) OR (modelling W/5 approach) OR (modelling W/5 method) OR (modelling W/5 technique) OR (modelling W/5 model) OR (modelling W/5 process) OR (specification W/5 approach) OR (specification W/5 method) OR (specification W/5 technique) OR (specification W/5 model) OR (specification W/5 process) OR (design W/5 approach) OR (design W/5 method) OR (design W/5 technique) OR (design W/5 model) OR (design W/5 process) OR (definition W/5 approach) OR (definition W/5 method) OR (definition W/5 technique) OR (definition W/5 model) OR (definition W/5 process) OR (orchestration W/5 approach) OR (orchestration W/5 method) OR (orchestration W/5 technique) OR (orchestration W/5 model) OR (orchestration W/5 process) OR (conception W/5 approach) OR (conception W/5 method) OR (conception W/5 technique) OR (conception W/5 model) OR (conception W/5 process) OR (description W/5 approach) OR (description W/5 method) OR (description W/5 technique) OR (description W/5 model) OR (description W/5 process) OR (analysis W/5 approach) OR (analysis W/5 method) OR (analysis W/5 technique) OR (analysis W/5 model) OR (analysis W/5 process) OR (representation W/5 approach) OR (representation W/5 method) OR (representation W/5 technique) OR (representation W/5 model) OR (representation W/5 process)))

String para questão secundária:

((scientific workflow OR scientific workflows OR scientific experiment OR scientific experiments OR science process OR science processes OR scientific process OR scientific processes OR workflow in e-science OR e-science workflow OR e-science workflows) **AND** ((development approach) OR (development method) OR (development technique) OR (development model) OR (development process) OR (composition approach) OR (composition method) OR (composition technique) OR (composition model) OR (composition process) OR (modeling approach) OR (modeling method) OR (modeling technique) OR (modeling model) OR (modeling process) OR (modelling approach) OR (modelling method) OR (modelling technique) OR (modelling model) OR (modelling process) OR (specification approach) OR (specification method) OR (specification technique) OR (specification model) OR (specification process) OR (design approach) OR (design method) OR (design technique) OR (design model) OR (design process) OR (definition approach) OR (definition method) OR (definition technique) OR (definition model) OR (definition process) OR (orchestration approach) OR (orchestration method) OR (orchestration technique) OR (orchestration model) OR (orchestration process) OR (conception approach) OR (conception method) OR (conception technique) OR (conception model) OR (conception process) OR (description approach) OR (description method) OR (description technique) OR (description model) OR (description process) OR (analysis

approach) OR (analysis method) OR (analysis technique) OR (analysis model) OR (analysis process) OR (representation approach) OR (representation method) OR (representation technique) OR (representation model) OR (representation process))) AND (inspection OR review OR verification OR validation OR reading OR revise)

String da questão secundária para a base Scopus:

TITLE-ABS-KEY((((scientific workflow) OR (scientific workflows) OR (scientific experiment) OR (scientific experiments) OR (science process) OR (science processes) OR (scientific process) OR (scientific processes) OR (workflow in e-science) OR (workflows in e-science) OR (e-science workflow) OR (e-science workflows))) AND ((development W/5 approach) OR (development W/5 method) OR (development W/5 technique) OR (development W/5 model) OR (development W/5 process) OR (composition W/5 approach) OR (composition W/5 method) OR (composition W/5 technique) OR (composition W/5 model) OR (composition W/5 process) OR (modeling W/5 approach) OR (modeling W/5 method) OR (modeling W/5 technique) OR (modeling W/5 model) OR (modeling W/5 process) OR (modelling W/5 approach) OR (modelling W/5 method) OR (modelling W/5 technique) OR (modelling W/5 model) OR (modelling W/5 process) OR (specification W/5 approach) OR (specification W/5 method) OR (specification W/5 technique) OR (specification W/5 model) OR (specification W/5 process) OR (design W/5 approach) OR (design W/5 method) OR (design W/5 technique) OR (design W/5 model) OR (design W/5 process) OR (definition W/5 approach) OR (definition W/5 method) OR (definition W/5 technique) OR (definition W/5 model) OR (definition W/5 process) OR (orchestration W/5 approach) OR (orchestration W/5 method) OR (orchestration W/5 technique) OR (orchestration W/5 model) OR (orchestration W/5 process) OR (conception W/5 approach) OR (conception W/5 method) OR (conception W/5 technique) OR (conception W/5 model) OR (conception W/5 process) OR (description W/5 approach) OR (description W/5 method) OR (description W/5 technique) OR (description W/5 model) OR (description W/5 process) OR (analysis W/5 approach) OR (analysis W/5 method) OR (analysis W/5 technique) OR (analysis W/5 model) OR (analysis W/5 process) OR (representation W/5 approach) OR (representation W/5 method) OR (representation W/5 technique) OR (representation W/5 model) OR (representation W/5 process))) AND (inspection OR review OR verification OR validation OR reading OR revise))

String da questão secundária para a base IEEE Xplore:

((('scientific workflow' OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((development <near/5> approach) OR (development <near/5> method) OR (development <near/5> technique) OR (development <near/5> model) OR (development <near/5> process) OR (composition <near/5> approach) OR (composition <near/5> method) OR (composition <near/5> technique) OR (composition

<near/5> model) OR (composition <near/5> process) OR (modeling <near/5> approach) OR (modeling <near/5> method) OR (modeling <near/5> technique) OR (modeling <near/5> model) OR (modeling <near/5> process) OR (modelling <near/5> approach) OR (modelling <near/5> method) OR (modelling <near/5> technique) OR (modelling <near/5> model) OR (modelling <near/5> process) OR (specification <near/5> approach) OR (specification <near/5> method) OR (specification <near/5> technique) OR (specification <near/5> model) OR (specification <near/5> process) OR (design <near/5> approach) OR (design <near/5> method) OR (design <near/5> technique) OR (design <near/5> model) OR (design <near/5> process) OR (definition <near/5> approach) OR (definition <near/5> method) OR (definition <near/5> technique) OR (definition <near/5> model) OR (definition <near/5> process) OR (orchestration <near/5> approach) OR (orchestration <near/5> method) OR (orchestration <near/5> technique) OR (orchestration <near/5> model) OR (orchestration <near/5> process) OR (conception <near/5> approach) OR (conception <near/5> method) OR (conception <near/5> technique) OR (conception <near/5> model) OR (conception <near/5> process) OR (description <near/5> approach) OR (description <near/5> method) OR (description <near/5> technique) OR (description <near/5> model) OR (description <near/5> process) OR (analysis <near/5> approach) OR (analysis <near/5> method) OR (analysis <near/5> technique) OR (analysis <near/5> model) OR (analysis <near/5> process) OR (representation <near/5> approach) OR (representation <near/5> method) OR (representation <near/5> technique) OR (representation <near/5> model) OR (representation <near/5> process))) AND (inspection OR review OR verification OR validation OR reading OR revise)) <in> (metadata)

Devido a limitações na base de busca, é necessário dividir a string em três sub-strings de forma que a base seja capaz de interpretá-la corretamente. A String de busca original é dividida em 3 partes, descritas abaixo:

Parte 1: (((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((development <near/5> approach) OR (development <near/5> method) OR (development <near/5> technique) OR (development <near/5> model) OR (development <near/5> process) OR (composition <near/5> approach) OR (composition <near/5> method) OR (composition <near/5> technique) OR (composition <near/5> model) OR (composition <near/5> process) OR (modeling <near/5> approach) OR (modeling <near/5> method) OR (modeling <near/5> technique) OR (modeling <near/5> model) OR (modeling <near/5> process) OR (modelling <near/5> approach) OR (modelling <near/5> method) OR (modelling <near/5> technique) OR (modelling <near/5> model) OR (modelling <near/5> process) OR (specification <near/5> approach) OR (specification <near/5> method) OR (specification <near/5> technique) OR (specification <near/5> model) OR (specification <near/5> process))) AND (inspection OR review OR verification OR validation OR reading OR revise)) <in> (metadata)

Parte 2: (((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((design <near/5> approach) OR (design <near/5> method) OR (design <near/5> technique) OR (design <near/5> model) OR (design <near/5> process) OR (definition <near/5> approach) OR (definition <near/5> method) OR (definition <near/5> technique) OR (definition <near/5> model) OR (definition <near/5> process) OR (orchestration <near/5> approach) OR (orchestration <near/5> method) OR (orchestration <near/5> technique) OR (orchestration <near/5> model) OR (orchestration <near/5> process) OR (conception <near/5> approach) OR (conception <near/5> method) OR (conception <near/5> technique) OR (conception <near/5> model) OR (conception <near/5> process))) AND (inspection OR review OR verification OR validation OR reading OR revise)) <in> (metadata)

Parte3: (((('scientific workflow') OR ('scientific workflows') OR ('scientific experiment') OR ('scientific experiments') OR ('science process') OR ('science processes') OR ('scientific process') OR ('scientific processes') OR ('workflow in e-science') OR ('workflows in e-science') OR ('e-science workflow') OR ('e-science workflows')) AND ((description <near/5> approach) OR (description <near/5> method) OR (description <near/5> technique) OR (description <near/5> model) OR (description <near/5> process) OR (analysis <near/5> approach) OR (analysis <near/5> method) OR (analysis <near/5> technique) OR (analysis <near/5> model) OR (analysis <near/5> process) OR (representation <near/5> approach) OR (representation <near/5> method) OR (representation <near/5> technique) OR (representation <near/5> model) OR (representation <near/5> process))) AND (inspection OR review OR verification OR validation OR reading OR revise)) <in> (metadata)

Obs.: Na base IEEE Xplore não foi possível efetuar filtro na string de busca para procurar somente no *abstract*, *title* e *keywords*. Desta maneira foi executada uma busca mais ampla em todos os campos, se incorporado assim o ônus da separação manual dos trabalhos retornados.

String da questão secundária para a base Compendex EI:

(((((scientific workflow}) OR ({scientific workflows}) OR ({scientific experiment}) OR ({scientific experiments}) OR ({science process}) OR ({science processes}) OR ({scientific process}) OR ({scientific processes}) OR ({workflow in e-science}) OR ({workflows in e-science}) OR ({e-science workflow}) OR ({e-science workflows})) AND ((development ONEAR/5 approach) OR (development ONEAR/5 method) OR (development ONEAR/5 technique) OR (development ONEAR/5 model) OR (development ONEAR/5 process) OR (composition ONEAR/5 approach) OR (composition ONEAR/5 method) OR (composition ONEAR/5 technique) OR (composition ONEAR/5 model) OR (composition ONEAR/5 process) OR (modeling ONEAR/5 approach) OR (modeling ONEAR/5 method) OR (modeling ONEAR/5 technique) OR (modeling ONEAR/5 model) OR (modeling ONEAR/5 process) OR (modelling

ONEAR/5 approach) OR (modelling ONEAR/5 method) OR (modelling ONEAR/5 technique) OR (modelling ONEAR/5 model) OR (modelling ONEAR/5 process) OR (specification ONEAR/5 approach) OR (specification ONEAR/5 method) OR (specification ONEAR/5 technique) OR (specification ONEAR/5 model) OR (specification ONEAR/5 process) OR (design ONEAR/5 approach) OR (design ONEAR/5 method) OR (design ONEAR/5 technique) OR (design ONEAR/5 model) OR (design ONEAR/5 process) OR (definition ONEAR/5 approach) OR (definition ONEAR/5 method) OR (definition ONEAR/5 technique) OR (definition ONEAR/5 model) OR (definition ONEAR/5 process) OR (orchestration ONEAR/5 approach) OR (orchestration ONEAR/5 method) OR (orchestration ONEAR/5 technique) OR (orchestration ONEAR/5 model) OR (orchestration ONEAR/5 process) OR (conception ONEAR/5 approach) OR (conception ONEAR/5 method) OR (conception ONEAR/5 technique) OR (conception ONEAR/5 model) OR (conception ONEAR/5 process) OR (description ONEAR/5 approach) OR (description ONEAR/5 method) OR (description ONEAR/5 technique) OR (description ONEAR/5 model) OR (description ONEAR/5 process) OR (analysis ONEAR/5 approach) OR (analysis ONEAR/5 method) OR (analysis ONEAR/5 technique) OR (analysis ONEAR/5 model) OR (analysis ONEAR/5 process) OR (representation ONEAR/5 approach) OR (representation ONEAR/5 method) OR (representation ONEAR/5 technique) OR (representation ONEAR/5 model) OR (representation ONEAR/5 process))) AND (inspection OR review OR verification OR validation OR reading OR revise)) wn KY

String da questão secundária para a base Science Direct:

tak((((scientific workflow)) OR (scientific workflows)) OR (scientific experiment)) OR (scientific experiments)) OR (science process)) OR (science processes)) OR (scientific process)) OR (scientific processes)) OR (workflow in e-science)) OR (workflows in e-science)) OR (e-science workflow)) OR (e-science workflows))) AND ((development W/5 approach) OR (development W/5 method) OR (development W/5 technique) OR (development W/5 model) OR (development W/5 process) OR (composition W/5 approach) OR (composition W/5 method) OR (composition W/5 technique) OR (composition W/5 model) OR (composition W/5 process) OR (modeling W/5 approach) OR (modeling W/5 method) OR (modeling W/5 technique) OR (modeling W/5 model) OR (modeling W/5 process) OR (modelling W/5 approach) OR (modelling W/5 method) OR (modelling W/5 technique) OR (modelling W/5 model) OR (modelling W/5 process) OR (specification W/5 approach) OR (specification W/5 method) OR (specification W/5 technique) OR (specification W/5 model) OR (specification W/5 process) OR (design W/5 approach) OR (design W/5 method) OR (design W/5 technique) OR (design W/5 model) OR (design W/5 process) OR (definition W/5 approach) OR (definition W/5 method) OR (definition W/5 technique) OR (definition W/5 model) OR (definition W/5 process) OR (orchestration W/5 approach) OR (orchestration W/5 method) OR (orchestration W/5 technique) OR (orchestration W/5 model) OR (orchestration W/5 process) OR (conception W/5 approach) OR (conception W/5 method) OR (conception W/5 technique) OR (conception W/5 model) OR (conception W/5 process) OR (description W/5 approach) OR (description W/5 method) OR

(description W/5 technique) OR (description W/5 model) OR (description W/5 process) OR (analysis W/5 approach) OR (analysis W/5 method) OR (analysis W/5 technique) OR (analysis W/5 model) OR (analysis W/5 process) OR (representation W/5 approach) OR (representation W/5 method) OR (representation W/5 technique) OR (representation W/5 model) OR (representation W/5 process))) AND (inspection OR review OR verification OR validation OR reading OR revise))

ANEXO B – LISTA DE TODAS AS REFERÊNCIAS

Legenda: “E” – Excluído; “I” – Incluído; “-” – Não recuperado.

Referência	Questão Pri.	Questão Sec.
Arenas, A., Aziz, B., Bicarregui, J. & Matthews, B., 2008. Managing Conflicts of Interest in Virtual Organisations. Electron. Notes Theor. Comput. Sci., Vol. 197(2), pp. 45-56 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Bosin, A., Dessì, N., Fugini, M. & Pes, B., 2008. Cooperative e-organizations for distributed bioinformatics experiments. Lect. Notes Comput. Sci., Vol. 5326 LNCS9th International Conference on Intelligent Data Engineering and Automated Learning, IDEAL 2008, pp. 306-313 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Bowers, S. & Ludäscher, B., 2005. Actor-oriented design of scientific workflows. Lect. Notes Comput. Sci., Vol. 3716 LNCS24th International Conference on Conceptual Modeling - ER 2005, pp. 369-384 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Curcin, V., Ghanem, M.M. & Guo, Y., 2009. Analysing scientific workflows with computational tree logic. Cluster Computing, Vol. 12(4), pp. 399-419 . Base: CompendexEI; Recuperado: 2010.09.29 12:50:14.	E	E
Fei, X., Lu, S. & Lin, C., 2009. A mapreduce-enabled scientific workflow composition framework. 2009 IEEE International Conference on Web Services, ICWS 2009, pp. 663-670 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Jamil, H. & Islam, A., 2009. The power of declarative languages: A comparative exposition of scientific workflow design using BioFlow and Taverna. (PART 1)SERVICES 2009 - 5th 2009 World Congress on Services, pp. 322-329 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Leng, Y., El-Gayyar, M. & Cremers, A.B., 2010. Semantics Enhanced Composition Planner for Distributed Resources. Distributed Computing and Applications to Business Engineering and Science (DCABES), 2010 Ninth International Symposium on DOI - 10.1109/DCABES.2010.19, pp. 61-65 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Lin, Y., Libourel, T. & Mougnot, I., 2009. A workflow language for the experimental sciences. , Vol. ISASICEIS 2009 - 11th International Conference on Enterprise Information Systems, pp. 372-375 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ngu, A., Bowers, S., Haasch, N., McPhillips, T. & Critchlow, T., 2008. Flexible scientific workflow modeling using frames, templates, and dynamic embedding. Lect. Notes Comput. Sci., Vol. 5069 LNCS20th International Conference on Scientific and Statistical Database Management, SSDBM 2008, pp. 566-572 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ogasawara, E., Paulino, C., Murta, L., Werner, C. & Mattoso, M., 2009. Experiment line: Software reuse in scientific workflows. Lect. Notes Comput. Sci., Vol. 5566 LNCS21st International Conference on Scientific and Statistical Database Management, SSDBM 2009, pp. 264-272 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ogasawara, E., Rangel, P., Murta, L., Werner, C. & Mattoso, M., 2009. Comparison and versioning of scientific workflows. Comparison and Versioning of Software Models, 2009. CVSM '09. ICSE Workshop on DOI - 10.1109/CVSM.2009.5071718, pp. 25-30 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Pignotti, E., Edwards, P., Polhill, G., Gotts, N. & Preece, A., 2008. A semantic workflow mechanism to realise experimental goals and constraints. 2008 3rd Workshop on Workflows in Support of Large-Scale Science, WORKS 2008, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Slominski, A., 2010. Flexible Creation and Adaptive Execution of Scientific Workflows in Cloud and Grid Environments by Using Web 2.0-Based Electronic Lab Notebook Metaphor. Services (SERVICES-1), 2010 6th World Congress on DOI - 10.1109/SERVICES.2010.28, pp. 326-327 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-

Sroka, J., WÅ,odarczyk, P., Krupa, L. & Hidders, J., 2010. DFL designer - Collection-oriented scientific workflows with Petri nets and nested relational calculus. 1st International Workshop on Workflow Approaches to New Data-centric Science, Wands '10, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Sdor, B., Baksa, A., BaiÅjzs, A., Szalai, S. & TrÅznai, G., 2008. New approach to modelling spacecraft modules. , Vol. 1159th International Astronautical Congress 2008, IAC 2008, pp. 7135-7145 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Zhang, J., Kuc, D. & Lu, S., 2010. Confucius: A Scientific Collaboration System Using Collaborative Scientific Workflows. Web Services (ICWS), 2010 IEEE International Conference on DOI - 10.1109/ICWS.2010.103, pp. 575-583 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Abramson, D., Bethwaite, B., Enticott, C., Garic, S. & Peachey, T., 2009. Parameter space exploration using scientific workflows. Lect. Notes Comput. Sci., Vol. 5544 LNCS(PART 1)9th International Conference on Computational Science, ICCS 2009, pp. 104-113 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Abramson, D., Bethwaite, B., Enticott, C., Garic, S., Peachey, T., Michailova, A. & Amirriazi, S., 2010. Automatic design optimization using parallel workflows. Procedia Computer Science, Vol. 1(1)ICCS 2010, pp. 2159-2168 . Base: Science Direct; Recuperado: 2010.09.29 01:24:08.	E	-
Altintas, I., Berkley, C., Jaeger, E., Jones, M., Ludascher, B. & Mock, S., 2004. Kepler: an extensible system for design and execution of scientific workflows. Scientific and Statistical Database Management, 2004. Proceedings. 16th International Conference on, pp. 423-424 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Altintas, I., Bhagwanani, S., Buttler, D., Chandra, S., Cheng, Z., Coleman, M., Critchlow, T., Gupta, A., Han, W., Liu, L., Ludascher, B., Pu, C., Moore, R., Shoshani, A. & Vouk, M., 2003. A modeling and execution environment for distributed scientific workflows. Conference on Scientific and Statistical Database Management, 2003. 15th International, pp. 247-250 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Altintas, I., Lin, A.W., Chen, J., Churas, C., Gujral, M., Sun, S., Li, W., Manansala, R., Sedova, M., Grethe, J.S. & Ellisman, M., 2010. CAMERA 2.0: A Data-centric Metagenomics Community Infrastructure Driven by Scientific Workflows. Services (SERVICES-1), 2010 6th World Congress on DOI - 10.1109/SERVICES.2010.89, pp. 352-359 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Ayadi, N., Lacroix, Z., Vidal, M.-E. & Ruckhaus, E., 2007. Deductive web services: An ontology-driven approach for service interoperability in life science. Lect. Notes Comput. Sci., Vol. 4806 LNCS(PART 2)OTM Confederated International Workshops and Posters AWeSOMe, CAMS, OTM Academy Doctoral Consortium, MONET, OnToContent, ORM, PerSys, PPN, RDDS, SWS, and SWWS 2007, pp. 1338-1347 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Belhajjame, K., Wolstencroft, K., Corcho, O., Oinn, T., Tanoh, F., William, A. & Goble, C., 2008. Metadata Management in the Taverna Workflow System. Cluster Computing and the Grid, 2008. CCGRID '08. 8th IEEE International Symposium on, pp. 651-656 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Bharathi, S., Chervenak, A., Deelman, E., Mehta, G., Su, M.-H. & Vahi, K., 2008. Characterization of scientific workflows. Workflows in Support of Large-Scale Science, 2008. WORKS 2008. Third Workshop on DOI - 10.1109/WORKS.2008.4723958, pp. 1-10 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Bourgne, G. & Corruble, V., 2008. A framework for knowledge discovery in a society of agents. Lect. Notes Comput. Sci., Vol. 5255 LNAI11th International Conference on Discovery Science, DS 2008, pp. 172-184 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Brazier, P., Chebotko, A., Gates, A. & Salayandia, L., 2009. GEO-SEED: A Metadata Repository for Geosciences Web Service Discovery. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.43, pp. 356-359 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Buckingham, L., Hogan, J., Roe, P., Sumitomo, J. & Towsey, M., 2008. Comparative Studies Made Simple in GPFlow. Cluster Computing and the Grid, 2008. CCGRID '08. 8th IEEE International Symposium on DOI - 10.1109/CCGRID.2008.91, pp. 699-699 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Burns, G. & Russ, T., 2009. Biomedical knowledge engineering tools based on experimental design: A case study based on neuroanatomical tract-tracing experiments. 5th International Conference on Knowledge Capture, K-CAP'09, pp. 173-174 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Cardoso, L., de Souza, J. & Marques, C., 2002. A collaborative approach to the reuse of scientific experiments in the Bill of Experiments tool. Computer Supported Cooperative Work in Design, 2002. The 7th International Conference on DOI - 10.1109/CSCWD.2002.1047704, pp. 296-301 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Chase, J., Schuchardt, K., Chin Jr., G., Daily, J. & Scheibe, T., 2008. Iterative workflows for numerical simulations in subsurface sciences. , Vol. PART 12008 IEEE Congress on Services, SERVICES 2008, pp. 461-464 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Chebotko, A., Lu, S., Chang, S., Fotouhi, F. & Yang, P., 2010. Secure Abstraction Views for Scientific Workflow Provenance Querying. Services Computing, IEEE Transactions on DOI - 10.1109/TSC.2010.38, Vol. PP(99)Services Computing, IEEE Transactions on, pp. 1-1 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Chen, I.-M. & Markowitz, V., 1995. Modeling scientific experiments with an object data model. Data Engineering, 1995. Proceedings of the Eleventh International Conference on, pp. 391-400 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Chen, J. & Yang, Y., 2010. Localising temporal constraints in scientific workflows. J. Comput. Syst. Sci., Vol. 76(6), pp. 464-474 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Chen, Y., Zhang, R., Fan, X., Zhang, C. & Li, L., 2008. A visual scientific workflow designer for chemists in grid environment. 3rd ChinaGrid Annual Conference, ChinaGrid 2008, pp. 211-217 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chiu, D., Deshpande, S., Agrawal, G. & Li, R., 2009. A Dynamic Approach toward QoS-Aware Service Workflow Composition. Web Services, 2009. ICWS 2009. IEEE International Conference on DOI - 10.1109/ICWS.2009.36, pp. 655-662 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Curcin, V., Ghanem, M. & Guo, Y., 2010. Polymorphic type framework for scientific workflows with relational data model. Int. J. Bus. Process Integr. Manage., Vol. 5(1), pp. 45-62 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Curcin, V., Ghanem, M., Molokhia, M., Guo, Y. & Darlington, J., 2008. Mining Adverse Drug Reactions with E-Science Workflows. Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International DOI - 10.1109/CIBEC.2008.4786100, pp. 1-5 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Curran, O., Downes, P., Cunniffe, J. & Shearer, A., 2008. Fine-Grained Workflow in Heterogeneous Environments. Parallel, Distributed and Network-Based Processing, 2008. PDP 2008. 16th Euromicro Conference on, pp. 115-119 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Dadi, U. & Di, L., 2009. Creating web service interfaces and scientific workflows using command line tools: A GRASS example. Geoinformatics, 2009 17th International Conference on DOI - 10.1109/GEOINFORMATICS.2009.5293464, pp. 1-6 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Dou, W., Zhao, J. & Fan, S., 2010. A collaborative scheduling approach for service-driven scientific workflow execution. J. Comput. Syst. Sci., Vol. 76(6), pp. 416-427 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Du, N., Li, Q., Liang, Y. & Du, L., 2008. Embracing Scientific Workflow Knowledge into IDE. Semantics, Knowledge and Grid, 2008. SKG '08. Fourth International Conference on DOI - 10.1109/SKG.2008.79, pp. 384-387 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Du, N., Liang, Y., Li, Q. & Du, L., 2008. Transition System Semantics of DFL as a Scientific Workflow Language. Semantics, Knowledge and Grid, 2008. SKG '08. Fourth International Conference on DOI - 10.1109/SKG.2008.78, pp. 241-246 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Ezenwoye, O., Busi, S. & Sadjadi, S., 2010. Dynamically reconfigurable data-intensive service composition. , Vol. 16th International Conference on Web Information Systems and Technologies, WEBIST 2010, pp. 125-130 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Fei, X. & Lu, S., 2010. A Collectional Data Model for Scientific Workflow Composition. Web Services (ICWS), 2010 IEEE International Conference on DOI - 10.1109/ICWS.2010.93, pp. 567-574 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Gooding, D. & Addis, T., 2008. Modelling experiments as mediating models. Found. Sci., Vol. 13(1), pp. 17-35 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Jablonski, S., Cure, O., Rehman, M. & Volz, B., 2008. Architecture of the DaltOn Data Integration System for Scientific Applications. Cluster Computing and the Grid, 2008. CCGRID '08. 8th IEEE International Symposium on, pp. 701-701 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Jamil, H. & El-Hajj-Diab, B., 2008. BioFlow: A Web-Based Declarative Workflow Language for Life Sciences. Services - Part I, 2008. IEEE Congress on DOI - 10.1109/SERVICES-1.2008.73, pp. 453-460 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Jamil, H., Islam, A. & Hossain, S., 2010. A declarative language and toolkit for scientific workflow implementation and execution. Int. J. Bus. Process Integr. Manage., Vol. 5(1), pp. 3-17 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-

Jung, J.-Y., Lee, W. & Kang, S.-H., 2006. Process decomposition and choreography for distributed scientific workflow enactment. Lect. Notes Comput. Sci., Vol. 3984 LNCSICCSA 2006: International Conference on Computational Science and Its Applications, pp. 942-951 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kandaswamy, G., Fang, L., Huang, Y., Shirasuna, S., Marru, S. & Gannon, D., 2006. Building web services for scientific grid applications. IBM Journal of Research and Development DOI - 10.1147/rd.502.0249, Vol. 50(2.3)IBM Journal of Research and Development, pp. 249-260 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Kang, Z., Li, Y., De Garis, H. & Kang, L.-S., 2002. A multi-level and multi-scale evolutionary modeling system for scientific data. , Vol. 12002 International Joint Conference on Neural Networks (IJCNN '02), pp. 737-742 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kapetanios, E., 1996. Extracting and providing knowledge within an object-centered scientific information system for atmospheric research. Scientific and Statistical Database Systems, 1996. Proceedings., Eighth International Conference on, pp. 198-207 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
King, R. & Soldatova, L., 2008. Formalising phylogenetic experiments: Ontologies and logical inference. , Vol. SS-08-072008 AAAI Spring Symposium, pp. 59-62 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kinsy, M. & Lacroix, Z., 2007. Storing efficiently bioinformatics workflows. 7th IEEE International Conference on Bioinformatics and Bioengineering, BIBE, pp. 1328-1332 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lacroix, Z., Vidal, M.-E. & Legendre, C., 2009. Customized and optimized service selection with protocolDB. Lect. Notes Comput. Sci., Vol. 5697 LNCS2nd International Conference on Data Management in Grid and Peer-to-Peer Systems, Globe 2009, pp. 112-123 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Lee, K., Sakellariou, R., Paton, N. & Fernandes, A., 2007. Workflow adaptation as an autonomic computing problem. Proc. Workshop Workflows Support Large-scale Sci.16th International Symposium on High Performance Distributed Computing 2007, HPDC'07 and Co-Located Workshops, pp. 29-34 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lin, C., Lu, S., Fei, X., Pai, D. & Hua, J., 2009. A task abstraction and mapping approach to the shimming problem in scientific workflows. SCC 2009 - 2009 IEEE International Conference on Services Computing, pp. 284-291 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Lin, C., Lu, S., Lai, Z., Chebotko, A., Fei, X., Hua, J. & Fotouhi, F., 2008. Service-Oriented Architecture for VIEW: A Visual Scientific Workflow Management System. , Vol. 1Services Computing, 2008. SCC '08. IEEE International Conference on DOI - 10.1109/SCC.2008.118, pp. 335-342 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Lin, J., Zhao, Z., Li, S., Zhang, H., Li, L., Li, C. & Li, L., 2009. Application-Objected Workflow Management System Based on Abstract Service. Grid and Cooperative Computing, 2009. GCC '09. Eighth International Conference on DOI - 10.1109/GCC.2009.32, pp. 144-152 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Liu, C., Zhang, W., Luo, Z., Cao, X. & Liu, H., 2009. Managing large-scale scientific computing in ensemble prediction using BPEL. 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2009, pp. 94-101 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ludäscher, B., Lin, K., Bowers, S., Jaeger-Frank, E., Brodaric, B. & Baru, C., 2006. Managing scientific data: From data integration to scientific workflows. Spec. Pap. Geol. Soc. Am.(397), pp. 109-129 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Maheshwari, K. & Montagnat, J., 2010. Scientific Workflow Development Using Both Visual and Script-Based Representation. Services (SERVICES-1), 2010 6th World Congress on DOI - 10.1109/SERVICES.2010.14, pp. 328-335 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Marru, S., Pierce, M., Herath, C. & Perera, S., 2008. Open Grid Computing Environment's Workflow Suite for E-Science Projects. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.149, pp. 332-333 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
McFerren, G., van Zyl, T., van der Merwe, M. & du Preez, M., 2008. User Requirements for Sensor Web based Scientific Workflows in the Cholera Research Domain. , Vol. 5Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International DOI - 10.1109/IGARSS.2008.4780046, pp. V - 136-V - 139 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
McPhillips, T., Bowers, S. & Ludäscher, B., 2006. Collection-oriented scientific workflows for integrating and analyzing biological data. Lect. Notes Comput. Sci., Vol. 4075 LNBI3rd International Workshop on Data Integration in the Life Sciences, DILS 2006, pp. 248-263 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
McPhillips, T., Bowers, S., Zinn, D. & Ludäscher, B., 2009. Scientific workflow design for mere mortals. Future Generation Computer Systems, Vol. 25(5), pp. 541-551 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-

Miles, S., Deelman, E., Groth, P., Vahi, K., Mehta, G. & Moreau, L., 2007. Connecting scientific data to scientific experiments with provenance. E-Science 2007, 3rd IEEE International Conference on E-Science and Grid Computing, pp. 179-186 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Murray, D. & Consi, T., 2007. Monitoring and Controlling System Performance in the Pioneer Buoy. OCEANS 2007 DOI - 10.1109/OCEANS.2007.4449237, pp. 1-4 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Ngu, A., Jamnagarwala, A., Chin Jr., G., Sivaramakrishnan, C. & Critchlow, T., 2010. Context-aware scientific workflow systems using KEPLER. Int. J. Bus. Process Integr. Manage., Vol. 5(1), pp. 18-31 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
de Oliveira, D., Cunha, L., Tomaz, L., Pereira, V. & Mattoso, M., 2009. Using Ontologies to Support Deep Water Oil Exploration Scientific Workflows. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.17, pp. 364-367 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Pahwa, J., White, R., Jones, A., Burgess, M., Gray, W., Fiddian, N., Sutton, T., Brewer, P., Yesson, C., Caithness, N., Culham, A., Bisby, F., Scoble, M., Williams, P. & Bhagwat, S., 2006. Accessing biodiversity resources in computational environments from workflow applications. Workflows in Support of Large-Scale Science, 2006. WORKS '06. Workshop on DOI - 10.1109/WORKS.2006.5282352, pp. 1-10 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Pandey, S., Voorsluys, W., Rahman, M., Buyya, R., Dobson, J. & Chiu, K., 2009. Brain Image Registration Analysis Workflow for fMRI Studies on Global Grids. Advanced Information Networking and Applications, 2009. AINA '09. International Conference on DOI - 10.1109/AINA.2009.13, pp. 435-442 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Paventhana, A., Takeda, K., Cox, S. & Nicole, D., 2006. Leveraging Windows Workflow Foundation for Scientific Workflows in Wind Tunnel Applications. Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on DOI - 10.1109/ICDEW.2006.71, pp. 65-65 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Pinto, G., Strauch, J., De Souza, J., Oliveira, J., Cardoso, L., Botelho, L., Martorano, L., Morais, E., De Castro Jr., M. & Da Justa Medeiros, S., 2002. A Framework to Support Scientific Knowledge Management: A Case Study in Agro-meteorology. , Vol. 7Proceedings of the Seventh International Conference on CSCW in Design, pp. 320-324 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ramakrishnan, L., Gannon, D. & Plale, B., 2010. WORKEM: Representing and Emulating Distributed Scientific Workflow Execution State. Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference on DOI - 10.1109/CCGRID.2010.89, pp. 283-292 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Rimal, B. & El-Refaey, M., 2010. A Framework of Scientific Workflow Management Systems for Multi-tenant Cloud Orchestration Environment. Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE), 2010 19th IEEE International Workshop on DOI - 10.1109/WETICE.2010.20, pp. 88-93 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Rygg, A., Roe, P. & Wong, O., 2006. GPFlow: An Intuitive Environment for Web Based Scientific Workflow. Grid and Cooperative Computing Workshops, 2006. GCCW '06. Fifth International Conference on, pp. 204-211 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Sfakianakis, S., Koumakis, L., Zacharioudakis, G. & Tsiknakis, M., 2009. Web-Based Authoring and Secure Enactment of Bioinformatics Workflows. Grid and Pervasive Computing Conference, 2009. GPC '09. Workshops at the DOI - 10.1109/GPC.2009.14, pp. 88-95 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Simmhan, Y., van Ingen, C., Szalay, A., Barga, R. & Heasley, J., 2009. Building Reliable Data Pipelines for Managing Community Data Using Scientific Workflows. e-Science, 2009. e-Science '09. Fifth IEEE International Conference on DOI - 10.1109/e-Science.2009.52, pp. 321-328 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Stockinger, H., Donno, F., Eulisse, G., Mazzucato, M. & Steenberg, C., 2005. Matchmaking, datasets and physics analysis. Parallel Processing, 2005. ICPP 2005 Workshops. International Conference Workshops on, pp. 21-28 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Storti, E., 2010. Semantic-driven design and management of KDD processes. 2010 International Symposium on Collaborative Technologies and Systems, CTS 2010, pp. 647-649 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Tan, W., Chard, K., Sulakhe, D., Madduri, R., Foster, I., Soiland-Reyes, S. & Goble, C., 2009. Scientific Workflows as Services in caGrid: A Taverna and gRAVI Approach. Web Services, 2009. ICWS 2009. IEEE International Conference on DOI - 10.1109/ICWS.2009.19, pp. 413-420 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Tan, W., Foster, I. & Madduri, R., 2008. Combining the Power of Taverna and caGrid: Scientific Workflows that Enable Web-Scale Collaboration. Internet Computing, IEEE DOI - 10.1109/MIC.2008.120, Vol. 12(6)Internet Computing, IEEE, pp. 61-68 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-

Wang, L., Huang, Z. & Luo, M., 2009. Supporting dynamic workflow adaptation in a dataflow-constrained workflow net. 2009 International Conference on New Trends in Information and Service Science, NISS 2009, pp. 1000-1005 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	
Wei, J., Zhang, R., Fan, X., Zhang, C. & Gong, Z., 2008. Applying BPEL to Chemical Research on Grid. ChinaGrid Annual Conference, 2008. ChinaGrid '08. The Third DOI - 10.1109/ChinaGrid.2008.30, pp. 235-242 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Wu, Q., Zhu, M., Lu, X., Brown, P., Lin, Y., Gu, Y., Cao, F. & Reuter, M., 2010. Automation and management of scientific workflows in distributed network environments. Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW), 2010 IEEE International Symposium on DOI - 10.1109/IPDPSW.2010.5470720, pp. 1-8 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Wu, W., Uram, T., Wilde, M., Hereld, M. & Papka, M., 2010. A Web 2.0-Based Scientific Application Framework. Web Services (ICWS), 2010 IEEE International Conference on DOI - 10.1109/ICWS.2010.107, pp. 642-643 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Yang, B., Bundy, A., Smail, A. & Dixon, L., 2005. Deductive synthesis of workflows for e-science. , Vol. 12005 IEEE International Symposium on Cluster Computing and the Grid, CCGrid 2005, pp. 168-175 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Yang, X., 2009. Developing an End-to-End Scientific Workflow: a Case Study of Using a Reliable, Lightweight, and Comprehensive Workflow Platform in e-Science. Computing in Science & Engineering DOI - 10.1109/MCSE.2009.211, Vol. PP(99)Computing in Science & Engineering, pp. 1-1 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Yildiz, U., Mouallem, P., Vouk, M., Crawl, D. & Altintas, I., 2010. Fault-Tolerance in Dataflow-Based Scientific Workflow Management. Services (SERVICES-1), 2010 6th World Congress on DOI - 10.1109/SERVICES.2010.93, pp. 336-343 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Zacharioudakis, G., Koumakis, L., Sfakianakis, S. & Tsiknakis, M., 2009. A Semantic Infrastructure for the Integration of Bioinformatics Services. Intelligent Systems Design and Applications, 2009. ISDA '09. Ninth International Conference on DOI - 10.1109/ISDA.2009.221, pp. 367-372 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Zargari, A. & Grise, W., 1999. A comparison of traditional methods of setting-up the experiments with the utilization of computer-based DOE software: validation of a scientific experiment. Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Conference, 1999. Proceedings DOI - 10.1109/EEIC.1999.826221, pp. 275-280 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	E
Zhang, H., Fan, X., Zhang, R., Lin, J., Zhao, Z. & Li, L., 2008. Extending BPEL2.0 for Grid-Based Scientific Workflow Systems. Asia-Pacific Services Computing Conference, 2008. APSCC '08. IEEE, pp. 757-762 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Zhang, J., 2010. Co-Taverna: A Tool Supporting Collaborative Scientific Workflows. Services Computing (SCC), 2010 IEEE International Conference on DOI - 10.1109/SCC.2010.99, pp. 41-48 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Zhang, J., Pennington, D. & Michener, W., 2006. Automatic transformation from geospatial conceptual workflow to executable workflow using GRASS GIS command line modules in Kepler. Lect. Notes Comput. Sci., Vol. 3993 LNCS - IIIICCS 2006: 6th International Conference on Computational Science, pp. 912-919 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Zhang, J., Pennington, D. & Michener, W., 2005. Using web services and scientific workflow for species distribution prediction modeling. Lect. Notes Comput. Sci., Vol. 3739 LNCS6th International Conference on Advances in Web-Age Information Management, WAIM 2005, pp. 610-617 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Zhao, Z., Belloum, A., Yakali, H., Sloot, P. & Hertzberger, B., 2005. Dynamic Work.ow in a Grid Enabled Problem Solving Environment. Computer and Information Technology, 2005. CIT 2005. The Fifth International Conference on DOI - 10.1109/CIT.2005.101, pp. 339-345 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Zhao, Z., Lin, J., Zhang, H., Zhang, R., Chen, Y. & Li, L., 2008. An improved visual BPEL-based environment for scientific workflow. Proceedings - 7th International Conference on Grid and Cooperative Computing, GCC 2008, pp. 435-441 . Base: Compendex EI; Recuperado: 2009.05.08 11:15:00.	E	-
Zhou, Q. & Prasanna, V., 2010. Workflow management of simulation based computation processes in transportation domain. Information Reuse and Integration (IRI), 2010 IEEE International Conference on DOI - 10.1109/IRI.2010.5558970, pp. 19-24 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Zinn, D., 2008. Modeling and optimization of scientific workflows. EDBT 2008 Ph.D. Workshop, held in conjunction with the 11th International Conference on Extending Database Technology, EDBT, pp. 1-10 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Zinn, D., Bowers, S. & LudÄscher, B., 2010. XML-based computation for scientific workflows. 26th IEEE International Conference on Data Engineering, ICDE 2010, pp. 812-815 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Zinn, D., Bowers, S., McPhillips, T. & LudÄscher, B., 2009. Scientific workflow design with data assembly lines. 4th Workshop on Workflows in Support of Large-Scale Science, WORKS '09, in Conjunction with SC 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
van Zyl, T. & Vahed, A., 2009. Using sensorml to describe scientific workflows in distributed Web Service environments. , Vol. 5Geoscience and Remote Sensing Symposium,2009 IEEE International,IGARSS 2009 DOI - 10.1109/IGARSS.2009.5417654, pp. V-375-V-377 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Ailamaki, A., Ioannidis, Y.E. & Livny, M., 1998. Scientific workflow management by database management. Proceedings of the 1998 10th International Conference on Scientific and Statistical Database Management, pp. 190-199 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Shek, E., Kaestle, G. & Dao, S., 1999. ASSISS: an active semi-structured scientific information sharing system. Scientific and Statistical Database Management, 1999. Eleventh International Conference on, pp. 279-279 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Zhang, J., 2006. Ontology-Driven Composition and Validation of Scientific Grid Workflows in Kepler: a Case Study of Hyperspectral Image Processing. Grid and Cooperative Computing Workshops, 2006. GCCW '06. Fifth International Conference on, pp. 282-289 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	E
Abe, H., 1974. Scientific process in deriving diagnosis from the clinical data. Rinsho Byori, Vol. 22(10 suppl), pp. 13- . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Abraham, J., Brazier, P., Chebotko, A., Navarro, J. & Piazza, A., 2010. Distributed Storage and Querying Techniques for a Semantic Web of Scientific Workflow Provenance. Services Computing (SCC), 2010 IEEE International Conference on DOI - 10.1109/SCC.2010.14, pp. 178-185 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Abramson, D., Bethwaite, B., Dinh, M.N., Enticott, C., Firth, S., Garic, S., Harper, I., Lackmann, M., Nguyen, H., Ramdas, T., Russel, A., Schek, S. & Vail, M., 2009. Virtual Microscopy and Analysis Using Scientific Workflows. e-Science, 2009. e-Science '09. Fifth IEEE International Conference on DOI - 10.1109/e-Science.2009.41, pp. 239-246 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Abramson, D., Bethwaite, B., Enticott, C., Garic, S., Peachey, T., Michailova, A. & Amirriazi, S., 2010. Embedding optimization in computational science workflows. J. Comput. Sci., Vol. 1(1), pp. 41-47 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Abramson, D., Bethwaite, B., Enticott, C., Garic, S., Peachey, T., Michailova, A., Amirriazi, S. & Chitters, R., 2009. Robust workflows for science and engineering. 2nd ACM Workshop on Many-Task Computing on Grids and Supercomputers 2009, MTAGS '09, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Abramson, D., Enticott, C. & Altinas, I., 2008. Nimrod/K: Towards massively parallel dynamic Grid workflows. High Performance Computing, Networking, Storage and Analysis, 2008. SC 2008. International Conference for DOI - 10.1109/SC.2008.5215726, pp. 1-11 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Abramson, D., Enticott, C. & Peachey, T., 2008. Parameter Estimation Using Scientific Workflows. eScience, 2008. eScience '08. IEEE Fourth International Conference on, pp. 392-393 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Ahsant, M. & Basney, J., 2009. Workflows in Dynamic and Restricted Delegation. Availability, Reliability and Security, 2009. ARES '09. International Conference on DOI - 10.1109/ARES.2009.92, pp. 17-24 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Aickin, M., 2004. Inference and scientific exploration. J Sci Explor, Vol. 18(3), pp. 361-367 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Aish, A.-M., Ramberg, I.-L. & Wasserman, D., 2002. Measuring attitudes of mental health care staff toward suicidal patients. Arch. Suicide Res., Vol. 6(4), pp. 309-323 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Akau, R. & Larson, D., 1989. Thermal control of space X-ray experiment. J Spacecr Rockets, Vol. 26(5), pp. 297-302 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Akcay, H. & Yager, R., 2010. The Impact of a Science/Technology/Society Teaching Approach on Student Learning in Five Domains. J. Sci. Educ. Technol., pp. 1-10 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-

Akram, A., Meredith, D. & Allan, R., 2006. Evaluation of BPEL to Scientific Workflows. , Vol. 1Cluster Computing and the Grid, 2006. CCGRID 06. Sixth IEEE International Symposium on DOI - 10.1109/CCGRID.2006.44, pp. 269-274 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Alhiyafi, J., Lu, S. & Ram, J., 2009. Simulation of genomic recombination and the influence of DNA sequence diversity on cluster formation and detectability of recombination: A scientific workflow method. Int. J. Comp. Biol. Drug Des., Vol. 2(1), pp. 81-99 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Alich, J. & Castillo, B., 2007. Incorporation of measurement science into flight test instrumentation. U.S. Air Force T and E Days, 2007(Test and Evaluation), pp. 69-74 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Alparslan, C., Tekkaya, C. & Geban, Ö., 2003. Using the conceptual change instruction to improve learning. J. Biol. Educ., Vol. 37(3), pp. 133-137 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Altintas, I., 2008. Lifecycle of Scientific Workflows and their Provenance: A Usage Perspective. Services - Part I, 2008. IEEE Congress on, pp. 474-475 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Altintas, I., Barney, O. & Jaeger-Frank, E., 2006. Provenance collection support in the kepler scientific workflow system. Lect. Notes Comput. Sci., Vol. 4145 LNCSInternational Provenance and Annotation Workshop, IPAW 2006, pp. 118-132 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Altintas, I., Jaeger, E., Lin, K., Ludaescher, B. & Memon, A., 2004. A Web service composition and deployment framework for scientific workflows. Web Services, 2004. Proceedings. IEEE International Conference on DOI - 10.1109/ICWS.2004.1314956, pp. 814-815 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Amagasa, T., Kitagawa, H. & Komano, T., 2007. Constructing a Web Service System for Large-Scale Meteorological Grid Data. e-Science and Grid Computing, IEEE International Conference on, pp. 118-124 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
An, E., Coulson, R., Webb, J., Williams, D., Reibman, R. & Dunn, S., 1999. Shallow water multiplexer development for the South Florida Ocean Measurement Center (SFOMC). , Vol. 2OCEANS '99 MTS/IEEE. Riding the Crest into the 21st Century DOI - 10.1109/OCEANS.1999.805011, pp. 1009-1016 vol.2 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Anand, M., Bowers, S., McPhillips, T. & LudÄscher, B., 2009. Efficient provenance storage over nested data collections. 12th International Conference on Extending Database Technology: Advances in Database Technology, EDBT'09, pp. 958-969 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Anderson, E., Silva, C., Ahrens, J., Heitmann, K. & Habib, S., 2008. Provenance in comparative analysis: A study in cosmology. Comput. Sci. Eng., Vol. 10(3), pp. 30-36 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Andrews, P., Sherwin, T. & Banister, B., 2001. Large-Scale Flexible Storage with SAN Technology Phil Andrews, Tom Sherwin, and Bryan Banister. Mass Storage Systems and Technologies, 2001. MSS '01. Eighteenth IEEE Symposium on, pp. 291-291 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Anwar, Z., Wang, Z., Wang, C., Ni, D., Xu, Y. & Yan, Y., 2009. An integer programming model and heuristic algorithm for automatic scheduling in synchrotron facilities. Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on DOI - 10.1109/ICSMC.2009.5346681, pp. 4048-4053 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Arikawa, M. & Maruno, S., 1998. Analysis of mental models of junior high school students on learning the process of heating using electricity. Jpn. J. Educ. Psychol., Vol. 46(1), pp. 58-67 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ashford, J.W., Sherman, K.A. & Kumar, V., 1989. Advances in Alzheimer therapy: Cholinesterase inhibitors. Neurobiology of Aging, Vol. 10(1), pp. 99-105 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Autenrieth, R., Butler-Purry, K., Page, C., Hurtado, L. & Welch, J., 2009. Enrichment Experiences in Engineering (E3) for teachers summer research program. 2009 ASEE Annual Conference and Exposition, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Avila-Garcia, M., Trefethen, A., Joshi, N., Gleeson, F. & Ba-alawi, W., 2009. Sharing and Reusing Cancer Image Segmentation Algorithms Using Scientific Workflows: Pros and Cons. e-Science, 2009. e-Science '09. Fifth IEEE International Conference on DOI - 10.1109/e-Science.2009.22, pp. 94-101 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-

Ayadi, N. & Lacroix, Z., 2007. Resolving Scientific Service Interoperability With Schema Mapping. Bioinformatics and Bioengineering, 2007. BIBE 2007. Proceedings of the 7th IEEE International Conference on DOI - 10.1109/BIBE.2007.4375600, pp. 448-455 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Azimi, A. & Parsa, S., 2010. Architectural plan for constructing fault tolerable workflow engines based on grid service. , Vol. 4Education Technology and Computer (ICETC), 2010 2nd International Conference on DOI - 10.1109/ICETC.2010.5529657, pp. V4-384-V4-389 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Aziz, K., 2007. The FDA's critical path initiative for medical products. J. Clin. Ligand Assay, Vol. 29(4), pp. 171-176 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Baker, N., McClatchey, R. & Le Goff, J.-M., 1997. Scientific workflow management in a distributed production environment. Enterprise Distributed Object Computing Workshop [1997]. EDOC '97. Proceedings. First International DOI - 10.1109/EDOC.1997.628370, pp. 291-299 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Balasoorya, J., 2010. Cloud Computing Infrastructure for Biological Echo-Systems. Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on DOI - 10.1109/CLOUD.2010.80, pp. 526-527 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Baldrige, K., Greenberg, J., Sudholt, W., Mock, S., Altintas, I., Amoreira, C., Potier, Y., Birnbaum, A., Bhatia, K. & Taufer, M., 2005. The Computational Chemistry Prototyping Environment. Proceedings of the IEEE DOI - 10.1109/JPROC.2004.842747, Vol. 93(3)Proceedings of the IEEE, pp. 510-521 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Balis, B. & Bubak, M., 2008. Monitoring infrastructure for Grid scientific workflows. Workflows in Support of Large-Scale Science, 2008. WORKS 2008. Third Workshop on, pp. 1-10 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Balis, B., Bubak, M. & Pelczar, M., 2007. From Monitoring Data to Experiment Information ? Monitoring of Grid Scientific Workflows. e-Science and Grid Computing, IEEE International Conference on, pp. 77-84 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Bao, Z., Cohen-Boulakia, S., Davidson, S., Eyal, A. & Khanna, S., 2009. Differencing Provenance in Scientific Workflows. Data Engineering, 2009. ICDE '09. IEEE 25th International Conference on DOI - 10.1109/ICDE.2009.103, pp. 808-819 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Barga, R., Jackson, J., Araujo, N., Guo, D., Gautam, N. & Simmhan, Y., 2008. The Trident Scientific Workflow Workbench. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.126, pp. 317-318 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Barnes, B., Menon, A., Mills, R., Bruyns, C., Twombly, A., Smith, J., Montgomery, K. & Boyle, R., 2003. Virtual reality extensions into surgical training and teleoperation. Information Technology Applications in Biomedicine, 2003. 4th International IEEE EMBS Special Topic Conference on DOI - 10.1109/ITAB.2003.1222493, pp. 142-145 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Baron, R.C. & Monahan, J.E., 1966. The Mariner-Mars Data Automation System. Aerospace and Electronic Systems, IEEE Transactions on DOI - 10.1109/TAES.1966.4501997, Vol. AES-2(6-Suppl)Aerospace and Electronic Systems, IEEE Transactions on, pp. 108-113 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Barreto, R., Klasky, S., Podhorszki, N., Mouallem, P. & Vouk, M., 2009. Collaboration portal for petascale simulations. Collaborative Technologies and Systems, 2009. CTS '09. International Symposium on DOI - 10.1109/CTS.2009.5067505, pp. 384-393 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Baser, M., 2006. Effects of conceptual change and traditional confirmatory simulations on pre-service teachers' understanding of direct current circuits. J. Sci. Educ. Technol., Vol. 15(5-6), pp. 367-381 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Batl, K., Ertürk, G. & Kaptan, F., 2010. The awareness levels of pre-school education teachers regarding science process skills. Procedia - Social and Behavioral Sciences, Vol. 2(2)Innovation and Creativity in Education, pp. 1993-1999 . Base: Science Direct; Recuperado: 2010.09.29 01:24:08.	E	-
Benton, T., Solan, M., Travis, J. & Sait, S., 2007. Microcosm experiments can inform global ecological problems. Trends Ecol. Evol., Vol. 22(10), pp. 516-521 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Best, B., Halpin, P., Fujioka, E., Read, A., Qian, S., Hazen, L. & Schick, R., 2007. Geospatial web services within a scientific workflow: Predicting marine mammal habitats in a dynamic environment. Ecol. Informatics, Vol. 2(3 SPEC. ISS.), pp. 210-223 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Bethell, B. & Bellward, G., 1974. Drug education: a problem in moral philosophy. II. A proposal. J. DRUG EDUC., Vol. 4(4), pp. 439-455 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Beyene, N., 2004. The art of space flight exercise hardware: Design and implementation. , Vol. 1A Collection of Technical Papers - AIAA Space 2004 Conference and Exposition, pp. 239-249 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Bhanu, B. & Tan, X., 2003. Fingerprint indexing based on novel features of minutiae triplets. Pattern Analysis and Machine Intelligence, IEEE Transactions on DOI - 10.1109/TPAMI.2003.1195995, Vol. 25(5)Pattern Analysis and Machine Intelligence, IEEE Transactions on, pp. 616-622 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Biau, D., Jolles, B. & Porcher, R., 2010. P value and the theory of hypothesis testing: An explanation for new researchers. Clin. Orthop. Relat. Res., Vol. 468(3), pp. 885-892 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Bienz, R. & Schooley, L., 1990. Communication software design for telescience. Computers and Communications, 1990. Conference Proceedings., Ninth Annual International Phoenix Conference on DOI - 10.1109/PCCC.1990.101733, pp. 892- . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Bilgin, I. & Geban, Ö., 2006. The effect of cooperative learning approach based on conceptual change condition on students' understanding of chemical equilibrium concepts. J. Sci. Educ. Technol., Vol. 15(1), pp. 31-46 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Binnekamp, M. & Ingenbleek, P., 2008. Do "good" food products make others look "bad"? Spin-off effects of labels for sustainable food production in the consumer perception. Br. Food J., Vol. 110(9), pp. 843-864 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Bittencourt, L. & Madeira, E., 2007. Fulfilling Task Dependence Gaps for Workflow Scheduling on Grids. Signal-Image Technologies and Internet-Based System, 2007. SITIS '07. Third International IEEE Conference on, pp. 468-475 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Blangero, J., 2004. Localization and identification of human quantitative trait loci: King Harvest has surely come. Curr. Opin. Genet. Dev., Vol. 14(3), pp. 233-240 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Bonitz, R., Nguyen, T. & Kim, W., 2000. The Mars Surveyor '01 Rover and Robotic Arm. , Vol. 7Aerospace Conference Proceedings, 2000 IEEE DOI - 10.1109/AERO.2000.879291, pp. 235-246 vol.7 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Bosin, A., Dessi, N., Bairappan, M. & Pes, B., 2010. Will SOA accommodate the next step of e-science?. New Technologies of Distributed Systems (NOTERE), 2010 10th Annual International Conference on DOI - 10.1109/NOTERE.2010.5536603, pp. 303-308 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Bowers, S., McPhillips, T., Wu, M. & Ludascher, B., 2007. Project histories: Managing data provenance across collection-oriented scientific workflow runs. , Vol. 4544 LNBI Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 122-138 . Base: Compendex EI; Recuperado: 2009.05.08 11:15:00.	E	-
Braghetto, K., Ferreira, J. & Pu, C., 2007. Using control-flow patterns for specifying business processes in cooperative environments. Proc ACM Symp Appl Computing 2007 ACM Symposium on Applied Computing, pp. 1234-1241 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Brawley, L. & Paskevich, D., 1997. Conducting team building research in the context of sport and exercise. J. Appl. Sport Psychol., Vol. 9(1), pp. 11-40 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Bremer, P.-T., Weber, G., Tierny, J., Pascucci, V., Day, M. & Bell, J., 2009. A Topological Framework for the Interactive Exploration of Large Scale Turbulent Combustion. e-Science, 2009. e-Science '09. Fifth IEEE International Conference on DOI - 10.1109/e-Science.2009.42, pp. 247-254 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Brennan, P.J., 1998. Meta-analysis in chemical toxicity is more often abused than used (against). Toxicology Letters, Vol. 95(Supplement 1), pp. 3-3 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Bretherton, D., Blower, J., Haines, K. & Smith, G., 2009. Running climate models on grids using G-REX. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., Vol. 367(1890), pp. 847-853 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Briscoe, C. & Wells, E., 2002. Reforming Primary Science Assessment Practices: A Case Study of One Teacher's Professional Development Through Action Research. Sci. Educ., Vol. 86(3), pp. 417-435 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Britsch, S., 2009. Differential discourses: The contribution of visual analysis to defining scientific literacy in the early years classroom. Vis. Commun., Vol. 8(2), pp. 207-228 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-

Brochhausen, M., Spear, A., Cocos, C., Weiler, G., Martiñ, L., Anguita, A., Stenzhorn, H., Daskalaki, E., Schera, F., Schwarz, U., Sfakianakis, S., Kiefer, S., DÄrr, M., Graf, N. & Tsiknakis, M., . The ACGT Master Ontology and its applications - Towards an ontology-driven cancer research and management system. J. Biomed. Informatics, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Brodaric, B. & Gahegan, M., 2001. Learning geoscience categories In Situ: Implications for geographic knowledge representation. ACM-GIS 2001: Proceedings of the Ninth ACM International Symposium on: Advances in Geographic Information Systems, pp. 130-135 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Brown, S., 2008. Mesmerizing marketing: A compact cultural history. Eur. Bus. Rev., Vol. 20(4), pp. 350-363 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Brown, W., Bartlett, M. & Dicus, R., 2002. Impact of windowing and subsampling algorithms on acoustic scattering strength databases. , Vol. 2OCEANS '02 MTS/IEEE DOI - 10.1109/OCEANS.2002.1192145, pp. 1258-1263 vol.2 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Buchman, S., Turneaure, J., Lipa, J., Dong, M., Cumbermack, K. & Wang, S., 1998. A superconducting microwave oscillator clock for use on the Space Station. Frequency Control Symposium, 1998. Proceedings of the 1998 IEEE International DOI - 10.1109/FREQ.1998.717950, pp. 534-539 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Burgess, P., Alderman, N., Forbes, C., Costello, A., Coates, L.-A., Dawson, D., Anderson, N., Gilbert, S., Dumontheil, I. & Channon, S., 2006. The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. J. Int. Neuropsychol. Soc., Vol. 12(2), pp. 194-209 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Burke, H., 2000. Discovering patterns in microarray data. Mol. Diagn., Vol. 5(4), pp. 349-357 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Butler, R., 1997. Stories and experiments in social inquiry. Organ. Stud., Vol. 18(6), pp. 927-948 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Butz, B., Miller, S., Duarte, M., Wlodarczyk, B. & Cooper, R., 2006. Work in Progress: An Intelligent Tutoring System for Forensic Biology. Frontiers in Education Conference, 36th Annual DOI - 10.1109/FIE.2006.322736, pp. 19-20 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Büchs, W., Harenberg, A., Zimmermann, J. & Weiß, B., 2003. Biodiversity, the ultimate agri-environmental indicator? Potential and limits for the application of faunistic elements as gradual indicators in agroecosystems. Agric. Ecosyst. Environ., Vol. 98(1-3), pp. 99-123 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Cadopi, M., 1998. La pratique mentale dans les stratégies de préparation des sportifs. Neuro-Psy, Vol. 13(3), pp. 91-96 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Caeiro, M., Nemeth, Z. & Priol, T., 2008. A Chemical Workflow Engine for scientific workflows with dynamicity support. Workflows in Support of Large-Scale Science, 2008. WORKS 2008. Third Workshop on, pp. 1-10 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Cai, Z., Kumar, V. & Schwan, K., 2006. IQ-Paths: Predictably High Performance Data Streams across Dynamic Network Overlays. High Performance Distributed Computing, 2006 15th IEEE International Symposium on DOI - 10.1109/HPDC.2006.1652132, pp. 18-29 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Canuto, E., 2002. Nanoradian digital stabilization of a suspended table for scientific experiments. , Vol. 6American Control Conference, 2002. Proceedings of the 2002 DOI - 10.1109/ACC.2002.1025445, pp. 4946-4947 vol.6 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Canuto, E. & Tonani, S., 2001. Nanoradian digital stabilization of a suspended table for scientific experiments. , Vol. 2Emerging Technologies and Factory Automation, 2001. Proceedings. 2001 8th IEEE International Conference on DOI - 10.1109/ETFA.2001.997748, pp. 645-648 vol.2 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Cao, X.-Z., Tang, X.-Y., Liu, Z. & Zhang, G.-S., 2008. Methodology for determining technical requirements of the standard for ecological protection & rehabilitation in development and construction. J. Ecol. Rural Environ., Vol. 24(1), pp. 77-82 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Carusi, A., Clark, T. & Marshall, M.S., 2009. Web semantics in action: Web 3.0 in e-Science. E-Science Workshops, 2009 5th IEEE International Conference on DOI - 10.1109/ESCIW.2009.5407976, pp. 192-193 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Castanho, M., 2003. Pop-science: Facts or fiction? Friend or foe?. Interciencia, Vol. 28(11), pp. 665-668 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Castano, A., Fukunaga, A., Biesiadecki, J., Neakrase, L., Whelley, P., Greeley, R., Lemmon, M., Castano, R. & Chien, S., 2006. Autonomous Detection of Dust Devils and Clouds on Mars. Image Processing, 2006 IEEE International Conference on DOI - 10.1109/ICIP.2006.313120, pp. 2765-2768 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Cerioti, M., Colombo, C., Scari, E. & Vasile, M., 2005. Halo orbit determination in the mission analysis of the Hevelius - Lunar microsatellite mission. , Vol. 4International Astronautical Federation - 56th International Astronautical Congress 2005, pp. 2754-2764 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chabeli, M., 1999. Student nurses' learning needs & expectations in the clinical learning units.. Curationis, Vol. 22(4), pp. 24-28 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chahal, M., 2010. Off-trial access to experimental cancer agents for the terminally ill: Balancing the needs of individuals and society. J. Med. Ethics, Vol. 36(6), pp. 367-370 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Chappell, S. & Dunlap, K., 2006. Incorporating operator situation awareness into the design Process: A key ingredient for future combat unmanned rotorcraft. , Vol. IIIAHS Internaitonal 62nd Annual Forum, pp. 1927-1936 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chebotko, A., Chang, S., Lu, S., Fotouhi, F. & Yang, P., 2008. Scientific Workflow Provenance Querying with Security Views. Web-Age Information Management, 2008. WAIM '08. The Ninth International Conference on DOI - 10.1109/WAIM.2008.41, pp. 349-356 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Chebotko, A., Lin, C., Fei, X., Lai, Z., Lu, S., Hua, J. & Fotouhi, F., 2007. VIEW: a Visual sciEntificWorkflow management system. Services, 2007 IEEE Congress on, pp. 207-208 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Chen, K. & Chang, T.-Y., 2006. A study on the cognitive thresholds of formal styles. Concurrent Eng Res Appl, Vol. 14(3), pp. 207-218 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chen, K., Zhao, G., Meng, Z. & Lu, H., 2009. Transfer alignment for experiment's INS on space moving platform. Industrial Electronics and Applications, 2009. ICIEA 2009. 4th IEEE Conference on DOI - 10.1109/ICIEA.2009.5138513, pp. 1823-1827 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Chen, S., 2004. A prototype of virtual geographical environment (VGE) for the Tibet Plateau and its applications. , Vol. 5Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International DOI - 10.1109/IGARSS.2004.1370286, pp. 2849-2852 vol.5 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Chin Jr., G. & Lansing, C., 2004. Capturing and supporting contexts for scientific data sharing via the biological sciences collaboratory. Computer Supported Cooperative Work - Conference Proceedings, CSCW 2004, pp. 409-418 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chiu, C., 2002. The effects of collaborative teamwork on secondary science. J. Comput. Assisted Learn., Vol. 18(3), pp. 262-271 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Chiu, D. & Agrawal, G., 2009. Hierarchical Caches for Grid Workflows. Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on DOI - 10.1109/CCGRID.2009.10, pp. 228-235 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Choi, C., Jung, M., Hong, H.-G., Chae, H. & Jeong, D., 2008. Analysis of selection items test for selecting scientifically gifted students in chemistry class. J. Korean Chem. Soc., Vol. 52(3), pp. 295-302 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Civetta, J., 1992. Critical care: How should we evaluate our progress?. CRIT. CARE MED., Vol. 20(12), pp. 1714-1720 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Clarke, K. & Hess, D., 1996. A 1000 A/20 kV/25 kHz-500 kHz Volt-Ampere-Wattmeter for loads with power factors from 0.001 to 1.00. Instrumentation and Measurement, IEEE Transactions on DOI - 10.1109/19.481326, Vol. 45(1)Instrumentation and Measurement, IEEE Transactions on, pp. 142-145 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	E
Classen, J., Young, J., Bottcher, R. & Westerman, P., 2000. Design and analysis of a pilot scale biofiltration system for odorous air. Trans. Am. Soc. Agric. Eng., Vol. 43(1), pp. 111-118 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Clementi, L., Krishnan, S., Goodman, W., Ren, J., Li, W., Arzberger, P., Vareille, G., Dallakyan, S. & Sanner, M., 2008. Services Oriented Architecture for Managing Workflows of Avian Flu Grid. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.37, pp. 582-589 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-

Cloeren, P., 1993. Overview of the latest developments in flat die coextrusion. Proceedings of the Conference on Advances in Extrusion Technology, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Cohen, A., 2010. Developing drug prototypes: pharmacology replaces safety and tolerability?. Nat. Rev. Drug Discov., pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Coppola, R.K. & Toth, E.E., 1995. Developing computational science curricula: the earth vision experience. , Vol. 1Proceedings of the 1995 ACM/IEEE Supercomputing Conference. Part 2 (of 2), pp. 246-254 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Costantini, D., Villepreux, P. & Mandigout, S., 2008. Interest of the scientific contributions in the research of the performance in handball and rugby. Sci. Sports, Vol. 23(1)Intérêt des apports scientifiques dans la recherche de la performance en handball et en rugby, pp. 1-5 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Costantini, D., Villepreux, P. & Mandigout, S., 2008. Intérêt des apports scientifiques dans la recherche de la performance en handball et en rugby. Science & Sports, Vol. 23(1), pp. 1-5 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Coster, D., Basiuk, V., Pereverzev, G., Kalupin, D., ZagÅ³rksi, R., Stankiewicz, R., Huynh, P. & Imbeaux, F., 2010. The European Transport Solver. IEEE Trans Plasma Sci, Vol. 38(9 PART 1), pp. 2085-2092 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Cote, C., 1970. Interrogation, Recording, and Location System. Geoscience Electronics, IEEE Transactions on DOI - 10.1109/TGE.1970.271398, Vol. 8(4)Geoscience Electronics, IEEE Transactions on, pp. 243-245 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Crandell, J. & Kochkin, V., 2005. Scientific damage assessment methodology and practical applications. 2005 Structures Congress and the 2005 Forensic Engineering Symposium - Metropolis and Beyond, pp. 2531-2542 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Curcin, V. & Ghanem, M., 2008. Scientific workflow systems - can one size fit all?. Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International DOI - 10.1109/CIBEC.2008.4786077, pp. 1-9 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	E
Curcin, V., Ghanem, M., Guo, Y., Rowe, A., He, W., Pei, H., Qiang, L. & Li, Y., 2004. IT service infrastructure for integrative systems biology. Services Computing, 2004. (SCC 2004). Proceedings. 2004 IEEE International Conference on DOI - 10.1109/SCC.2004.1357998, pp. 123-131 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Czarnul, P., 2009. A JEE-Based Modelling and Execution Environment for Workflow Applications with Just-in-Time Service Selection. Grid and Pervasive Computing Conference, 2009. GPC '09. Workshops at the DOI - 10.1109/GPC.2009.24, pp. 50-57 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Czuba, K. & Sikora, D., 2010. Phase drift versus temperature measurements of coaxial cables. Microwave Radar and Wireless Communications (MIKON), 2010 18th International Conference on DOI -, pp. 1-3 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Da Costa, G., Dikaiakos, M. & Orlando, S., 2007. Nine months in the life of EGEE: a look from the South. Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 2007. MASCOTS '07. 15th International Symposium on, pp. 281-287 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Da Cruz, S., Campos, M. & Mattoso, M., 2009. Towards a taxonomy of provenance in Scientific Workflow Management Systems. (PART 1)SERVICES 2009 - 5th 2009 World Congress on Services, pp. 259-266 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Dai, Q., 2006. On relation of manufacturing system, manufacturing mode and manufacturing technology. Technology and Innovation Conference, 2006. ITIC 2006. International DOI -, pp. 889-893 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Dale, B., 2001. The sedimentary record of dinoflagellate cysts: Looking back into the future of phytoplankton blooms. Sci. Mar., Vol. 65(SUPPLEMENT 2), pp. 257-272 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Darling, R., 2007. Squirrel foraging preferences: Gone nuts?. Bioscene, Vol. 33(2), pp. 15-18 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Das, S., Spall, J. & Ghanem, R., 2007. Efficient Monte Carlo computation of fisher information matrix using prior information. 2007 Performance Metrics for Intelligent Systems Workshop, PerMIS'07, pp. 256-263 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
De Roure, D. & Goble, C., 2009. Software design for empowering scientists. IEEE Software, Vol. 26(1), pp. 88-95 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

De Roure, D., Goble, C., Bhagat, J., Cruickshank, D., Goderis, A., Michaelides, D. & Newman, D., 2008. myExperiment: Defining the Social Virtual Research Environment. eScience, 2008. eScience '08. IEEE Fourth International Conference on, pp. 182-189 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
De Roure, D., Goble, C. & Stevens, R., 2007. Designing themyexperiment virtual research environment for the social sharing of workflows. E-Science 2007, 3rd IEEE International Conference on E-Science and Grid Computing, pp. 603-610 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
De Sousa Pedrosa, T., Pereira, A., Cardoso, N., Costa, M., Carita, T., Nobre, E., Pedrosa, P., Capeta, N. & EstÃvÃo, S., 2005. 3D environments as social learning tools: The VIRTU@LIS experience. ACM SIGGRAPH 2005 Educators Program, SIGGRAPH '05, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
De Stasio, A., Ertelt, M., Kemmner, W., Leser, U. & Ceccarelli, M., 2009. Exploiting scientific workflows for large-scale gene expression data analysis. Computer and Information Sciences, 2009. ISCIS 2009. 24th International Symposium on DOI - 10.1109/ISCIS.2009.5291850, pp. 448-453 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
De Wet, A., Manduca, C., Wobus, R. & Bettison-Varga, L., 2009. Twenty-two years of undergraduate research in the geosciences - The Keck experience. Spec. Pap. Geol. Soc. Am.(461), pp. 163-172 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Deary, V., Deary, I., McKenna, H., McCance, T., Watson, R. & Hoogbruin, A., 2002. Elisions in the field of caring. J. Adv. Nurs., Vol. 39(1), pp. 96-102 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
DeBurman, S., 2002. Learning how scientists work: Experiential research projects to promotecell biology learning and scientific process skills. Cell Biol. Educ., Vol. 1(4), pp. 154-172 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Deelman, E. & Gil, Y., 2006. Managing Large-Scale Scientific Workflows in Distributed Environments: Experiences and Challenges. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on DOI - 10.1109/E-SCIENCE.2006.261077, pp. 144-144 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Del Basso, S.D., 1994. Microgravity research in a space station environment. Proceedings of the 4th International Conference on Engineering, Construction, and Operations in Space, pp. 473-482 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Demissie, M., Keefer, L., Slowikowski, J. & Stevenson, K., 2006. Evaluating the effectiveness of the Illinois River Conservation Reserve Enhancement Program in reducing sediment delivery. (306)IAHS-AISH Publ., pp. 295-303 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Dent, J., Edwards-Jones, G. & McGregor, M., 1995. Simulation of ecological, social and economic factors in agricultural systems. Agric. Syst., Vol. 49(4), pp. 337-351 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Derevnin, V., Morozov, A. & Stepin, V., 1998. Oceanology controlled buoyancy autonomous buoys. , Vol. 3OCEANS '98 Conference Proceedings DOI - 10.1109/OCEANS.1998.726399, pp. 1812-1816 vol.3 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Dessi, N. & Pes, B., 2009. Towards Scientific Dataspaces. , Vol. 3Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT '09. IEEE/WIC/ACM International Joint Conferences on DOI - 10.1109/WI-IAT.2009.353, pp. 575-578 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Devadze, S. & Sudnitson, A., 2005. FSM Decomposition Software for Education and Research. , Vol. 1Computer as a Tool, 2005. EUROCON 2005.The International Conference on DOI - 10.1109/EURCON.2005.1630063, pp. 839-842 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Devi, T. & Veeraraghavan, P., 2004. Integrated approach in harnessing and reuse of knowledge acquired over the four decades in ISRO. , Vol. 5International Astronautical Federation - 55th International Astronautical Congress 2004, pp. 3234-3243 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Diehl, S. & Gorg, C., 2003. Experiencing natural phenomena with virtual, constructed and mathematical models. , Vol. 1Simulation Conference, 2003. Proceedings of the 2003 Winter, pp. 778-781 Vol.1-778-781 Vol.1 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Dikmenli, M., 2009. Biology student teachers' ideas about purpose of laboratory work. Asia-Pac. Forum Sci. Learn. Teach., Vol. 10(2), pp. 1-15 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Domange, N. & Gregoire, C., 2006. Quality of in situ data about pollutant concentration. TrAC Trends Anal. Chem., Vol. 25(2), pp. 179-189 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E

Dong, H., Devlin, S. & Ebede, C., 1991. Measuring customers' perceptions of service quality. Global Telecommunications Conference, 1991. GLOBECOM '91. 'Countdown to the New Millennium. Featuring a Mini-Theme on: Personal Communications Services, pp. 1753-1756 vol.3-1753-1756 vol.3 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Doren, R., Trexler, J., Gottlieb, A. & Harwell, M., 2009. Ecological indicators for system-wide assessment of the greater everglades ecosystem restoration program. Ecol. Indic., pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Drewnowski, A. & Fulgoni III, V., 2008. Nutrient profiling of foods: Creating a nutrient-rich food index. Nutr. Rev., Vol. 66(1), pp. 23-39 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Du, K., Yuan, Z., Yang, S. & Wang, H., 2008. ArchDB: Towards Parallelized Recovery in Massive Archived Databases. , Vol. 2Future Generation Communication and Networking Symposia, 2008. FGCNS '08. Second International Conference on DOI - 10.1109/FGCNS.2008.79, pp. 1-4 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Dubey, A., Neema, S., Kowalkowski, J. & Singh, A., 2008. Scientific Computing Autonomic Reliability Framework. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.113, pp. 352-353 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Dubey, A., Piccoli, L., Kowalkowski, J., Simone, J., Sun, X.-H., Karsai, G. & Neema, S., 2009. Using Runtime Verification to Design a Reliable Execution Framework for Scientific Workflows. Engineering of Autonomic and Autonomous Systems, 2009. EASE 2009. Sixth IEEE Conference and Workshops on DOI - 10.1109/EASE.2009.13, pp. 87-96 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	E
Dzierzbicka-G?owacka, L. & Zieli?ski, A., 1998. An algorithm for calculating the concentration of phytoplankton in a stratified sea with respect to the daily migration of zooplankton. Part 2. Numerical simulation. Oceanologia, Vol. 40(4), pp. 371-398 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ekanayake, J., Pallickara, S. & Fox, G., 2008. A collaborative framework for scientific data analysis and visualization. 2008 International Symposium on Collaborative Technologies and Systems, CTS'08, pp. 339-346 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
El-Gayyar, M., Leng, Y. & Cremers, A., 2010. Distributed Management of Scientific Workflows in SWIMS. Distributed Computing and Applications to Business Engineering and Science (DCABES), 2010 Ninth International Symposium on DOI - 10.1109/DCABES.2010.73, pp. 327-331 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
El-Gayyar, M., Leng, Y., Shumilov, S. & Cremers, A., 2009. New Execution Paradigm for Data-Intensive Scientific Workflows. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.22, pp. 334-339 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Elliott, D. & Hollins, B., 1995. Product Evaluation: Theoretical and Practical Considerations. Aust. Crit. Care, Vol. 8(2), pp. 14-19 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ellqvist, T., Koop, D., Freire, J., Silva, C. & Stromback, L., 2008. Using Mediation to Achieve Provenance Interoperability (Extended Abstract). eScience, 2008. eScience '08. IEEE Fourth International Conference on, pp. 398-399 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Ellqvist, T., Koop, D., Freire, J., Silva, C. & Stromback, L., 2009. Using Mediation to Achieve Provenance Interoperability. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.68, pp. 291-298 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Emereole, H., 2008. LEARNERS' AND TEACHERS' CONCEPTUAL KNOWLEDGE OF SCIENCE PROCESSES: THE CASE OF BOTSWANA. Int. J. Sci. Math. Educ., pp. 1-24 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Emmerich, W., Butchart, B., Chen, L., Wassermann, B. & Price, S., 2005. Grid service orchestration using the Business Process Execution Language (BPEL). J. Grid Comput., Vol. 3(3-4), pp. 283-304 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Engelmann, H.D. & Severin, D., 1973. Mass Spectrometric Analysis with the Aid of a Process Computer.. Chem-Ing-Tech, Vol. 45(19)MASSENSPEKTROMETRISCHE ANALYSE MIT HILFE EINES PROZESSRECHNERS., pp. 1166-1170 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Enke, H., Steinmetz, M., Adorf, H.-M., Beck-Ratzka, A., Breitling, F., BrÄ¼semeister, T., Carlson, A., Ensslin, T., HÄ¼ggqvist, M., Nickelt, I., Radke, T., Reinefeld, A., Reiser, A., Scholl, T., Spurzem, R., Steinacker, J., Voges, W., WambsganÄ¼, J. & White, S., . AstroGrid-D: Grid technology for astronomical science. New Astron, Vol. 16(2), pp. 79-93 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-

Erb, A. & van der Heijden, A., 2005. The STAR Program: an innovative approach to teaching the scientific method to high school students. OCEANS, 2005. Proceedings of MTS/IEEE DOI - 10.1109/OCEANS.2005.1639812, pp. 549-554 Vol. 1 . Base: IEEEXplore; Recuperado: 2010.09.29 12:58:54.	E	-
Esquembre, F., 2004. Easy Java Simulations: A software tool to create scientific simulations in Java. Comput Phys Commun, Vol. 156(2), pp. 199-204 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Etkina, E., Karelina, A. & Ruibal-Villasenor, M., 2008. How long does it take? A study of student acquisition of scientific abilities. Phys. Rev. Spec. Top. Phys. Educ. Res., Vol. 4(2), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Fagard, R., 1996. Advantages and disadvantages of the meta-analysis approach. J. HYPERTENS. SUPPL., Vol. 14(2), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Fan, S.-K. & Del Castillo, E., 1999. Calculation of an optimal region of operation for dual response systems fitted from experimental data. J Oper Res Soc, Vol. 50(8), pp. 826-836 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Fan, W.-H., 2009. Is computer simulation independent life. Xitong Fangzhen Xuebao, Vol. 21(24), pp. 7982-7984 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Faust, D. & Meehl, P., 1992. Using scientific methods to resolve questions in the history and philosophy of science: Some illustrations. BEHAV. THER., Vol. 23(2), pp. 195-211 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Fedorochko, J., 1984. STANDARDIZATION REQUIRED FOR CAD/CAM SYSTEM.. Thirty-third Annual Conference Proceedings - Standards Engineering Society: Standards in the Age of Automation., pp. 99-130 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Feinstein, A., 1978. Clinical biostatistics. XLII. The architecture of cross-sectional research (Part I). CLIN. PHARMACOL. THER., Vol. 23(1), pp. 81-96 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ferraris, A., 2003. European microgravity facilities in the Columbus laboratory: Verification of the operational scenario and new designs. , Vol. 354th International Astronautical Congress of the International Astronautical Federation (IAF), the International Academy of Astronautics and the International Institute of Space Law, pp. 971-981 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Ferreira, J.E., Wu, Q., Malkowski, S. & Pu, C., 2010. Towards Flexible Event-Handling in Workflows through Data States. Services (SERVICES-1), 2010 6th World Congress on DOI - 10.1109/SERVICES.2010.60, pp. 344-351 . Base: IEEEXplore; Recuperado: 2010.09.29 12:58:23.	E	-
Fialho, I.J. & Thampi, S.K., 2000. Interplay between hardware and control system design in the development of the active rack isolation system. , Vol. 1(I)41st AIAA/ASME/AHS/ASC Structures, Structural Dynamics, and Materials Conference and ExhibitAIAA/ASME/AHS Adaptive Structures ForumAIAA Non-Deterministic Approaches ForumAIAA Space Inflatables Forum, pp. 575-583 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Fichter, L., 1988. Process-response modeling and the scientific process. Journal of Geological Education, Vol. 36(2), pp. 72-78 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Forkert, T., Kloss, G., Krause, C. & Schreiber, A., 2004. Techniques for wrapping scientific applications to CORBA components. High-Level Parallel Programming Models and Supportive Environments, 2004. Proceedings. Ninth International Workshop on, pp. 100-108 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Fraser, W. & Maguvhe, M., 2008. Teaching life sciences to blind and visually impaired learners. J. Biol. Educ., Vol. 42(2), pp. 84-89 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Frazier, B., Stone, R. & Thompson, P., 1986. SELECTION OF ORBITS FOR THE CRRES DUAL MISSION SATELLITE.. J Astronaut Sci, Vol. 34(3), pp. 315-329 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Freire, J., Silva, C., Callahan, S., Santos, E., Scheidegger, C. & Vo, H., 2006. Managing rapidly-evolving scientific workflows. Lect. Notes Comput. Sci., Vol. 4145 LNCSInternational Provenance and Annotation Workshop, IPAW 2006, pp. 10-18 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Fritz, T. & Cessna, J., 1975. ATS-6 NOAA Low Energy Proton Experiment. Aerospace and Electronic Systems, IEEE Transactions on, Vol. AES-11(6), pp. 1145-1154 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-

Fujimoto, N., 2008. On Non-Approximability of Coarse-Grained Workflow Grid Scheduling. Parallel Architectures, Algorithms, and Networks, 2008. I-SPAN 2008. International Symposium on, pp. 127-132 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Gabor, B. & Kemme, B., 2006. Exp-WF:Workflow Support for Laboratory Information Systems. Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on, pp. 69-69 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Gadelha, L. & Mattoso, M., 2008. Kairos: An Architecture for Securing Authorship and Temporal Information of Provenance Data in Grid-Enabled Workflow Management Systems. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.161, pp. 597-602 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Garrison, T. & Schrock, S., 1992. QuickStar: Rapid access to space for small scientific payloads. Acta Astronaut, Vol. 26(8-10), pp. 729-740 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Gaston, L. & Gagnon, R., 1996. The role of process research in manual development. Clin. Psychol. Sci. Pract., Vol. 3(1), pp. 13-24 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Geddes, J., Mackay, C., Lloyd, S., Simpson, A., Power, D., Russell, D., Katarova, M., Rossor, M., Fox, N., Fletcher, J., Hill, D., McLeish, K., Hajnal, J., Lawrie, S., Job, D., McIntosh, A., Wardlaw, J., Sandercock, P., Palmer, J., Perry, D., Procter, R., Ure, J., Bath, P. & Watson, G., 2006. The Challenges of Developing a Collaborative Data and Compute Grid for Neurosciences. Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on DOI - 10.1109/CBMS.2006.156, pp. 81-86 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Geib, R., 2004. A Closer Look At ECDA Pre-Assessment. Pipeline Gas J, Vol. 231(3), pp. 68-70 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Germann, P., Aram, R. & Burke, G., 1996. Identifying Patterns and Relationships among the Responses of Seventh-Grade Students to the Science Process Skill of Designing Experiments. J. Res. Sci. Teach., Vol. 33(1), pp. 79-99 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Gertz, M., 2009. The role of security in scientific data management. 2nd SIGSPATIAL ACM GIS 2009 International Workshop on Security and Privacy in GIS and LBS, SPRINGL 2009, pp. 1- . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ghaedi, A. & Nekoui, M., 2008. 3-axes satellite attitude control based on biased angular momentum. Power Electronics and Motion Control Conference, 2008. EPE-PEMC 2008. 13th DOI - 10.1109/EPEPEMC.2008.4635407, pp. 1054-1057 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Ghosh, S. & Teschmacher, L., 2002. Comparisons for search designs using search probabilities. J. Stat. Plann. Inference, Vol. 104(2), pp. 439-458 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Gil, Y., 2009. From data to knowledge to discoveries: Artificial intelligence and scientific workflows. Sci Program, Vol. 17(3), pp. 231-246 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Gil, Y., Deelman, E., Ellisman, M., Fahringer, T., Fox, G., Gannon, D., Goble, C., Livny, M., Moreau, L. & Myers, J., 2007. Examining the Challenges of Scientific Workflows. Computer, Vol. 40(12), pp. 24-32 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Gil, Y., Kim, J., Florez, G., Ratnakar, V. & González-Calero, P., 2009. Workflow matching using semantic metadata. 5th International Conference on Knowledge Capture, K-CAP'09, pp. 121-128 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Glenn, S. & Schofield, O., 2002. The New Jersey Shelf Observing System. , Vol. 30 OCEANS '02 MTS/IEEE DOI - 10.1109/OCEANS.2002.1191887, pp. 1680-1687 vol.3 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Gobert, J., Montalvo, O., Toto, E., Pedro, M. & Baker, R., 2010. The science assistments project: Scaffolding scientific inquiry skills. Lect. Notes Comput. Sci., Vol. 6095 LNCS(PART 2)10th International Conference on Intelligent Tutoring Systems, ITS 2010, pp. 445- . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Goderis, A., Brooks, C., Altintas, I., Lee, E. & Goble, C., 2009. Heterogeneous composition of models of computation. Future Gener Comput Syst, Vol. 25(5), pp. 552-560 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Gopalakrishnan, V. & Buchanan, B., 1998. Representing and learning temporal relationships among experimental variables. Temporal Representation and Reasoning, 1998. Proceedings. Fifth International Workshop on DOI - 10.1109/TIME.1998.674144, pp. 148-155 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-

Grant Willhite, D. & Wright, S., 2009. Detergent-based isolation of yeast membrane rafts: An inquiry-based laboratory series for the undergraduate cell biology or biochemistry lab. <i>Biochem. Mol. Biol. Educ.</i> , Vol. 37(6), pp. 349-354 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Grant, C., 1984. Quality and quantity issues in environmental analysis. <i>ASTM STANDARDIZAT. NEWS</i> , Vol. 12(2), pp. 13-16 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Greenwald, M., 2004. Beyond benchmarking - How experiments and simulations can work together in plasma physics. <i>Comput Phys Commun</i> , Vol. 164(1-3)Proceedings of the 18th International Conference, pp. 1-8 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Grigori, D., Corrales, J., Bouzeghoub, M. & Gater, A., 2010. Ranking BPEL Processes for Service Discovery. <i>Services Computing</i> , IEEE Transactions on DOI - 10.1109/TSC.2010.6, Vol. 3(3)Services Computing, IEEE Transactions on, pp. 178-192 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Groth, P. & Gil, Y., 2009. Analyzing the Gap between Workflows and their Natural Language Descriptions. <i>Services - I</i> , 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.55, pp. 299-305 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Groth, P., Miles, S., Fang, W., Wong, S., Zauner, K.-P. & Moreau, L., 2005. Recording and using provenance in a protein compressibility experiment. <i>High Performance Distributed Computing</i> , 2005. HPDC-14. Proceedings. 14th IEEE International Symposium on DOI - 10.1109/HPDC.2005.1520960, pp. 201-208 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Guan, Z., Hernandez, F., Bangalore, P., Gray, J., Skjellum, A., Velusamy, V. & Liu, Y., 2006. Grid-flow: A Grid-enabled scientific workflow system with a Petri-net-based interface. <i>Concurrency Comput. Pract. Exper.</i> , Vol. 18(10), pp. 1115-1140 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Gudivada, V. & Raghavan, V., 1995. Content based image retrieval systems. <i>Computer DOI - 10.1109/2.410145</i> , Vol. 28(9)Computer, pp. 18-22 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Guru, S., Taylor, P., Neuhaus, H., Shu, Y., Smith, D. & Terhorst, A., 2008. Hydrological Sensor Web for the South Esk Catchment in the Tasmanian state of Australia. <i>eScience</i> , 2008. eScience '08. IEEE Fourth International Conference on, pp. 432-433 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Han, X., Song, Z., Sun, W., Xie, J. & Li, L., 2010. A Software Approach for Real-Time Control of Operation and Experiments at the WHMFC Pulsed High Magnetic Field Facility. <i>Applied Superconductivity</i> , IEEE Transactions on DOI - 10.1109/TASC.2009.2039789, Vol. 20(3)Applied Superconductivity, IEEE Transactions on, pp. 1697-1700 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Harris, H.M., 1986. INTERACTIVE MISSION PLANNING FOR A SPACE SHUTTLE FLIGHT EXPERIMENT - A CASE HISTORY.. , Vol. 16(2)Aerospace Simulation II, Proceedings of the Conference., pp. 143-151 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Harris, T., 2003. Data Models and the Acquisition and Manipulation of Data. <i>Philos. Sci.</i> , Vol. 70(5), pp. 1508-1517 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Harrower, M., MacEachren, A. & Griffin, A., 2000. Developing a geographic visualization tool to support Earth Science learning. <i>Cartogr. Geogr. Inf. Sci</i> , Vol. 27(4), pp. 279-293 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Hartsfield, C.H.W., 1976. The U.S. space shuttle mission. , Vol. 15Decision and Control including the 15th Symposium on Adaptive Processes, 1976 IEEE Conference on DOI - 10.1109/CDC.1976.267773, pp. 445-445 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Hashim, R. & Abidin, S., 2010. User interface development for a computer-based user study: The Universal Model approach. 24th IEEE International Conference on Advanced Information Networking and Applications Workshops, WAINA 2010, pp. 109-114 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Hatfield, A. & Hipel, K., 2002. Risk and systems theory. <i>Risk Anal.</i> , Vol. 22(6), pp. 1043-1057 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
He, J., Zhang, Y., Huang, G. & Pang, C., 2009. A novel time computation model based on algorithm complexity for data intensive scientific workflow design and scheduling. <i>Concurrency Comput. Pract. Exper.</i> , Vol. 21(16), pp. 2070-2083 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
He, M. & ping Du, Y., 2010. P-top-k queries in probabilistic framework from information extraction models. , Vol. 5Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on DOI - 10.1109/FSKD.2010.5569526, pp. 2376-2379 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-

Heim, S., 1992. Is cancer cytogenetics reducible to the molecular genetics of cancer cells?. GENES CHROMOSOMES CANCER, Vol. 5(3), pp. 188-196 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Heitmeyer, C. & Jeffords, R., 2007. Applying a Formal Requirements Method to Three NASA Systems: Lessons Learned. Aerospace Conference, 2007 IEEE DOI - 10.1109/AERO.2007.352764, pp. 1-10 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Helmreich, B., Schreff, D. & Wilderer, P., 2000. Full scale experiences with small sequencing batch reactor plants in Bavaria. Water Sci. Technol., Vol. 41(1)The 4th IAWQ International Specialised Conference on Small Wastewater Treatment Plants, pp. 89-96 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Hemler, R., 2000. Key elements of the user-friendly, GFDL SKYHI general circulation model. Sci Program, Vol. 8(1), pp. 39-47 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Henderson-Sellers, B., 2005. Evaluating the feasibility of method engineering for the creation of agent-oriented methodologies. Lect. Notes Comput. Sci., Vol. 3690 LNAI4th International Central and Eastern European Conference on Multi-Agent Systems, CEEMAS 2005, pp. 142-152 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Henry, A., Chipman, R. & Hu, T.-H.G., 1993. Quasi-steady state microgravity analysis for the Space Station Freedom using multibody dynamic simulator. (pt 4)34th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, pp. 2039-2043 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Herath, C. & Plale, B., 2010. Streamflow Programming Model for Data Streaming in Scientific Workflows. Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference on DOI - 10.1109/CCGRID.2010.116, pp. 302-311 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Ho, T. & Abramson, D., 2006. A Unified Data Grid Replication Framework. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on DOI - 10.1109/E-SCIENCE.2006.261136, pp. 52-52 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Hoban, S., Sachs, J., Laughlin, D., Yesha, Y. & Keating, J., 2003. Science investigation system for telescopes in education research. System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on DOI - 10.1109/HICSS.2003.1173738, pp. 7 pp.- . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Holland, S., Nunnally, C., Armstrong, S. & Laufer, G., 2002. University of Virginia suborbital infrared sensing experiment. , Vol. 4710Thermosense XXIV, pp. 638-648 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Holt, P. & Light, J., 2004. From fashion to focus: Automation as a practical art. Eur. Pharm. Contract.(SPRING), pp. 12-16 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Hongbiao, L., Feng, L. & Wanjun, Y., 2010. The research of scientific workflow engine. Software Engineering and Service Sciences (ICSESS), 2010 IEEE International Conference on DOI - 10.1109/ICSESS.2010.5552352, pp. 412-414 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Hoonstra, E. & Notermans, S., 2001. Quantitative microbiological risk assessment. Int. J. Food Microbiol., Vol. 66(1-2), pp. 21-29 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Horre, W., Lee, K., Hughes, D., Michiels, S. & Joosen, W., 2009. A Graph Based Approach to Supporting Reconfiguration in Wireless Sensor Networks. Networks and Communications, 2009. NETCOM '09. First International Conference on DOI - 10.1109/NetCoM.2009.18, pp. 326-331 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Howard, J., Rice, K., Coe, R. & McCollum, M., 2002. International Space Station communication systems. , Vol. 3Wireless Personal Multimedia Communications, 2002. The 5th International Symposium on DOI - 10.1109/WPMC.2002.1088363, pp. 1172-1176 vol.3 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Howe, B., Green-Fishback, H. & Maier, D., 2009. Scientific mashups: Runtime-configurable data product ensembles. Lect. Notes Comput. Sci., Vol. 5566 LNCS21st International Conference on Scientific and Statistical Database Management, SSDBM 2009, pp. 19-36 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Howes, E., 2008. Educative experiences and early childhood science education: A Deweyan perspective on learning to observe. Teach. Teach. Educ., Vol. 24(3), pp. 536-549 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Hoyle, D., Delderfield, M., Kitching, L., Smith, G., Crowther, P. & Buchan, I., 2009. Shared genomics: A platform for emerging interpretation of genetic epidemiology. 2009 22nd IEEE International Symposium on Computer-Based Medical Systems, CBMS 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Hsu, C.-H., Chen, T.-L. & Lee, K.-H., 2009. QoS Based Parallel File Transfer for Grid Economics. , Vol. 1Multimedia Information Networking and Security, 2009. MINES '09. International Conference on DOI - 10.1109/MINES.2009.282, pp. 653-657 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-

Hsu, Y.-S., 2004. Using the Internet to develop students' capacity for scientific inquiry. J. Educ. Comput. Res., Vol. 31(2), pp. 137-161 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Huang, R., Casanova, H. & Chien, A.A., 2007. Automatic resource specification generation for resource selection. Supercomputing, 2007. SC '07. Proceedings of the 2007 ACM/IEEE Conference on DOI - 10.1145/1362622.1362638, pp. 1-11 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Huibin, X., Xinran, L. & Hong, Z., 2009. A Grid Workflow Process Engine: Architecture and Simulation. , Vol. 1Information Technology and Applications, 2009. IFITA '09. International Forum on DOI - 10.1109/IFITA.2009.87, pp. 245-248 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Husson, J.-L., 2005. Innovations technologiques : quels progrès ? Quelle chirurgie ? L'orthopédie à la croisée des chemins. Réflexion pluridisciplinaire. Revue de Chirurgie Orthopédique et Réparatrice de l'Appareil Moteur, Vol. 91(5, Supplement 1), pp. 137-138 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Husson, J.-L., Destailats, J.-M. & Husson, C., 2005. Technological innovations: New perspectives for " multidisciplinary" orthopedic surgery. Rev. Chir. Orthop. Repar. Appar. Mot., Vol. 91(5 SUPPL.)Innovations technologiques: Quels progrès? Quelle chirurgie? L'orthopédie à la croisée des chemins. réflexion pluridisciplinaire, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ibrahim, D., 2009. The Concept of Web Service Versioning in Provenance. Network-Based Information Systems, 2009. NBIS '09. International Conference on DOI - 10.1109/NBIS.2009.65, pp. 469-474 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Imlau, K.-P. & Heller, T., 2007. New steel solutions for the worldwide car industry. Steel Res. Int., Vol. 78(3), pp. 180-184 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ingram, M., 2007. Biology and beyond: The science of "back to nature" farming in the United States. Ann. Assoc. Am. Geogr., Vol. 97(2), pp. 298-312 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ivkovi?, M., Wu, H., Spencer, D. & McRae, T., 2007. Modelling the effects of stem sweep, branch size and wood stiffness of radiata pine on structural timber production. Aust Forest, Vol. 70(3), pp. 173-184 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Jackson, A., Hettinga, D., Mead, J. & Mercer, C., 2009. Using consensus methods in developing clinical guidelines for exercise in managing persistent low back pain. Physiotherapy, Vol. 95(4), pp. 302-311 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Jahanshahi, M.H., 1980. VOYAGER HIGH GAIN ANTENNA POINTING CALIBRATION.. J Astronaut Sci, Vol. 28(2), pp. 167-193 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Jararweh, Y., Hary, A., Al-Nashif, Y., Hariri, S., Akoglu, A. & Jenerette, D., 2009. Accelerated discovery through integration of Kepler with data turbine for ecosystem research. 7th IEEE/ACS International Conference on Computer Systems and Applications, AICCSA-2009, pp. 1005-1012 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Jarman, A., 2009. Scientific Workflows for the Enterprise. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.16, pp. 360-363 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Jiang, N. & Parashar, M., 2008. Programming support for sensor-based scientific applications. Parallel and Distributed Processing, 2008. IPDPS 2008. IEEE International Symposium on, pp. 1-5 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Johnson, J., 2003. Chemists may gain with DOE R&D plan. Chem Eng News, Vol. 81(49), pp. 28-31 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Jones, M. & Taylor, A., 2009. Developing a sense of scale: Looking backward. J. Res. Sci. Teach., Vol. 46(4), pp. 460-475 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Juve, G., Deelman, E., Vahi, K., Mehta, G., Berriman, B., Berman, B. & Maechling, P., 2009. Scientific workflow applications on Amazon EC2. E-Science Workshops, 2009 5th IEEE International Conference on DOI - 10.1109/ESCIW.2009.5408002, pp. 59-66 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Kaber, D., Stoll, N. & Thurow, K., 2007. Human-automation interaction strategies for life science applications: Implications and future research. 3rd IEEE International Conference on Automation Science and Engineering, IEEE CASE 2007, pp. 615-620 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kaber, D., Stoll, N., Thurow, K., Green, R., Kim, S.-H. & Mosaly, P., 2009. Human-automation interaction strategies and models for life science applications. Hum Fact Ergon Manuf, Vol. 19(6), pp. 601-621 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-

Kandaswamy, G., Mandal, A. & Reed, D., 2008. Fault Tolerance and Recovery of Scientific Workflows on Computational Grids. Cluster Computing and the Grid, 2008. CCGRID '08. 8th IEEE International Symposium on, pp. 777-782 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Kapetanios, E. & Kramer, R., 1995. Knowledge-based system approach for scientific data analysis and the notion of metadata. Proceedings of the 1995 14th IEEE Symposium on Mass Storage Systems, pp. 274-283 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kapetanios, E. & Kramer, R., 1995. A knowledge-based system approach for scientific data analysis and the notion of metadata. Mass Storage Systems, 1995. 'Storage - At the Forefront of Information Infrastructures', Proceedings of the Fourteenth IEEE Symposium on, pp. 274-283 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Kapetanios, E. & Norrie, M., 1997. Data mining and modeling in scientific databases. Scientific and Statistical Database Management, 1997. Proceedings., Ninth International Conference on, pp. 24-27 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Karavanic, K. & Miller, B., 1997. Experiment Management Support for Performance Tuning. Supercomputing, ACM/IEEE 1997 Conference DOI - 10.1109/SC.1997.10013, pp. 8-8 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Karsli, F., Sahin, Ç. & Ayas, A., 2009. Determining science teachers' ideas about the science process skills: a case study. Procedia - Social and Behavioral Sciences, Vol. 1(1)World Conference on Educational Sciences, Nicosia, North Cyprus, 4-7 February 2009 - New Trends and Issues in Educational Sciences, pp. 890-895 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Kaulemu, D., 2002. Moral theories as tools in programme evaluation. J. Soc. Dev. Afr., Vol. 17(1), pp. 7-18 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Keefe, D., 2010. Integrating Visualization and Interaction Research to Improve Scientific Workflows. Computer Graphics and Applications, IEEE DOI - 10.1109/MCG.2010.30, Vol. 30(2)Computer Graphics and Applications, IEEE, pp. 8-13 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Kelling, S., Hochachka, W., Fink, D., Riedewald, M., Caruana, R., Ballard, G. & Hooker, G., 2009. Data-intensive science: A new paradigm for biodiversity studies. BioScience, Vol. 59(7), pp. 613-620 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Kettler, B. & Darden, L., 1993. Protein sequencing experiment planning using analogy.. Proc Int Conf Intell Syst Mol Biol, Vol. 1, pp. 216-224 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Khan, F., Han, Y., Pillana, S. & Brezany, P., 2010. An Ant-Colony-Optimization Based Approach for Determination of Parameter Significance of Scientific Workflows. Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on DOI - 10.1109/AINA.2010.24, pp. 1241-1248 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Khan, F., Han, Y., Pillana, S. & Brezany, P., 2009. Estimation of Parameters Sensitivity for Scientific Workflows. Parallel Processing Workshops, 2009. ICPPW '09. International Conference on DOI - 10.1109/ICPPW.2009.9, pp. 457-462 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Khan, F., Han, Y., Pillana, S. & Brezany, P., 2008. Provenance Support for Grid-Enabled Scientific Workflows. Semantics, Knowledge and Grid, 2008. SKG '08. Fourth International Conference on DOI - 10.1109/SKG.2008.86, pp. 173-180 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Khan, M., Sing, T. & Mathews, E., 2004. Learning science beyond the classroom with technology. Proceedings of the Seventh IASTED International Conference on Computers and Advanced Technology in Education, pp. 197-201 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kim, K.-J., 1990. Survey of synchrotron radiation devices producing circular or variable polarization. , Vol. 1345Advanced X-Ray/EUV Radiation Sources and Applications, pp. 116-124 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kirby, R. & Silva, C., 2008. The need for verifiable visualization. IEEE Comput Graphics Appl, Vol. 28(5), pp. 78-83 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Kirk, S.J., 1990. Integrating VE into the design process. , Vol. 25SAVE Proceedings - 1990 International Conference of the Society of American Value Engineers, pp. 48-57 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Klasky, S., Ludaescher, B. & Parashar, M., 2006. The Center for Plasma Edge Simulation Workflow Requirements. Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on, pp. 73-73 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-

Klos, S., Henke, C., Kieren, C., Walpuski, M. & Sumfleth, E., 2008. Scientific experiments and chemical knowledge. Two different competences. Z. Padagog., Vol. 54(3)Naturwissenschaftliches experimentieren und chemisches fachwissen - Zwei verschiedene kompetenzen, pp. 304-321 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kohnen, W., 2008. Manned underwater vehicles: A relevant and cornerstone technology for the future assessment of ocean eco-systems. OCEANS 2008 DOI - 10.1109/OCEANS.2008.5152133, pp. 1-5 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Kommineni, J., Abramson, D. & Tan, J., 2006. Communication over a Secured Heterogeneous Grid with the GriddleS Runtime Environment. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on, pp. 21-21 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Kong, S., Yeung, Y. & Wu, X., 2009. An experience of teaching for learning by observation: Remote-controlled experiments on electrical circuits. Comput Educ, Vol. 52(3), pp. 702-717 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Koop, D., Scheidegger, C., Callahan, S., Freire, J. & Silva, C., 2008. VisComplete: Automating Suggestions for Visualization Pipelines. Visualization and Computer Graphics, IEEE Transactions on DOI - 10.1109/TVCG.2008.174, Vol. 14(6)Visualization and Computer Graphics, IEEE Transactions on, pp. 1691-1698 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Koumakis, L., Moustakis, V., Tsiknakis, M., Kafetzopoulos, D. & Potamias, G., 2009. Supporting genotype-to-phenotype association studies with grid-enabled knowledge discovery workflows. Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE DOI - 10.1109/IEMBS.2009.5333882, pp. 6958-6962 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Kramer, H., 2001. Microbore catheters: keys to successful design and manufacture, Part II.. Med Device Technol, Vol. 12(9), pp. 16-19 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kreylos, O., Bawden, G., Bernardin, T., Billen, M., Cowgill, E., Gold, R., Hamann, B., Jadamec, M., Kellogg, L., Staadt, O. & Sumner, D., 2006. Enabling scientific workflows in virtual reality. , Vol. 2006VRCIA 2006: ACM International Conference on Virtual Reality Continuum and its Applications, pp. 155-162 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Krishnan, G., 1997. Application of artificial intelligence (AI) concepts to the development of space flight parts approval model. Eur Space Agency Spec Publ ESA SP(395), pp. 147-154 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Kumar, V., Kurc, T., Ratnakar, V., Kim, J., Mehta, G., Vahi, K., Nelson, Y., Sadayappan, P., Deelman, E., Gil, Y., Hall, M. & Saltz, J., 2010. Parameterized specification, configuration and execution of data-intensive scientific workflows. Cluster Comput., Vol. 13(3), pp. 315-333 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Kumar, V., Sadayappan, P., Mehta, G., Vahi, K., Deelman, E., Ratnakar, V., Kim, J., Gil, Y., Hall, M., Kurc, T. & Saltz, J., 2009. An integrated framework for parameter-based optimization of scientific workflows. 18th ACM International Symposium on High Performance Distributed Computing, HPDC 09, Co-located with the 2009 International Symposium on High Performance Distributed Computing Conference, HPDC'09, pp. 177-186 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Kuo, R., Chang, M. & Heh, J.-S., 2001. Applying interactive mechanism to virtual experiment environment on WWW with experiment action language. Advanced Learning Technologies, 2001. Proceedings. IEEE International Conference on DOI - 10.1109/ICALT.2001.943925, pp. 289-290 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Kur?un, O. & Favorov, O., 2004. SINBAD automation of scientific discovery: From factor analysis to theory synthesis. Nat. Comput., Vol. 3(2), pp. 207-233 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Kurki-Suonio, K., 2010. Principles Supporting the Perceptual Teaching of Physics: A "Practical Teaching Philosophy". Sci. Educ., pp. 1-33 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Lam, T., Hauser, N., Götz, A., Hathaway, P., Franceschini, F., Rayner, H. & Zhang, L., 2006. GumTree-An integrated scientific experiment environment. Phys B Condens Matter, Vol. 385-386, pp. 1330-1332 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lampsas, P., Loukopoulos, T., Dimopoulos, F. & Athanasiou, M., 2006. Scheduling Independent Tasks in Heterogeneous Environments under Communication Constraints. Parallel and Distributed Computing, Applications and Technologies, 2006. PDCAT '06. Seventh International Conference on, pp. 373-378 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Lane, R., 1986. Important parameters in plasma fractionation. VOX SANG., Vol. 51(SUPPL. 1), pp. 49-51 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Lange, K., 1997. Modern metal forming technology for industrial production. Journal of Materials Processing Technology, Vol. 71(1)Proceedings of the 1996 3rd International Conference on Cold and Warm Forging Technology, October 07,1996 - October 09,1996, pp. 2-13 . Base: Compendex EI; Recuperado: 2009.05.08 11:15:00.	E	-
Langguth, C., Ranaldi, P. & Schuldt, H., 2009. Towards Quality of Service in Scientific Workflows by Using Advance Resource Reservations. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.39, pp. 251-258 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Larson, S. & Morrison, B., 2000. Managing software requirements in the context of the scientific enterprise. , Vol. 4Aerospace Conference Proceedings, 2000 IEEE, pp. 509-522 vol.4-509-522 vol.4 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Learmonth, R., 1997. Biological methods: A novel course in undergraduate biology. FASEB J., Vol. 11(9), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lederman, N., 1995. Suchting on the nature of scientific thought: Are we anchoring curricula in quicksand?. Sci Educ, Vol. 4(4), pp. 371-377 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
LeRouge, C. & Lisetti, C., 2006. Triangulating design science, behavioural science and practice for technological advancement in tele-home health. Int. J. Healthc. Technol. Manage., Vol. 7(5), pp. 348-363 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Levinson, R., Robinson, P. & Thompson, D., 1993. Integrated perception, planning and control for autonomous soil analysis. Artificial Intelligence for Applications, 1993. Proceedings., Ninth Conference on, pp. 249-255 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Li, H., Xiao, J.-Y., Gao, M.-T., Ye, L. & Dai, X., 2008. Preventive effects of soybean phospholipid on liver fibrosis in rats. J. Clin. Rehab. Tissue Eng. Res., Vol. 12(20), pp. 3914-3917 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Li, J. & Wang, J., 2008. Real-time computerized annotation of pictures. IEEE Trans Pattern Anal Mach Intell, Vol. 30(6), pp. 985-1002 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Li, J. & Wang, J., 2006. Real-time computerized annotation of pictures. 14th Annual ACM International Conference on Multimedia, MM 2006, pp. 911-920 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Li, N. & Hu, J., 2009. Scientific workflow model for provenance analysis based on many-sorted high-level nets. 8th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2009, pp. 512-515 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Li, X., Di, L., Han, W., Zhao, P. & Dadi, U., 2009. Sharing and reuse of service-based geospatial processing through a Web Processing Service. Geoinformatics, 2009 17th International Conference on DOI - 10.1109/GEOINFORMATICS.2009.5293431, pp. 1-5 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Ligetiné, A., Mihály, O. & Vajó, P., 1976. Planification a long terme du systeme pedagogique Hongrois. Int Rev Educ, Vol. 22(4), pp. 465-478 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Likharev, K., 1999. Single-electron devices and their applications. Proceedings of the IEEE DOI - 10.1109/5.752518, Vol. 87(4)Proceedings of the IEEE, pp. 606-632 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Lim, C., Lu, S., Chebotko, A. & Fotouhi, F., 2010. Prospective and Retrospective Provenance Collection in Scientific Workflow Environments. Services Computing (SCC), 2010 IEEE International Conference on DOI - 10.1109/SCC.2010.18, pp. 449-456 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Lin, C., Lu, S., Fei, X., Chebotko, A., Pai, D., Lai, Z., Fotouhi, F. & Hua, J., 2009. A Reference Architecture for Scientific Workflow Management Systems and the VIEW SOA Solution. Services Computing, IEEE Transactions on DOI - 10.1109/TSC.2009.4, Vol. 2(1)Services Computing, IEEE Transactions on, pp. 79-92 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Liqing, L., Zhensheng, Y. & Yiqian, L., 1986. SCIENTIFIC EXPERIMENTS OF A 176m STEEL SLANT-LEGGED RIGID FRAME RAILWAY BRIDGE.. Proc of the Pac Struct Steel Conf, pp. 71-80 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lisetti, C. & LeRouge, C., 2004. Affective computing in tele-home health. , Vol. 37Proceedings of the Hawaii International Conference on System Sciences, pp. 2317-2324 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Litkowski, K., 1997. Category development based on semantic principles. Soc. Sci. Comput. Rev., Vol. 15(4), pp. 394-409 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Litvak, A., 2004. High-power millimeter waves applications in plasmas and industrial technologies. Infrared and Millimeter Waves, 2004 and 12th International Conference on Terahertz Electronics, 2004. Conference Digest of the 2004 Joint 29th International Conference on DOI - 10.1109/ICIMW.2004.1422350, pp. 827-828 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	E
Litvak, A., 2002. Application of gyrotrons in research and industrial technologies. Infrared and Millimeter Waves, 2002. Conference Digest. Twenty Seventh International Conference on DOI - 10.1109/ICIMW.2002.1076048, pp. P15- . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	E
Liu, C., Zhang, W., Luo, Z., Lu, F. & Liu, H., 2009. BPEL-based workflow management and parallel job scheduling in Ensemble Prediction. 8th International Conference on Grid and Cooperative Computing, GCC 2009, pp. 409-414 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Liu, D., Abdulla, G., Franklin, M., Garlick, J. & Miller, M., 2006. Data-Preservation in Scientific Workflow Middleware. Scientific and Statistical Database Management, 2006. 18th International Conference on DOI - 10.1109/SSDBM.2006.18, pp. 49-58 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	E
Liu, M. & Fang, K., 2006. A CASE STUDY IN THE APPLICATION OF SUPERSATURATED DESIGNS TO COMPUTER EXPERIMENTS* * Research supported by the National Natural Science Foundation of China (10301015), the Science and Technology Innovation Fund of Nankai University, the Visiting Scholar Program at Chern Institute of Mathematics and a Hong Kong Research Grants Council Grant (RGC/HKBU 200804). Acta Math. Sci., Vol. 26(4), pp. 595-602 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Liu, M., He, H., Sun, X. & Yu, G., 2009. Scientific workflow approach (Kepler) for carbon flux data processing. , Vol. 12009 2nd International Conference on Intelligent Computing Technology and Automation, ICICTA 2009, pp. 694-697 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Liu, W., Duarte, V., Cunha, J. & Luo, T., 2008. Towards a framework for monitoring Grid scientific workflows. 2008 IEEE International Conference on Networking, Sensing and Control, ICNSC, pp. 1211-1216 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Liu, X., Chen, J., Wu, Z., Ni, Z., Yuan, D. & Yang, Y., 2010. Handling Recoverable Temporal Violations in Scientific Workflow Systems: A Workflow Rescheduling Based Strategy. Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference on DOI - 10.1109/CCGRID.2010.15, pp. 534-537 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Louisy, F., 1996. PHYSIOLAB project for CASSIOPEE: the space cardiovascular exploratory laboratory. Trav Sci Cherc Serv Sante Armees, Vol. 17Projet PHYSIOLAB pour CASSIOPEE: le laboratoire d'exploration fonctionnelle cardiovasculaire de l'espace., pp. 281-282 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lu, J., Wang, X.-F., Adjei, O. & Hussain, F., 2004. Sequential patterns graph and its construction algorithm. Jisuanji Xuebao, Vol. 27(6), pp. 782-788 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lu, S. & El-Sharkawi, M., 2006. NEPTUNE Power System: Detection and Location of Switch Malfunctions and High Impedance Faults. , Vol. 3Industrial Electronics, 2006 IEEE International Symposium on, pp. 1960-1965 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Lu, S., El-Sharkawi, M., Kirkham, H. & Howe, B., 2006. NEPTUNE power system: startup power supply for 10 kV to 400 V Dc-Dc converters. Applied Power Electronics Conference and Exposition, 2006. APEC '06. Twenty-First Annual IEEE DOI - 10.1109/APEC.2006.1620720, pp. 5 pp.- . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	E
Lu, S. & Zhang, J., 2009. Collaborative scientific workflows. 2009 IEEE International Conference on Web Services, ICWS 2009, pp. 527-534 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Lu, X., Wang, J., Mitra, P. & Giles, C., 2007. Automatic Extraction of Data from 2-D Plots in Documents. , Vol. 1Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on DOI - 10.1109/ICDAR.2007.4378701, pp. 188-192 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Lu, Y., Cohen, I., Zhou, X. & Tian, Q., 2007. Feature selection using principal feature analysis. 15th ACM International Conference on Multimedia, MM'07, pp. 301-304 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Luckenbill, L., Hintze, K., Ramakrishna, B. & Pizziconi, V., 1999. Interactive nano-visualization in science and engineering education: Conforming technology to transform education: The IN-VSEE project. Interact. Multimedia Electron. J. Comp. Enhanced Learn., Vol. 1, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Ludascher, B., Bowers, S., McPhillips, T. & Podhorszki, N., 2006. Scientific Workflows: More e-Science Mileage from Cyberinfrastructure. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on, pp. 145-145 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-

Ludwischer, B., 2009. What makes scientific workflows scientific?. Lect. Notes Comput. Sci., Vol. 5566 LNCS21st International Conference on Scientific and Statistical Database Management, SSDBM 2009, pp. 217- . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Lumme, M., Valtonen, E., Eronen, T., Torsti, J., Lewitowicz, M. & Bazin, D., 1995. Solar energetic particle abundance measurement by ERNE on board Soho. Adv. Space Res., Vol. 15(7), pp. 73-76 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Luo, L., Zhang, R., Zhang, C., Fan, X. & Li, L., 2008. Workflow Technology for Drug Discovery. Grid and Pervasive Computing Workshops, 2008. GPC Workshops '08. The 3rd International Conference on DOI - 10.1109/GPC.WORKSHOPS.2008.40, pp. 69-74 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Luo, Z., Cui, C.-Z., Nan, K. & Yan, B.-P., 2005. Chemical evolving research of the galaxy under grid environment. Ruan Jian Xue Bao, Vol. 16(8), pp. 1465-1473 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Lushbough, C., Bergman, M., Lawrence, C., Jennewein, D. & Brendel, V., 2010. BioExtract Server: An Integrated Workflow-Enabling System to Access and Analyze Heterogeneous, Distributed Biomolecular Data. Computational Biology and Bioinformatics, IEEE/ACM Transactions on DOI - 10.1109/TCBB.2008.98, Vol. 7(1)Computational Biology and Bioinformatics, IEEE/ACM Transactions on, pp. 12-24 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Ma, J., Zhu, Z., Lu, L., Liu, W. & Wu, D., 2010. Advancement of scientific knowledge output management in Chinese Academy of Sciences based on institutional repositories network. , Vol. 3Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on DOI - 10.1109/ICCSIT.2010.5564673, pp. 502-506 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	--
Ma, Y. & Gong, B., 2009. Double-layer Scheduling Strategy of Load Balancing in Scientific Workflow. Parallel and Distributed Systems (ICPADS), 2009 15th International Conference on DOI - 10.1109/ICPADS.2009.27, pp. 671-678 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Ma, Y., Gong, B. & Zou, L., 2009. Marginal Pricing Based Scheduling Strategy of Scientific Workflow Using Cost-Gradient Metric. Grid and Cooperative Computing, 2009. GCC '09. Eighth International Conference on DOI - 10.1109/GCC.2009.9, pp. 136-143 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Machado, K., Schroeder, E., Ruiz, D. & De Souza, O., 2007. Automating molecular docking with explicit receptor flexibility using scientific workflows. Lect. Notes Comput. Sci., Vol. 4643 LNBI2nd Brazilian Symposium on Bioinformatics, BSB 2007, pp. 1-11 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
MacLaren, J., Allen, G., Dekate, C., Huang, D., Hutanu, A. & Zhang, C., 2005. Shelter from the storm: Building a safe archive in a hostile world. Lect. Notes Comput. Sci., Vol. 3762 LNCSOTM Confederated International Workshops and Posters, AWeSOME, CAMS, GADA, MIOS+INTEROP, ORM, PhDS, SeBGIS, SWWS, and WOSE 2005, pp. 294-303 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Macy, A., 2001. Students as signal sources in the Biomedical Engineering Laboratory. , Vol. 4Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE DOI - 10.1109/IEMBS.2001.1019733, pp. 4040-4043 vol.4 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Madden, E., 2003. Seeing the science for the trees: Employing Daubert standards to assess the Adequacy of National Forest Management under the National Forest Management Act. J. Environ. Law Litig., Vol. 18(2), pp. 321-364 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Maheshwari, K., Goble, C., Missier, P. & Montagnat, J., 2009. Medical image processing workflow support on the EGEE grid with taverna. Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on DOI - 10.1109/CBMS.2009.5255332, pp. 1-7 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Mann, R.W., 1966. Electronics for a Low-Cost Orbiting Vehicle. Aerospace and Electronic Systems, IEEE Transactions on DOI - 10.1109/TAES.1966.4501724, Vol. AES-2(1)Aerospace and Electronic Systems, IEEE Transactions on, pp. 102-114 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Marasco, C., 2005. Technicians craft a promising future. Chem Eng News, Vol. 83(45), pp. 49-50 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Marinho, A., Werner, C., Da Cruz, S., Mattoso, M., Braganholo, V. & Murta, L., 2009. A strategy for provenance gathering in distributed scientific workflows. (PART 1)SERVICES 2009 - 5th 2009 World Congress on Services, pp. 344-347 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Marini, L., Kooper, R., Bajcsy, P. & Myers, J., 2008. Publishing Active Workflows to Problem-Focused Web Spaces. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.151, pp. 444-445 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-

Martínez, M., 2000. Education, the pharmacist and ethical aspects of pharmacy practice. Rev. Mex. Ciencia. Farm., Vol. 31(4)La educación, el farmacéutico y los aspectos éticos de la práctica profesional, pp. 18-22 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Matthies, L., Chen, B. & Petrescu, J., 1997. Stereo vision, residual image processing and Mars rover localization. , Vol. 3Image Processing, 1997. Proceedings., International Conference on DOI - 10.1109/ICIP.1997.632077, pp. 248-251 vol.3 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Mattoso, M., Werner, C., Travassos, G., Braganholo, V., Ogasawara, E., De Oliveira, D., Da Cruz, S., Martinho, W. & Murta, L., 2010. Towards supporting the life cycle of large scale scientific experiments. Int. J. Bus. Process Integr. Manage., Vol. 5(1), pp. 79-92 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Mauser, W., 2003. GLOWA-Danube - Integrative hydrologic modelling approaches to watershed management decision support. Petermanns Geogr. Mitt., Vol. 147(6)GLOWA-Danube: Integrative hydrologische Modellentwicklung zur Entscheidungsunterstützung beim Einzugsgebietsmanagement, pp. 68-75 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
McClatchey, R., Baker, N., Harris, W., Le Goff, J.-M., Kovacs, Z., Estrella, F., Bazan, A. & Le Flour, T., 1997. Version management in a distributed workflow application. Database and Expert Systems Applications, 1997. Proceedings., Eighth International Workshop on DOI - 10.1109/DEXA.1997.617223, pp. 10-15 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
McClatchey, R., Estrella, F., Le Goff, J.-M., Kovacs, Z. & Baker, N., 1997. Object databases in a distributed scientific workflow application. Information Technology, 1997. BIWIT '97., Proceedings of the Third Basque International Workshop on DOI - 10.1109/BIWIT.1997.614047, pp. 11-21 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
McClelland, M., 2004. Emerging standards for interoperable biological systems. Biotechnology and Bioinformatics, 2004. Proceedings. Technology for Life: North Carolina Symposium on, pp. 93-95 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
McCormick, P. & Ahrens, J., 2005. Large-Scale Data Visualization and Rendering: A Problem-Driven Approach. Visualization Handbook, pp. 533-549 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Mellot, A., 2008. Automation and robotics - An essential component of entrepreneurship and innovation in space exploration. , Vol. 1559th International Astronautical Congress 2008, IAC 2008, pp. 9652-9657 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Mercer, S., 2006. Bioinformatics at microsoft research. , Vol. 4316 LNBI1st International Workshop on Data Mining and Bioinformatics, VDMB 2006, pp. 1- . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Meyer, D., 1991. Misinterpretation of interaction effects: A reply to rosnow and rosenthal. Psychol. Bull., Vol. 110(3), pp. 571-573 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Miller, E., 1995. Patterns of knowledge from engineering to the arts, figuring out the relationships. IEEE Potentials, Vol. 14(4), pp. 6-9 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Missier, P., Embury, S., Greenwood, M., Preece, A. & Jin, B., 2007. Managing information quality in e-science: The quator workbench. SIGMOD 2007: ACM SIGMOD International Conference on Management of Data, pp. 1150-1152 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Mo, Z.-N., Qin, M., He, M., Li, M.-J., Yang, X.-L., Xuan, Q. & Mo, L.-J., 2008. Reconstruction and identification of inbred strain mouse somatic cell nuclear transplanted embryos. J. Clin. Rehab. Tissue Eng. Res., Vol. 12(31), pp. 6001-6006 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Money, C., Van Hemmen, J. & Vermeire, T., 2007. Scientific governance and the process for exposure scenario development in REACH. J. Expos. Sci. Environ. Epidemiol., Vol. 17(SUPPL. 1), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Mugaloglu, E. & Saribas, D., 2010. Pre-service science teachers' competence to design an inquiry based lab lesson. Procedia - Social and Behavioral Sciences, Vol. 2(2)Innovation and Creativity in Education, pp. 4255-4259 . Base: Science Direct; Recuperado: 2010.09.29 01:24:08.	E	-
Mulec, K., 2006. Positive professional leaders: Aspects to consider in leadership development. Leadersh. Organ. Dev. J., Vol. 27(1), pp. 66-81 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Murphy, R., Srinivasan, V., Rashidi, N., Duncan, B., Rice, A., Henkel, Z., Garza, M., Nass, C., Groom, V., Zourntos, T., Daneshwar, R. & Prasad, S., 2010. Survivor Buddy and SciGirls: Affect, outreach, and questions. Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on DOI - 10.1109/HRI.2010.5453233, pp. 127-128 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-

Muscat, A. & Allen, E., 1997. A new program in Microelectronics Process Engineering at San Jose State University. University/Government/Industry Microelectronics Symposium, 1997., Proceedings of the Twelfth Biennial, pp. 14-17 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Muscat, A.J. & Allen, E.L., 1997. New program in microelectronics process engineering at San Jose State University. Biennial University/Government/Industry Microelectronics Symposium - Proceedings, pp. 14-17 . Base: Compendex EI; Recuperado: 2009.05.08 11:15:00.	E	-
Myers, J., Fox-Dobbs, C., Laird, J., Le, D., Reich, D. & Curtz, T., 1996. Electronic laboratory notebooks for collaborative research. Enabling Technologies: Infrastructure for Collaborative Enterprises, 1996. Proceedings of the 5th Workshop on DOI - 10.1109/ENABL.1996.555050, pp. 47-51 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Myers, J., McLaren, T. & Wadsworth, A., 2008. Digital Agriculture: Learning to Feed a Hungry World. eScience, 2008. eScience '08. IEEE Fourth International Conference on, pp. 438-439 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Nadeem, F. & Fahringer, T., 2009. Using Templates to Predict Execution Time of Scientific Workflow Applications in the Grid. Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on DOI - 10.1109/CCGRID.2009.77, pp. 316-323 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Nair, M., Desai, A., Kumar, N. & Gopalakrishna, V., 2010. Temperature monitoring using Sun SPOTS applied to vermiculture. Wireless Communication and Sensor Computing, 2010. ICWCSC 2010. International Conference on DOI - 10.1109/ICWCSC.2010.5415907, pp. 1-6 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Nakahara, J., Yamaguchi, E., Nishimori, T., Mochizuki, T., Nakano, M., Furuta, Y., Sekine, S., Ofusa, J. & Yamauchi, Y., 2006. Oyako de Science (Learning about science with parents): Developing an Educational Program for Scientific Experiments Using Cell Phones. Wireless, Mobile and Ubiquitous Technology in Education, 2006. WMUTE '06. Fourth IEEE International Workshop on DOI - 10.1109/WMUTE.2006.261361, pp. 134-136 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Nam, D., Lee, J., Hwang, S., Suh, Y.-K. & Kim, B., 2008. Research Process Support with Organizational Flow in e-Science. Computer and Information Science, 2008. ICIS 08. Seventh IEEE/ACIS International Conference on DOI - 10.1109/ICIS.2008.76, pp. 655-660 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Neumann, A. & Kaluza, J., 1997. A method for spectral analysis of concentrated solar radiation. Renew. Energy, Vol. 10(2-3 SPEC. ISS.), pp. 197-201 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Nichols, B. & Karp, D., 1964. West Ford radar and microwave equipment. Proceedings of the IEEE DOI - 10.1109/PROC.1964.2998, Vol. 52(5)Proceedings of the IEEE, pp. 576-588 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Noble, M. & Drake-Holland, A., 1986. Discrepancies between scientific theory and practice in relation to physiological hypotheses. Theor Med Bioeth, Vol. 7(3), pp. 219-231 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Nock, L., Brunke, S., Jiang, H., Mai, J., Trahey, G. & Von Behren, P., 2002. A new medical ultrasound research interface. , Vol. 2Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE DOI - 10.1109/ULTSYM.2002.1192593, pp. 1571-1573 vol.2 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Norenberg, D., Crewell, S., Lohnert, U., Rose, T. & Martellucci, A., 2008. A Novel Ground-Based Microwave Radiometer for High Precision Atmospheric Observations between 10 and 90 GHz. , Vol. 3Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International DOI - 10.1109/IGARSS.2008.4779500, pp. III - 919-III - 922 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
O'Brien, G. & Peters, J., 1994. Effect of four instructional strategies on integrated science process skill achievement of preservice elementary teachers having different cognitive development levels. J Elem Sci Edu, Vol. 6(1), pp. 30-45 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
O'Connor, M., 2000. Pathways for environmental evaluation: A walk in the (Hanging) Gardens of Babylon. Ecol. Econ., Vol. 34(2), pp. 175-193 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Oinn, T., Addis, M., Ferris, J., Marvin, D., Greenwood, M., Goble, C., Wipat, A., Li, P. & Carver, T., 2004. Delivering web service coordination capability to users. Thirteenth International World Wide Web Conference Proceedings, WWW2004, pp. 1170-1171 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
de Oliveira, D., Ogasawara, E., Baia&x0303;o, F. & Mattoso, M., 2010. SciCumulus: A Lightweight Cloud Middleware to Explore Many Task Computing Paradigm in Scientific Workflows. Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on DOI - 10.1109/CLOUD.2010.64, pp. 378-385 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-

Ostermann, S., Prodan, R. & Fahringer, T., 2009. Extending Grids with cloud resource management for scientific computing. Grid Computing, 2009 10th IEEE/ACM International Conference on DOI - 10.1109/GRID.2009.5353075, pp. 42-49 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Osterweil, L., Wise, A., Clarke, L., Ellison, A., Hadley, J., Boose, E. & Foster, D., 2006. Process technology to facilitate the conduct of science. Lect. Notes Comput. Sci., Vol. 3840 LNCSInternational Software Process Workshop, SPW 2005, pp. 403-415 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Pahwa, J., Brewer, P., Sutton, T., Yesson, C., Burgess, M., Xu, X., Jones, A., White, R., Gray, W., Fiddian, N., Bisby, R., Culham, A., Caithness, N., Scoble, N., Williams, P. & Bhagwat, S., 2006. Biodiversity World: a problem-solving environment for analysing biodiversity patterns. , Vol. 1Cluster Computing and the Grid, 2006. CCGRID 06. Sixth IEEE International Symposium on, pp. 8 pp.-8 pp. . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Paille, C., Albiol, J., Curwy, R., Lasseur, C. & Godia, F., 2000. FEMME: A precursor experiment for the evaluation of bioregenerative life support systems. Planet. Space Sci., Vol. 48(5), pp. 515-521 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Pallickara, S., Plale, B., Jensen, S. & Sun, Y., 2005. Structure, sharing and preservation of scientific experiment data. Challenges of Large Applications in Distributed Environments, 2005. CLADE 2005. Proceedings DOI - 10.1109/CLADE.2005.1520912, pp. 107-114 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Pandey, S., Wu, L., Guru, S. & Buyya, R., 2010. A Particle Swarm Optimization-Based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments. Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on DOI - 10.1109/AINA.2010.31, pp. 400-407 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Paterson, C. & Britten, N., 2008. The patient's experience of holistic care: Insights from acupuncture research. Chronic Illn., Vol. 4(4), pp. 264-277 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Pawlak, R., Horman, R., Stapleton, R. & Headley, R., 1996. Design of a real-time multisensor search and track system. For the detection and tracking of low-flying anti-ship cruise missiles. Nav Eng J, Vol. 108(3), pp. 363-371 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Pejovic, V., Gomez, R., Bojanic, S. & Lalinde, G., 2008. A reconfigurable array based prototype of a specialised string lookup chip. Microelectronics, 2008. MIEL 2008. 26th International Conference on DOI - 10.1109/ICMEL.2008.4559312, pp. 425-428 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Pellegrini, S. & Giacomini, F., 2008. Design of a Petri net-based workflow engine. 3rd International Conference on Grid and Pervasive Computing Symposia/Workshops, GPC 2008, pp. 81-86 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Peppercorn, J., 2006. Ethical issues in phase I cancer clinical trials. Int. J. Pharm. Med., Vol. 20(4), pp. 233-242 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Perera, S. & Gannon, D., 2006. Enabling Web Service extensions for scientific workflows. Workflows in Support of Large-Scale Science, 2006. WORKS '06. Workshop on DOI - 10.1109/WORKS.2006.5282346, pp. 1-10 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Perry, C., 2004. Getting beyond technical rationality in developing health behavior programs with youth. Am. J. Health Behav., Vol. 28(6), pp. 558-568 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Petschel-Held, G., Block, A., Cassel-Gintz, M., Kropp, J., Lüdeke, M., Moldenhauer, O., Reusswig, F. & Schellnhuber, H., 1999. Syndromes of Global Change: A qualitative modelling approach to assist global environmental management. Environ. Model. Assess., Vol. 4(4), pp. 295-314 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Pierrot-Bults, A., 2002. A perspective on the Census of marine life. The role of natural history institutions for this programme. Oceanol. Acta, Vol. 25(5), pp. 187-191 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Podhorszki, N., Klasky, S., Liu, Q., Docan, C., Parashar, M., Abbasi, H., Lofstead, J., Schwan, K., Wolf, M., Zheng, F. & Cummings, J., 2009. Plasma fusion code coupling using scalable I/O services and scientific workflows. 4th Workshop on Workflows in Support of Large-Scale Science, WORKS '09, in Conjunction with SC 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Pond, F. & Pond, J., 2010. Scientific Authority in the Creation-Evolution Debates. Evolution: Education and Outreach, pp. 1-20 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-

Pop, F., Dobre, C. & Cristea, V., 2008. Decentralized dynamic resource allocation for workflows in grid environments. 2008 10th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2008, pp. 557-563 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Pottie, G., 2007. Design of trustworthy fielded sensor networks. Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on, pp. 24-28 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Preobrajensky, V., 1998. Pointed problems of landuse study at the millenium boundary. Izv. Akad. Nauk Ser. Geogr.(3), pp. 14-19 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Prodan, R. & Fahringer, T., 2008. Overhead Analysis of Scientific Workflows in Grid Environments. Parallel and Distributed Systems, IEEE Transactions on, Vol. 19(3), pp. 378-393 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Prodan, R. & Wiczorek, M., 2010. Bi-criteria scheduling of scientific grid workflows. IEEE Trans. Autom. Sci. Eng., Vol. 7(2), pp. 364-376 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Pullin, A., BÃ?ldi, A., Can, O., Dieterich, M., Kat?l, V., Livoreil, B., LÃ?vei, G., MihÃ?k, B., Nevin, O., Selva, N. & Sousa-Pinto, I., 2009. Conservation focus on europe: Major conservation policy issues that need to be informed by conservation science. Conserv. Biol., Vol. 23(4), pp. 818-824 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Rabhi, F., Rana, O., Guabtini, A. & Benatallah, B., 2009. A user-driven environment for financial market data analysis. Lect. Notes Bus. Inf. Process., Vol. 23 LNBIP4th International Workshop on Enterprise Applications and Services in the Finance Industry, FinanceCom 2008, pp. 64-77 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Rajbhandari, S. & Walker, D., 2006. Incorporating Provenance in Service Oriented Architecture. Next Generation Web Services Practices, 2006. NWeSP 2006. International Conference on, pp. 33-40 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	E
Ramakrishnan, A., Singh, G., Zhao, H., Deelman, E., Sakellariou, R., Vahi, K., Blackburn, K., Meyers, D. & Samidi, M., 2007. Scheduling Data-IntensiveWorkflows onto Storage-Constrained Distributed Resources. Cluster Computing and the Grid, 2007. CCGRID 2007. Seventh IEEE International Symposium on DOI - 10.1109/CCGRID.2007.101, pp. 401-409 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Ramakrishnan, L., Guok, C., Jackson, K., Kissel, E., Swany, D.M. & Agarwal, D., 2010. On-demand Overlay Networks for Large Scientific Data Transfers. Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference on DOI - 10.1109/CCGRID.2010.82, pp. 359-367 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Ramakrishnan, L. & Plale, B., 2010. A multi-dimensional classification model for scientific workflow characteristics. 1st International Workshop on Workflow Approaches to New Data-centric Science, Wands '10, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ranaldo, N. & Zimeo, E., 2009. Time and Cost-Driven Scheduling of Data Parallel Tasks in Grid Workflows. Systems Journal, IEEE DOI - 10.1109/JSYST.2008.2011299, Vol. 3(1)Systems Journal, IEEE, pp. 104-120 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Ranaldo, N. & Zimeo, E., 2007. A Time and Cost-Based Matching Strategy for Data Parallelizable Tasks of Grid Workflows. e-Science and Grid Computing, IEEE International Conference on, pp. 295-303 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Raskin, R., 2006. Development of ontologies for earth system science. Spec. Pap. Geol. Soc. Am.(397), pp. 195-199 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ravel, M., 2005. Integrating design and measurement. Instrumentation & Measurement Magazine, IEEE, Vol. 8(1), pp. 41-45 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	E
Re, E., Di Cintio, A., Busca, G., Giunta, D. & Sanchez, M., 2009. Novel Time Synchronization techniques for Deep Space Probes. Frequency Control Symposium, 2009 Joint with the 22nd European Frequency and Time forum. IEEE International DOI - 10.1109/FREQ.2009.5168170, pp. 205-210 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Reddy, M. & Chaudhary, S., 2007. Scheduling in Grid: Rescheduling MPI applications using a fault-tolerant MPI implementation. Communication Systems Software and Middleware, 2007. COMSWARE 2007. 2nd International Conference on DOI - 10.1109/COMSWA.2007.382446, pp. 1-8 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Regrain, R., 1988. Geographical analyses of satellite images of the North and Picardy (France): methods, results and problems. Hommes et Terres du Nord, Vol. 4Analyses geographiques d'images satellitaires du Nord et de la Picardie (France): methodes, resultats et problemes, pp. 238-245 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E

Renon, H., 1978. Qualities of models for evaluation, representation and prediction of fluid phase equilibrium data. Fluid Phase Equilib., Vol. 2(2), pp. 101-118 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Reuss, T., Vossen, G. & Weske, M., 1997. Modeling samples processing in laboratory environments as scientific workflows. Database and Expert Systems Applications, 1997. Proceedings., Eighth International Workshop on, pp. 49-54 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Reysenbach, A.-L. & Flores, G., 2008. Electron microscopy encounters with unusual thermophiles helps direct genomic analysis of Aciduliprofundum boonei. Geobiology, Vol. 6(3), pp. 331-336 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Richards, L., 1980. Cybernetics and the management science process. Omega, Vol. 8(1), pp. 71-80 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Richter, E., 1985. The importance of test conditions in scientific experiments, as demonstrated by the example of tooth mobility. Dtsch Zahnarztl Z, Vol. 40(4)Die Bedeutung der Versuchsbedingungen im wissenschaftlichen Experiment, dargestellt am Beispiel der Zahnbeweglichkeit., pp. 404-409 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Riedel, M., Mallmann, D. & Streit, A., 2005. Enhancing scientific workflows with secure shell functionality in UNICORE grids. , Vol. 20051st International Conference on e-Science and Grid Computing, e-Science 2005, pp. 132-139 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Rodríguez, J., Balch, J. & Rodríguez-Clark, K., 2007. Assessing extinction risk in the absence of species-level data: Quantitative criteria for terrestrial ecosystems. Biodiversity Conserv., Vol. 16(1), pp. 183-209 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Rosa, G., Monico, J. & Chaves, J., 2010. Time series analysis of coordinates estimated with GPS: A methodological approach to detect, remove and recover seasonal effects. Bol. de Cienc. Geod., Vol. 16(1)Análise de séries temporais de coordenadas estimadas com GPS: Uma proposta metodológica para detecção, remoção e recuperação de efeitos sazonais, pp. 51-72 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Rosa, G., Monico, J. & Chaves, J., 2009. Time series analysis of coordinates estimated with GPS: an approach to detect, remove and recover seasonal effects. Bol. de Cienc. Geod., Vol. 15(3)Análise de séries temporais de coordenadas estimadas com gps: Uma proposta metodológica para detecção, remoção e recuperação de efeitos sazonais, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ross, A., 2004. Why bad forecasts can be good for business...[demand forecasting]. Manufacturing Engineer, Vol. 83(3), pp. 26-29 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Ross, R., Peterka, T., Shen, H.-W., Hong, Y., Ma, K.-L., Yu, H. & Moreland, K., 2008. Visualization and parallel I/O at extreme scale. J. Phys. Conf. Ser., Vol. 125, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Ruan, D. & Lu, S., 2010. Task Exception Handling in the VIEW Scientific Workflow System. Services Computing (SCC), 2010 IEEE International Conference on DOI - 10.1109/SCC.2010.77, pp. 637-638 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Rygg, A., Sumitomo, J. & Roe, P., 2006. A Unified Model of Batch and Interactive Scientific Workflow and Its Implementation Using Windows Workflow. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on, pp. 36-36 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Sabbatini, M. & Roesgen, T., 1996. Survivability test of high performance image processing hardware. Eur Space Agency Spec Publ ESA SP(385), pp. 473-477 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Sacchi, C., Rossi, T., Menapace, M. & Granelli, F., 2008. Utilization of UWB Transmission Techniques for Broadband Satellite Connections Operating in W-Band. GLOBECOM Workshops, 2008 IEEE DOI - 10.1109/GLOCOMW.2008.ECP.13, pp. 1-6 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Sadin, S. & Davis, R., 1994. The smallsat revolution ... back to the future?. Acta Astronaut, Vol. 34(C), pp. 109-122 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Saka, A., 2010. Implementation of cooperative learning and guided discussion methods in science teaching to improve professional skills of student teachers. J. Turk. Sci. Educ., Vol. 7(2), pp. 30-51 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Salzer, R., 2008. How we assure quality in teaching analytical chemistry. Anal. Bioanal. Chem., Vol. 390(1), pp. 85-88 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Santer, B., 1985. The use of general circulation models in climate impact analysis - A preliminary study of the impacts of a CO ₂ - induced climatic change on West European agriculture. <i>Clim. Change</i> , Vol. 7(1), pp. 71-93 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Santiago, V., do Amaral, A., Vijaykumar, N., Mattiello-Francisco, M., Martins, E. & Lopes, O., 2006. A Practical Approach for Automated Test Case Generation using Statecharts. , Vol. 2Computer Software and Applications Conference, 2006. COMPSAC '06. 30th Annual International, pp. 183-188 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Santiago, V., Silva, W. & Vijaykumar, N., 2008. Shortening Test Case Execution Time for Embedded Software. <i>Secure System Integration and Reliability Improvement</i> , 2008. SSIRI '08. Second International Conference on, pp. 81-88 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Santiago, V., Vijaykumar, N., Guimaraes, D., Amaral, A. & Ferreira, E., 2008. An Environment for Automated Test Case Generation from Statechart-based and Finite State Machine-based Behavioral Models. <i>Software Testing Verification and Validation Workshop</i> , 2008. ICSTW '08. IEEE International Conference on, pp. 63-72 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	E
Santos, D., Silva, L. & Coelho, J., 2010. Thematic repository as alternative model for science communication. <i>Information Systems and Technologies (CISTI)</i> , 2010 5th Iberian Conference on DOI -, pp. 1-6 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Saribas, D. & Bayram, H., 2009. Is it possible to improve science process skills and attitudes towards chemistry through the development of metacognitive skills embedded within a motivated chemistry lab?: a self-regulated learning approach. <i>Procedia - Social and Behavioral Sciences</i> , Vol. 1(1)World Conference on Educational Sciences, Nicosia, North Cyprus, 4-7 February 2009 - New Trends and Issues in Educational Sciences, pp. 61-72 . Base: Science Direct; Recuperado: 2009.05.08 11:28:00.	E	-
Satava, R., 2005. The scientific method is dead - Long live the (new) scientific method. <i>Surg. Innov.</i> , Vol. 12(2), pp. 173-176 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Scari, E., Ceriotti, M., Colombo, C. & Vasile, M., 2005. Mission analysis of Hevelius - Lunar microsatellite mission. , Vol. 4International Astronautical Federation - 56th International Astronautical Congress 2005, pp. 2117-2126 . Base: Compendex EI; Recuperado: 2009.05.08 11:15:00.	E	-
Schiaffonati, V., 2006. Patrick Suppes' view on models: Thinking about the details of science. <i>Epistemologia</i> , Vol. 29(2)Riflettere sui dettagli della scienza: Patrick Suppes e i modelli, pp. 239-266 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Schooley, L., Zeigler, B., Cellier, F. & Wang, F.-Y., 1993. High-autonomy control of space resource processing plants. <i>Control Systems Magazine</i> , IEEE DOI - 10.1109/37.214942, Vol. 13(3)Control Systems Magazine, IEEE, pp. 29-39 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Schweik, C., Evans, T. & Grove, J., 2005. Open source and open content: A framework for global collaboration in social-ecological research. <i>Ecol. Soc.</i> , Vol. 10(1), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Serra da Cruz, S., da Silva, F., Gadelha, L., Reis Cavalcanti, M., Campos, M. & Mattoso, M., 2008. A Lightweight Middleware Monitor for Distributed Scientific Workflows. <i>Cluster Computing and the Grid</i> , 2008. CCGRID '08. 8th IEEE International Symposium on DOI - 10.1109/CCGRID.2008.89, pp. 693-698 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Shao, Q., Kinsy, M. & Chen, Y., 2007. Storing and Discovering Critical Workflows from Log in Scientific Exploration. <i>Services</i> , 2007 IEEE Congress on, pp. 209-212 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Shaw, J., Blake, S. & Cliatt, M., 1992. Science activities for young children. <i>Early Childhood Educ J</i> , Vol. 20(1), pp. 15-17 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Shen, C., Wu, Z., Shen, F., Luo, J., Lv, B., Zhu, J., Song, C. & Zou, J., 2008. Key Design of Control-Monitoring System of Steady High Magnetic Field Facility. , Vol. 1Computational Intelligence and Industrial Application, 2008. PACIIA '08. Pacific-Asia Workshop on DOI - 10.1109/PACIIA.2008.39, pp. 81-85 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Shi, L., Jones, W., Jensen, R., Harris, S., Perkins, R., Goodsaid, F., Guo, L., Croner, L., Boysen, C., Fang, H., Qian, F., Amur, S., Bao, W., Barbacioru, C., Bertholet, V., Cao, X., Chu, T.-M., Collins, P., Fan, X.-H., Frueh, F., Fuscoe, J., Guo, X., Han, J., Herman, D., Hong, H., Kawasaki, E., Li, Q.-Z., Luo, Y., Ma, Y., Mei, N., Peterson, R., Puri, R., Shippy, R., Su, Z., Sun, Y., Sun, H., Thorn, B., Turpaz, Y., Wang, C., Wang, S., Warrington, J., Willey, J., Wu, J., Xie, Q., Zhang, L., Zhang, L., Zhong, S., Wolfinger, R. & Tong, W., 2008. The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. <i>BMC Bioinform.</i> , Vol. 9(SUPPL. 9), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Shi, X. & Stevens, R., 2008. SWARM: A scientific workflow for supporting bayesian approaches to improve metabolic models. 6th International Workshop on Challenges of Large Applications in Distributed Environments 2008, CLADE'08, pp. 25-34 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Shimizu, Y., Suzuki, T., Suzuki, K., Maeda, T., Kiyokawa, H., Matsuzaki, T. & Mori, K., 1997. Studies on development of fluid drag measurement method with high precision and development of new athlete swimming suit with low fluid drag. , Vol. 5Proceedings of the 1997 ASME Fluids Engineering Division Summer Meeting, FEDSM'97. Part 16 (of 24), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Shumilov, S., Leng, Y., El-Gayyar, M. & Cremers, A., 2008. Distributed scientific workflow management for data-intensive applications. 12th IEEE International Workshop on Future Trends of Distributed Computing Systems, FTDCS 2008, pp. 65-73 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
da Silva, F., Cavalcanti, M. & Davila, A., 2006. In Services: Data Management for In Silico Workflows. Database and Expert Systems Applications, 2006. DEXA '06. 17th International Workshop on DOI - 10.1109/DEXA.2006.72, pp. 206-210 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Siman-Tov, M., Gambill, W., Nelson, W., Ruggles, A. & Yoder, G., 1991. Thermal-hydraulic correlations for the advanced neutron source reactor fuel element design and analysis. , Vol. 190Winter Annual Meeting of the American Society of Mechanical Engineers, pp. 63-78 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Simmhan, Y., Barga, R., van Ingen, C., Lazowska, E. & Szalay, A., 2009. Building the Trident Scientific Workflow Workbench for Data Management in the Cloud. Advanced Engineering Computing and Applications in Sciences, 2009. ADVCOMP '09. Third International Conference on DOI - 10.1109/ADVCOMP.2009.14, pp. 41-50 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Simmhan, Y., Barga, R., van Ingen, C., Lazowska, E. & Szalay, A., 2008. On Building Scientific Workflow Systems for Data Management in the Cloud. eScience, 2008. eScience '08. IEEE Fourth International Conference on, pp. 434-435 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Simmhan, Y., Plale, B. & Gannon, D., 2006. A Framework for Collecting Provenance in Data-Centric Scientific Workflows. Web Services, 2006. ICWS '06. International Conference on DOI - 10.1109/ICWS.2006.5, pp. 427-436 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Sindrilaru, E., Costan, A. & Cristea, V., 2010. Fault Tolerance and Recovery in Grid Workflow Management Systems. Complex, Intelligent and Software Intensive Systems (CISIS), 2010 International Conference on DOI - 10.1109/CISIS.2010.113, pp. 475-480 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Skamp, K., 1987. Preservice teachers: Process skill entry behaviour and opinions about teaching primary science. Res. Sci. Educ., Vol. 17(1), pp. 76-86 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Slota, R., Zieba, J., Kryza, B. & Kitowski, J., 2006. Knowledge Evolution Supporting Automatic Workflow Composition. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on, pp. 37-37 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Smith, A., Nugent, C. & McClean, S., 2001. Intelligent decision support systems for medicine: inherent performance evaluation. , Vol. 4Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE DOI - 10.1109/IEMBS.2001.1019652, pp. 3746-3749 vol.4 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Smith, M., Dasher, H. & Klingborg, D., 2005. A model for recruiting and training youth development volunteers in urban areas. J. Ext., Vol. 43(5), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Soldatova, L., Clare, A., Sparkes, A. & King, R., 2006. An ontology for a Robot Scientist. Bioinformatics, Vol. 22(14), pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Soma, R., Bakshi, A. & Prasanna, V., 2007. An Architecture of a Workflow System for Integrated Asset Management in the Smart Oil Field Domain. Services, 2007 IEEE Congress on, pp. 191-198 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Staicu, A.-M., Crainiceanu, C. & Carroll, R., 2010. Fast methods for spatially correlated multilevel functional data. Biostatistics, Vol. 11(2), pp. 177-194 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Starr, H. & Siverson, R., 1998. Cumulation, evaluation and the research process: Investigating the diffusion of conflict. J. Peace Res., Vol. 35(2), pp. 231-237 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Stevens, R., Zhao, J. & Goble, C., 2007. Using provenance to manage knowledge of In Silico experiments. <i>Brief. Bioinform.</i> , Vol. 8(3), pp. 183-194 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Stoev, A. & Stoeva, P., 2008. Amateur observations of solar eclipses and derivation of scientific data. <i>Adv. Space Res.</i> , Vol. 42(11), pp. 1806-1813 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Strydom, W., Funke, N., Nienaber, S., Nortje, K. & Steyn, M., 2010. Evidence-based policymaking: A review. <i>S. Afr. J. Sci.</i> , Vol. 106(5-6), pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Subramanian, S., Puntervoll, P. & Sztromwasser, P., 2010. Optimizing the Data-Traffic of Centrally Coordinated Scientific Workflow Systems. <i>Web Services (ICWS), 2010 IEEE International Conference on DOI - 10.1109/ICWS.2010.71</i> , pp. 685-688 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Sussman, S., 1991. Curriculum development in school-based prevention research. <i>HEALTH EDUC. RES.</i> , Vol. 6(3), pp. 339-351 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Svakhine, N., Ebert, D. & Stredney, D., 2005. Illustration motifs for effective medical volume illustration. <i>Computer Graphics and Applications, IEEE DOI - 10.1109/MCG.2005.60</i> , Vol. 25(3)Computer Graphics and Applications, IEEE, pp. 31-39 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Svensson, J. & Werner, A., 2007. Large scale Bayesian data analysis for nuclear fusion experiments. <i>2007 IEEE International Symposium on Intelligent Signal Processing, WISP</i> , pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Swaen, G., Carmichael, N. & Doe, J., . Strengthening the reliability and credibility of observational epidemiology studies by creating an Observational Studies Register. <i>J. Clin. Epidemiol.</i> , pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Swap, R. & Privette, J., 1999. Overview of the Southern African Regional Science Initiative-SAFARI 2000. , Vol. 1Geoscience and Remote Sensing Symposium, 1999. <i>IGARSS '99 Proceedings. IEEE 1999 International</i> , pp. 595-597 vol.1-595-597 vol.1 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	E
Tan, W., Missier, P., Foster, I., Madduri, R., De Roure, D. & Goble, C., 2010. A comparison of using Taverna and BPEL in building scientific workflows: The case of caGrid. <i>Concurrency Comput. Pract. Exper.</i> , Vol. 22(9), pp. 1098-1117 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Tan, W., Missier, P., Madduri, R. & Foster, I., 2009. Building scientific workflow with taverna and BPEL: A comparative study in cagrid. <i>Lect. Notes Comput. Sci.</i> , Vol. 5472 <i>LNCSInternational Conference on Service-Oriented Computing, ICSOC 2008</i> , pp. 118-129 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Tan, W., Zhang, J. & Foster, I., 2010. Network Analysis of Scientific Workflows: A Gateway to Reuse. <i>Computer DOI - 10.1109/MC.2010.262</i> , Vol. 43(9)Computer, pp. 54-61 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Targino, R., Cavalcanti, M. & Mattoso, M., 2005. An environment to define and execute in-silico workflows using Web Services. , Vol. 3615 <i>Second International Workshop on Data Integration in the Life Sciences, DILS 2005</i> , pp. 288-291 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Tasoglu, A.K. & Bakaç, M., 2010. The effects of problem based learning and traditional teaching methods on students' academic achievements, conceptual developments and scientific process skills according to their graduated high school types. <i>Procedia - Social and Behavioral Sciences</i> , Vol. 2(2) <i>Innovation and Creativity in Education</i> , pp. 2409-2413 . Base: Science Direct; Recuperado: 2010.09.29 01:24:08.	E	-
Teodoro, G., Tavares, T., Ferreira, R., Kurc, T., Jr., W.M., Guedes, D., Pan, T. & Saltz, J., 2006. A Run-time System for Efficient Execution of Scientific Workflows on Distributed Environments. <i>Computer Architecture and High Performance Computing, 2006. SBAC-PAD '06. 18TH International Symposium on DOI - 10.1109/SBAC-PAD.2006.6</i> , pp. 81-90 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Terpstra, F. & Adriaans, P., 2006. Designing workflow components for e-Science. <i>e-Science 2006 - 2nd IEEE International Conference on e-Science and Grid Computing</i> , pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Terpstra, F., Zhao, Z., Mulder, W. & Adriaans, P., 2007. Towards a formal foundation for aggregating scientific workflows. <i>Lect. Notes Comput. Sci.</i> , Vol. 4489 <i>LNCS(PART 3)7th International Conference on Computational Science, ICCS 2007</i> , pp. 216-219 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Thaler, J. & Battler, M., 2007. In situ peer collaboration improves research and exploration output at Mars analog sites. , Vol. 358th International Astronautical Congress 2007, pp. 1517-1524 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Thelwall, M. & Hasler, L., 2007. Blog search engines. Online Info. Rev., Vol. 31(4), pp. 467-479 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Tripicchio, P., Ruffaldi, E., Avizzano, C. & Bergamasco, M., 2008. Virtual Laboratory: a virtual distributed platform to share and perform experiments. Haptic interfaces for virtual environment and teleoperator systems, 2008. haptics 2008. symposium on DOI - 10.1109/HAPTICS.2008.4479963, pp. 311-318 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Tuot, C., Sintek, M. & Dengel, A., 2008. IVIP - A scientific workflow system to support experts in spatial planning of crop production. Lect. Notes Comput. Sci., Vol. 5069 LNCS20th International Conference on Scientific and Statistical Database Management, SSDBM 2008, pp. 586-591 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Tylianakis, G. & Cotronis, Y., 2009. Data Provenance and Reproducibility in Grid Based Scientific Workflows. Grid and Pervasive Computing Conference, 2009. GPC '09. Workshops at the DOI - 10.1109/GPC.2009.16, pp. 42-49 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Tytler, R. & Peterson, S., 2003. Tracing young children's scientific reasoning. Res. Sci. Educ., Vol. 33(4), pp. 433-465 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Valerio, M., Sahoo, S., Barga, R. & Jackson, J., 2008. Capturing workflow event data for monitoring, performance analysis, and management of scientific workflows. 4th IEEE International Conference on eScience, eScience 2008, pp. 626-633 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Van Der Aalst, W., Bratosin, C., Sidorova, N. & TrÅ?ka, N., 2010. A reference model for grid architectures and its validation. Concurrency Comput. Pract. Exper., Vol. 22(11), pp. 1365-1385 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
Vanden Berg, A.J., 1977. ADVANTAGES OF DISADVANTAGES IN THE APPLICATION OF SAND MOLDINGS.. SME Tech Pap Ser CM for Cast Conf, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Vantassel-Baska, J., 2008. Curriculum development for gifted learners in science at the primary level. Rev. Esp. Pedagog., Vol. 66(240), pp. 283-296 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Varde, A., Ma, S., Maniruzzaman, M., Brown, D., Rundensteiner, E. & Sisson Jr., R., 2008. Comparing mathematical and heuristic approaches for scientific data analysis. Artif Intell Eng Des Anal Manuf, Vol. 22(1), pp. 53-69 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Vessot, R.F., 1991. Applications of highly stable oscillators to scientific measurements. Proc IEEE, Vol. 79(7), pp. 1040-1053 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Vigder, M., Vinson, N., Singer, J., Stewart, D. & Mews, K., 2008. Supporting Scientists' Everyday Work: Automating Scientific Workflows. Software, IEEE DOI - 10.1109/MS.2008.97, Vol. 25(4)Software, IEEE, pp. 52-58 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Vittorini, P. & De Orto, F., 2006. Data management in medicine: The EPIweb information system, a case study and some open issues. Lect. Notes Comput. Sci., Vol. 4254 LNCS10th International Conference on Extending Database Technology, EDBT 2006, pp. 423-432 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Vouk, M., 2003. Integration of heterogeneous scientific data using workflows - a case study in bioinformatics. Information Technology Interfaces, 2003. ITI 2003. Proceedings of the 25th International Conference on, pp. 25-28 . Base: IEEE Explore; Recuperado: 2009.05.08 11:07:00.	E	-
Vukmirovic, S., Erdeljan, A., Imre, L. & Nedic, N., 2010. Hierarchical neural model for workflow scheduling in Utility Management Systems. Soft Computing Applications (SOFA), 2010 4th International Workshop on DOI - 10.1109/SOFA.2010.5565626, pp. 51-56 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
VÅ?hÅ?kÅ?lÅ?, T., 2009. High quality factor integrated passive devices for microwave and millimeter wave applications. , Vol. 7291Nanosensors, Biosensors, and Info-Tech Sensors and Systems 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Wang, D. & Gong, B., 2009. Manageable Provenance Architecture in Scientific Workflow Systems. Grid and Cooperative Computing, 2009. GCC '09. Eighth International Conference on DOI - 10.1109/GCC.2009.17, pp. 131-135 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	E
Wang, F., Deng, H., Dai, W., Zheng, S. & Liang, B., 2010. The Security Issues of Scientific-Workflow Technique in Network Grid Computing. Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on DOI - 10.1109/IITSI.2010.12, pp. 285-288 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-

Wang, F., Deng, H., Ji, K., Liang, B., Deng, Y. & Ren, X., 2009. A study on scientific workflow-based astronomical data dissemination system. ICIC Express Lett., Vol. 3(4), pp. 903-908 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Wang, F., Luo, J., Deng, H., Liang, B. & Ji, K., 2009. C-SWF: A Lightweight Scientific Workflow System for Astronomical Data Processing. , Vol. 2Computer Science and Engineering, 2009. WCSE '09. Second International Workshop on DOI - 10.1109/WCSE.2009.767, pp. 64-67 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Wang, F., Rabsch, C., Kling, P., Liu, P. & Pearson, J., 2007. Web-based Collaborative Information Integration for Scientific Research. Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on DOI - 10.1109/ICDE.2007.368982, pp. 1232-1241 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Wang, J., Altintas, I., Berkley, C., Gilbert, L. & Jones, M., 2008. A High-Level Distributed Execution Framework for Scientific Workflows. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.166, pp. 634-639 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Wang, J., Altintas, I., Hosseini, P., Barseghian, D., Crawl, D., Berkley, C. & Jones, M., 2009. Accelerating Parameter Sweep Workflows by Utilizing Ad-hoc Network Computing Resources: An Ecological Example. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.9, pp. 267-274 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Wang, J., Korambath, P., Kim, S., Johnson, S., Jin, K., Crawl, D., Altintas, I., Smallen, S., Labate, B. & Houk, K.N., 2010. Theoretical enzyme design using the Kepler scientific workflows on the Grid. Procedia Computer Science, Vol. 1(1)ICCS 2010, pp. 1169-1178 . Base: Science Direct; Recuperado: 2010.09.29 01:24:08.	E	-
Wang, K., Tian, H., Wang, Y. & Guo, M., 2008. Fibronectin expression during mouse kidney development. J. Clin. Rehab. Tissue Eng. Res., Vol. 12(11), pp. 2087-2091 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Wang, Y., Cao, J. & Li, M., 2009. Change sequence mining in context-aware scientific workflow. 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2009, pp. 635-640 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Wang, Z. & Tyo, J., 2006. Data Interpretation for Spectral Sensors with Arbitrary Spectral Responses. Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on DOI - 10.1109/IGARSS.2006.353, pp. 1366-1369 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Wassink, I., van der Vet, P., Wolstencroft, K., Neerinx, P., Roos, M., Rauwerda, H. & Breit, T., 2009. Analysing Scientific Workflows: Why Workflows Not Only Connect Web Services. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.48, pp. 314-321 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Wattering, C., Nguyen, D., Fornaro, P., Guggisberg, M., Gyalog, T. & Burkhart, H., 2006. Problem-Based Learning Using Mobile Devices. Advanced Learning Technologies, 2006. Sixth International Conference on DOI - 10.1109/ICALT.2006.1652571, pp. 835-839 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Weber, B., Adair, A. & McGreal, S., 2008. Solutions to the five key brownfield valuation problems. J. Prop. Invest. Finan., Vol. 26(1), pp. 8-37 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Wei, Z., Liu, F., Wei, A. & Cui, X., 2009. Fencing training decision support system based on Bayesian network. 2009 International Conference on Computational Intelligence and Software Engineering, CiSE 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Wellman, N., Scarbrough, F., Ziegler, R. & Lyle, B., 1999. Do we facilitate the scientific process and the development of dietary guidance when findings from single studies are publicized? An American Society for Nutritional Sciences Controversy Session Report. Am. J. Clin. Nutr., Vol. 70(5), pp. 802-805 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	E
Wendelberger, J., Moore, L. & Hamada, M., 2009. Making tradeoffs in designing scientific experiments: A case study with multi-level factors. Qual Eng, Vol. 21(2), pp. 143-155 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Wieczorek, M., Podlipnig, S., Prodan, R. & Fahringer, T., 2008. Bi-criteria scheduling of scientific workflows for the grid. CCGRID 2008 - 8th IEEE International Symposium on Cluster Computing and the Grid, pp. 9-16 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Wilson, M., 1990. Analysis vs. presentation in data visualization software. Compcn Spring '90. Intellectual Leverage. Digest of Papers. Thirty-Fifth IEEE Computer Society International Conference. DOI - 10.1109/CMPCON.1990.63714, pp. 396- . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Winer, G. & Cottrell, J., 1996. Effects of Drawing on Directional Representations of the Process of Vision. J. Educ. Psychol., Vol. 88(4), pp. 704-714 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Wombacher, A., 2008. Composable Data Processing in Environmental Science - A Process View. , Vol. 2Services Computing, 2008. SCC '08. IEEE International Conference on, pp. 499-502 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Wood, B., 2009. Calculating economic values for Turkeys using a deterministic production model. Can. J. Anim. Sci., Vol. 89(2)Calculer les valeurs Économiques de la dinde en utilisant un model de production approprié, pp. 201-213 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Woollard, D., Medvidovic, N., Gil, Y. & Mattmann, C., 2008. Scientific software as workflows: From discovery to distribution. IEEE Software, Vol. 25(4), pp. 37-43 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Wootten, I., Rana, O. & Rajbhandari, S., 2006. Recording actor state in scientific workflows. Lect. Notes Comput. Sci., Vol. 4145 LNCSInternational Provenance and Annotation Workshop, IPAW 2006, pp. 109-117 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Xi, X.-C., Poo, A.-N. & Chou, S.-K., 2007. Support vector regression model predictive control on a HVAC plant. Control Eng. Pract., Vol. 15(8), pp. 897-908 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Xiao, H., Ding, H., Pan, Y., Peng, T. & Li, L., 2008. Study on a highly-stabilized pulse power supply for high magnetic field. Industrial Technology, 2008. ICIT 2008. IEEE International Conference on DOI - 10.1109/ICIT.2008.4608315, pp. 1-5 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Xiao, H., Li, L., Ding, H., Peng, T. & Pan, Y., 2010. Study on a Highly-Stabilized Pulsed Power Supply for High Magnetic Fields. Power Electronics, IEEE Transactions on DOI - 10.1109/TPEL.2009.2039954, Vol. PP(99)Power Electronics, IEEE Transactions on, pp. 1-1 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:54.	E	-
Xie, J., Du, L., Xia, T., Wang, M.-Y., Diao, X. & Li, Y.-P., 2008. Analysis of 1861 wounded inpatients and deaths in West China Hospital of Sichuan University after the Wenchuan earthquake. Chin. J. Evid.-Based Med., Vol. 8(8), pp. 591-596 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Xiuquan, W., Xiaoliu, S. & Zhongfu, T., 2009. Building a web-based decision support system for sustainable development of energy, economy and environment. 2009 International Conference on Signal Processing Systems, ICSPS 2009, pp. 81-84 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
Xu, K., Wang, Y. & Wu, C., 2006. Service Provenance based Abstraction of Grid Application Knowledge. Semantics, Knowledge and Grid, 2006. SKG '06. Second International Conference on DOI - 10.1109/SKG.2006.94, pp. 50-50 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:59:12.	E	-
Yan, Y. & Chapman, B., 2007. Scientific workflow scheduling in computational grids - Planning, reservation, and data/network-awareness. 8th IEEE/ACM International Conference on Grid Computing, GRID 2007, pp. 18-25 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Yang, C.-T., Huang, C.-J. & Hsiao, T.-C., 2008. A Data Grid File Replication Maintenance Strategy Using Bayesian Networks. , Vol. 1Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on DOI - 10.1109/ISDA.2008.357, pp. 456-461 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:58:23.	E	-
Yang, C.-T., Tsai, W.-C., Chen, T.-T. & Hsu, C.-H., 2007. A One-Way File Replica Consistency Model in Data Grids. Asia-Pacific Service Computing Conference, The 2nd IEEE, pp. 364-373 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Yang, H. & Parthasarathy, S., 2006. Mining Spatial and Spatio-Temporal Patterns in Scientific Data. Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on, pp. x146-x146 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Yang, Y., Liu, K., Chen, J., Lignier, J. & Jin, H., 2007. Peer-to-peer based grid workflow runtime environment of SwinDeW-G. E-Science 2007, 3rd IEEE International Conference on E-Science and Grid Computing, pp. 51-58 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Yang, Z., Lu, S. & Yang, P., 2007. Itinerary-Based Access Control for Mobile Tasks in Scientific Workflows. , Vol. 2Advanced Information Networking and Applications Workshops, 2007, AINAW '07. 21st International Conference on, pp. 506-511 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Yildiz, U., Guabtini, A. & Ngu, A., 2009. Business versus Scientific Workflows: A Comparative Study. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.60, pp. 340-343 . Base: IEEE Xplore; Recuperado: 2010.09.29 12:57:27.	E	-
Yu, J., Buyya, R. & Tham, C.K., 2005. Cost-based scheduling of scientific workflow applications on utility grids. e-Science and Grid Computing, 2005. First International Conference on, pp. 8 pp.-147-8 pp.-147 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-

Yu, S., Li, X.-M. & Guo, M., 2007. Expression of angiotensin II subtype 2 receptor in early renal development of mice. <i>J. Clin. Rehab. Tissue Eng. Res.</i> , Vol. 11(8), pp. 1481-1484 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Yuan, D., Yang, Y., Liu, X. & Chen, J., 2010. A cost-effective strategy for intermediate data storage in scientific cloud workflow systems. <i>Parallel & Distributed Processing (IPDPS)</i> , 2010 IEEE International Symposium on DOI - 10.1109/IPDPS.2010.5470453, pp. 1-12 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Zamboulis, L., Martin, N. & Poulouvassilis, A., 2007. Bioinformatics service reconciliation by heterogeneous schema transformation. , Vol. 4544 LNBI4th International Workshop on Data Integration in the Life Sciences, DILS 2007, pp. 89-104 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Zhang, J., 2006. Tracking Dynamics of Geospatial Phenomena in Distributed and Heterogeneous Environments Using Scientific Workflow and Web Services Technologies. <i>Grid and Cooperative Computing, 2006. GCC 2006. Fifth International Conference DOI - 10.1109/GCC.2006.95</i> , pp. 474-481 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Zhang, J., Altintas, I., Tao, J., Liu, X., Pennington, D. & Michener, W., 2006. Integrating Data Grid and Web Services for E-Science Applications: A Case Study of Exploring Species Distributions. <i>e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on DOI - 10.1109/E-SCIENCE.2006.261115</i> , pp. 31-31 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Zhang, J., Pennington, D. & Michener, W., 2007. Performance Evaluations of Geospatial Web Services Composition and Invocation. <i>Web Services, 2007. ICWS 2007. IEEE International Conference on</i> , pp. 1128-1135 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Zhang, J., Pennington, D. & Michener, W., 2005. Validating compositions of geospatial processing Web services in a scientific workflow environment. <i>Web Services, 2005. ICWS 2005. Proceedings. 2005 IEEE International Conference on</i> , pp. 822-822 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Zhang, Q. & Qian, Y.-F., 1999. Monthly mean surface albedo estimated from NCEP/NCAR reanalysis radiation data. <i>Acta Geogr. Sin.</i> , Vol. 54(4), pp. 309-317 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Zhang, W., Cao, J., Zhong, Y., Liu, L. & Wu, C., 2009. Block-Based Concurrent and Storage-Aware Data Streaming for Grid Applications with Lots of Small Files. <i>Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on DOI - 10.1109/CCGRID.2009.26</i> , pp. 538-543 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
qun Zhang, W., ming Wang, W., Cai, Y., dong Wu, Y., jun Zhang, Y., Liu, J., Chen, X. & Yang, H., 2009. Research of the active atomic hydrogen clocks at Shanghai Astronomical Observatory. Piezoelectricity, Acoustic Waves, and Device Applications (SPAWDA) and 2009 China Symposium on Frequency Control Technology, Joint Conference of the 2009 Symposium on DOI - 10.1109/SPAWDA.2009.5428944, pp. 58-58 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Zhang, W.-J., Zhang, G.-H. & Yang, J.-Y., 2007. Generalized extended interpolation method. <i>Fangzhi Gaoxiao Jichukexue Xuebao</i> , Vol. 20(3), pp. 281-285 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Zhang, Y., Mandal, A., Koebel, C. & Cooper, K., 2009. Combined Fault Tolerance and Scheduling Techniques for Workflow Applications on Computational Grids. <i>Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on DOI - 10.1109/CCGRID.2009.59</i> , pp. 244-251 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Zhao, P., Di, L., Yu, G., Yue, P., Wei, Y. & Yang, W., 2009. Semantic Web-based geospatial knowledge transformation. <i>Comput. Geosci.</i> , Vol. 35(4), pp. 798-808 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Zhao, W., Xu, X., Xiong, N. & Wang, Z., 2008. A Weight-Based Dynamic Replica Replacement Strategy in Data Grids. <i>Asia-Pacific Services Computing Conference, 2008. APSCC '08. IEEE DOI - 10.1109/APSCC.2008.41</i> , pp. 1544-1549 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Zhao, Y., Raicu, I. & Foster, I., 2008. Scientific Workflow Systems for 21st Century, New Bottle or New Wine?. <i>Services - Part I, 2008. IEEE Congress on DOI - 10.1109/SERVICES-1.2008.79</i> , pp. 467-471 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Zhao, Z., Belloum, A., De Laat, C., Adriaans, P. & Hertzberger, B., 2007. Using Jade agent framework to prototype an e-Science workflow bus. <i>7th IEEE International Symposium on Cluster Computing and the Grid, CCGrid 2007</i> , pp. 655-660 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

Zhao, Z., Belloum, A., Sloom, P. & Hertzberger, B., 2005. Agent technology and scientific workflow management in an e-science environment. Tools with Artificial Intelligence, 2005. ICTAI 05. 17th IEEE International Conference on DOI - 10.1109/ICTAI.2005.29, pp. 5 pp.-23 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
Zhao, Z., Booms, S., Belloum, A., de Laat, C. & Hertzberger, B., 2006. VLE-WFBus: A Scientific Workflow Bus for Multi e-Science Domains. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on DOI - 10.1109/E-SCIENCE.2006.261095, pp. 11-11 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E	-
Zherlitsyn, S., Herrmannsdoerfer, T., Skourski, Y., Sytcheva, A. & Wosnitza, J., 2006. Pulsed-magnet design at the Dresden High Magnetic Field Laboratory. J. Phys. Conf. Ser., Vol. 51(1), pp. 583-586 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
Zhou, J., Cui, Y., Davis, S., Guest, C.C. & Maechling, P., 2010. Workflow-Based High Performance Data Transfer and Ingestion to Support Petascale Simulations on TeraGrid. , Vol. 1 Computational Science and Optimization (CSO), 2010 Third International Joint Conference on DOI - 10.1109/CSO.2010.235, pp. 343-347 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Ziemer, F. & Dittmer, J., 1994. A system to monitor ocean wave fields. , Vol. 2 OCEANS '94. 'Oceans Engineering for Today's Technology and Tomorrow's Preservation.' Proceedings DOI - 10.1109/OCEANS.1994.364010, pp. II/28-II/31 vol.2 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
Zwoll, K., Muller, K.D., Becks, B., Erven, W. & Sauer, M., 1977. Direct Numerical Control of Machine Tools in a Nuclear Research Center by the CAMAC System. Nuclear Science, IEEE Transactions on DOI - 10.1109/TNS.1977.4328721, Vol. 24(1) Nuclear Science, IEEE Transactions on, pp. 446-451 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
Ā-zdilek, Z. & Ā-zkan, M., 2009. The effect of applying elements of instructional design on teaching material for the subject of classification of matter. Turk. Onl. J. Edu. Tech., Vol. 8(1), pp. 84-93 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
, 1971. AMERICAN ASTRONAUTICAL SOCIETY ANNUAL MEETING, 17th, PROCEEDINGS, 1971: THE OUTER SOLAR SYSTEM.. Adv Astronaut Sci, Vol. 2 Jun 28-30 1971, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 2009. Proceedings of the 2009 ICSE Workshop on Comparison and Versioning of Software Models, CVSM 2009. 2009 ICSE Workshop on Comparison and Versioning of Software Models, CVSM 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
, 2009. 8th International Conference on Grid and Cooperative Computing, GCC 2009. 8th International Conference on Grid and Cooperative Computing, GCC 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
, 2009. SERVICES 2009 - 5th 2009 World Congress on Services. (PART 2) SERVICES 2009 - 5th 2009 World Congress on Services, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
, 2009. OSSC-2009 - Proceedings of 2009 IEEE International Workshop on Open-source Software for Scientific Computation. 2009 IEEE International Workshop on Open-source Software for Scientific Computation, OSSC-2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	-
, 2009. Proceedings - 2009 IEEE Congress on Services, SERVICES 2009. (PART 1) SERVICES 2009 - 5th 2009 World Congress on Services, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E	E
, 2009. Title Page i. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.1, pp. i-i . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
, 2009. Cover Art. Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on DOI - 10.1109/CCGRID.2009.100, pp. C1-C1 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E	-
, 2008. Proceedings - 12th IEEE International Workshop on Future Trends of Distributed Computing Systems, FTDCS 2008. 12th IEEE International Workshop on Future Trends of Distributed Computing Systems, FTDCS 2008, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 2008. Business Process Management - 6th International Conference, BPM 2008, Proceedings. Lect. Notes Comput. Sci., Vol. 5240 LNCS 6th International Conference on Business Process Management, BPM 2008, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 2008. Post Workshop Proceedings - EDBT 2008 Ph.D. Workshop. EDBT 2008 Ph.D. Workshop, held in conjunction with the 11th International Conference on Extending Database Technology, EDBT, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-

, 2008. Call for Papers - Special Issue on Scientific Workflow Management and Applications. Robotics & Automation Magazine, IEEE DOI - 10.1109/M-RA.2007.914991, Vol. 15(1)Robotics & Automation Magazine, IEEE, pp. 121-121 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E	-
, 2007. Third IEEE International Conference on e-Science and Grid Computing - Title. e-Science and Grid Computing, IEEE International Conference on DOI - 10.1109/E-SCIENCE.2007.76, pp. i-iii . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	E	-
, 2006. Proceedings - SCW 2006: IEEE Services Computing Workshops. SCW 2006: IEEE Services Computing Workshops, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 2006. The Semantic Web - ISWC 2006 - 5th International Semantic Web Conference, ISWC 2006, Proceedings. Lect. Notes Comput. Sci., Vol. 4273 LNCS5th International Semantic Web Conference, ISWC 2006, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 2005. Conceptual Modeling - ER 2005 - 24th International Conference on Conceptual Modeling, Proceedings. Lect. Notes Comput. Sci., Vol. 3716 LNCS24th International Conference on Conceptual Modeling - ER 2005, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 1997. Proceedings of the 1997 3rd Basque International Workshop on Information Technology, BIWIT. Proceedings of the 1997 3rd Basque International Workshop on Information Technology, BIWIT, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E	-
, 1995. 1995 IEEE International SOI Conference Proceedings. SOI Conference, 1995. Proceedings., 1995 IEEE International, pp. - . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E	-
Leake, D. & Kendall-Morwick, J., 2008. Towards case-based support for e-Science workflow generation by mining provenance. Lect. Notes Comput. Sci., Vol. 5239 LNAI9th European Conference on Case-Based Reasoning, ECCBR 2008, pp. 269-283 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E (*)	-
Garcia, D. & De Toledo, M., 2008. Quality of service management for web service compositions. 2008 IEEE 11th International Conference on Computational Science and Engineering, CSE 2008, pp. 189-196 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E (*)	-
Martin, F., Baggett, T., Wolfe, T. & Mita, R., 2000. Using sampling theory to build a more universal forest vegetation database. Comput. Electron. Agric., Vol. 28(2)International Conference on Integrated Tools for Natural Resources Inventories in the 21st Century, pp. 119-132 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E (I)	-
Bowers, S., Ludascher, B., Ngu, A. & Critchlow, T., 2006. Enabling ScientificWorkflow Reuse through Structured Composition of Dataflow and Control-Flow. Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on, pp. 70-70 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	E (N)	-
Salayandia, L., Roach, S. & Gates, A., 2008. Program synthesis from workflow-driven ontologies. 2008 Annual Meeting of the North American Fuzzy Information Processing Society, NAFIPS 2008, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E (N)	-
Xiang, X. & Madey, G., 2007. Improving the reuse of scientific workflows and their by-products. 2007 IEEE International Conference on Web Services, ICWS 2007, pp. 792-799 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	E (N)	-
Chinthaka, E., Ekanayake, J., Leake, D. & Plale, B., 2009. CBR based workflow composition assistant. (PART 1)SERVICES 2009 - 5th 2009 World Congress on Services, pp. 352-355 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E**	-
Ludascher, B., Altintas, I. & Gupta, A., 2003. Compiling abstract scientific workflows into Web service workflows. Conference on Scientific and Statistical Database Management, 2003. 15th International DOI - 10.1109/SSDM.2003.1214990, pp. 251-254 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E**	-
McIver, R., Jones, A. & White, R., 2009. A Framework for Supporting the Composition of Biodiversity Informatics Resources. Database and Expert Systems Application, 2009. DEXA '09. 20th International Workshop on DOI - 10.1109/DEXA.2009.88, pp. 350-354 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E**	-
Robinson, A., Rahayu, W. & Dillon, T., 2009. WAD Workflow System: Data-Centric Workflow System. Software Engineering Conference, 2009. ASWEC '09. Australian DOI - 10.1109/ASWEC.2009.26, pp. 337-344 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E**	-
Chase, J., Gorton, I., Sivaramakrishnan, C., Almquist, J., Wynne, A., Chin, G. & Critchlow, T., 2009. Kepler + MeDiCi Service-Oriented Scientific Workflow Applications. Services - I, 2009 World Conference on DOI - 10.1109/SERVICES-I.2009.70, pp. 275-282 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E**	-

Cheng, J., Lin, X., Zhou, Y. & Li, J., 2009. A web based workflow system for distributed atmospheric data processing. 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2009, pp. 584-588 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E**	-
Eckman, B., Gaasterland, T., Lacroix, Z., Raschid, L., Snyder, B. & Vidal, M., 2006. Implementing a Bioinformatics Pipeline (BIP) on a Mediator Platform: Comparing Cost and Quality of Alternate Choices. Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on DOI - 10.1109/ICDEW.2006.66, pp. 67-67 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E**	-
Gosney, A., Oehmen, C., Wynne, A. & Almquist, J., 2010. An adaptive middleware framework for Scientific Computing at extreme scales. Information Reuse and Integration (IRI), 2010 IEEE International Conference on DOI - 10.1109/IRI.2010.5558934, pp. 232-238 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E**	-
Gubala, T., Herezłak, D., Bubak, M. & Malawski, M., 2006. Semantic Composition of Scientific Workflows Based on the Petri Nets Formalism. e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference on DOI - 10.1109/E-SCIENCE.2006.261096, pp. 12-12 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E**	-
Johnson, D., Meacham, K. & Kornmayer, H., 2009. A middleware independent Grid workflow builder for scientific applications. E-Science Workshops, 2009 5th IEEE International Conference on DOI - 10.1109/ESCIV.2009.5407993, pp. 86-91 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E**	-
Laforenza, D., Lombardo, R., Scarpellini, M., Serrano, M., Silvestri, F. & Faccioli, P., 2007. Biological Experiments on the Grid: A Novel Workflow Management Platform. Computer-Based Medical Systems, 2007. CBMS '07. Twentieth IEEE International Symposium on DOI - 10.1109/CBMS.2007.29, pp. 489-494 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E**	-
Lee, J., Nam, D., Hwang, S. & hwan Byeon, O., 2008. A Grid-Enabled Problem Solving Environment for Supporting Collaborative Aerodynamic Engineering Process. eScience, 2008. eScience '08. IEEE Fourth International Conference on DOI - 10.1109/eScience.2008.13, pp. 770-777 . Base: IEEEExplore; Recuperado: 2010.09.29 12:59:12.	E**	-
Sun, D., Li, Z., Zhao, Y. & Ma, D., 2009. Orchestra Designer: an open-source tool for scientific workflow modeling. Open-source Software for Scientific Computation (OSSC), 2009 IEEE International Workshop on DOI - 10.1109/OSSC.2009.5416916, pp. 39-43 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E**	-
Sun, Z. & Yue, P., 2010. The use of Web 2.0 and geoprocessing services to support geoscientific workflows. Geoinformatics, 2010 18th International Conference on DOI - 10.1109/GEONFORMATICS.2010.5567702, pp. 1-5 . Base: IEEEExplore; Recuperado: 2010.09.29 12:57:27.	E**	-
Yang, X., Bruin, R. & Dove, M., 2010. Developing an End-to-End Scientific Workflow:. Computing in Science & Engineering DOI - 10.1109/MCSE.2010.61, Vol. 12(3)Computing in Science & Engineering, pp. 52-61 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:23.	E**	-
LudÅrscher, B., Weske, M., McPhillips, T. & Bowers, S., 2009. Scientific workflows: Business as usual?. Lect. Notes Comput. Sci., Vol. 5701 LNCS7th International Conference on Business Process Management, BPM 2009, pp. 31-47 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E**	-
Yildiz, U., Guabtni, A. & Ngu, A., 2009. Towards scientific workflow patterns. 4th Workshop on Workflows in Support of Large-Scale Science, WORKS '09, in Conjunction with SC 2009, pp. - . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	E**	-
Bosin, A., Dessi, N., Fugini, M., Liberati, D. & Pes, B., 2005. Applying enterprise models to design cooperative scientific environments. Lect. Notes Comput. Sci., Vol. 3812 LNCSBPM 2005 International Workshops, BPI, BPD, ENEI, BPRM, WSCOBPM, BPS, pp. 281-292 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	I	-
Chard, K., Onyuksel, C., Tan, W., Sulakhe, D., Madduri, R. & Foster, I., 2008. Build Grid Enabled Scientific Workflows Using gRAVI and Taverna. eScience, 2008. eScience '08. IEEE Fourth International Conference on, pp. 614-619 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.	I	-
Ellison, A., Osterweil, L., Clarke, L., Hadley, J., Wise, A., Boose, E., Foster, D., Hanson, A., Jensen, D., Kuzeja, P., Riseman, E. & Schultz, H., 2006. Analytic webs support the synthesis of ecological data sets. Ecology, Vol. 87(6), pp. 1345-1358 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.	I	E
Jablonski, S., Volz, B. & Rehman, M., 2007. A conceptual modeling and execution framework for process based scientific applications. ACM 1st Workshop on CyberInfrastructure: Information Management in eScience, CIMS '07, Co-located with the 16th ACM Conference on Information and Knowledge Management, CIKM 2007, pp. 23-30 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.	I	-
Kalyanam, R., Zhao, L., Park, T. & Goasguen, S., 2007. A Web Service-Enabled Distributed Workflow System for Scientific Data Processing. Future Trends of Distributed Computing Systems, 2007. FTDCS '07. 11th IEEE International Workshop on DOI - 10.1109/FTDCS.2007.9, pp. 7-14 . Base: IEEEExplore; Recuperado: 2010.09.29 12:58:54.	I	-

Li, Q., Shan, Z., Hung, P., Chiu, D. & Cheung, S., 2006. Flows and views for scalable scientific process integration. , Vol. 1521st International Conference on Scalable Information Systems, InfoScale '06, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Liu, X., Dou, W., Chen, J., Fan, S., Cheung, S. & Cai, S., 2007. On design, verification, and dynamic modification of the problem-based scientific workflow model. Simul. Model. Pract. Theory, Vol. 15(9), pp. 1068-1088 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Liu, X., Dou, W., Fan, S. & Cai, S., 2006. The problem-based scientific workflow design and performance in Grid environments. 5th International Conference on Grid and Cooperative Computing, GCC 2006 - Workshops, pp. 267-274 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Osterweil, L., Clarke, L., Ellison, A., Boose, E., Podorozhny, R. & Wise, A., 2010. Clear and precise specification of ecological data management processes and dataset provenance. IEEE Trans. Autom. Sci. Eng., Vol. 7(1), pp. 189-195 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.		-
Osterweil, L., Clarke, L., Ellison, A., Podorozhny, R., Wise, A., Boose, E. & Hadley, J., 2008. Experience in using a process language to define scientific workflow and generate dataset provenance. 16th ACM SIGSOFT International Symposium on the Foundations of Software Engineering, SIGSOFT 2008/FSE-16, pp. 319-329 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.		-
Sorde, S., Aggarwal, S., Song, J., Koh, M. & See, S., 2007. Modeling and verifying non-DAG workflows for computational grids. 2007 IEEE Congress on Services, SERVICES 2007, pp. 237-244 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		
Verdi, K., Ellis, H. & Gryk, M., 2007. Conceptual-level workflow modeling of scientific experiments using NMR as a case study. BMC Bioinform., Vol. 8, pp. - . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		
Yang, P., Yang, Z. & Lu, S., 2007. Formal modeling and analysis of scientific workflows using hierarchical state machines. E-Science 2007, 3rd IEEE International Conference on E-Science and Grid Computing, pp. 619-620 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		
Amigoni, F., Fugini, M. & Liberati, D., 2007. A virtual laboratory for web and grid enabled scientific experiments. , Vol. SAIC9th International Conference on Enterprise Information Systems, ICEIS 2007, pp. 227-230 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.		
Boose, E., Ellison, A., Osterweil, L., Clarke, L., Podorozhny, R., Hadley, J., Wise, A. & Foster, D., 2007. Ensuring reliable datasets for environmental models and forecasts. Ecol. Informatics, Vol. 2(3 SPEC. ISS.), pp. 237-247 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		
Gil, Y., Ratnakar, V., Deelman, E., Mehta, G. & Kim, J., 2007. Wings for Pegasus: Creating large-scale scientific applications using semantic representations of computational workflows. , Vol. 2AAAI-07/IAAI-07 Proceedings: 22nd AAAI Conference on Artificial Intelligence and the 19th Innovative Applications of Artificial Intelligence Conference, pp. 1767-1774 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Kaestle, G., Shek, E. & Dao, S., 1999. Sharing experiences from scientific experiments. Scientific and Statistical Database Management, 1999. Eleventh International Conference on, pp. 168-177 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.		-
Kaletas, E., Afsarmanesh, H. & Hertzberger, L., 2004. A methodology for integrating new scientific domains and applications in a virtual laboratory environment. ICEIS 2004 - Proceedings of the Sixth International Conference on Enterprise Information Systems, pp. 265-272 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Kaletas, E., Afsarmanesh, H. & Hertzberger, L., 2003. Modelling multi-disciplinary scientific experiments and information. Lect. Notes Comput. Sci., Vol. 2869, pp. 75-82 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Lacroix, Z., Legendre, C. & Tuzmen, S., 2009. Reasoning on scientific workflows. (PART 1)SERVICES 2009 - 5th 2009 World Congress on Services, pp. 306-313 . Base: Scopus; Recuperado: 2010.09.29 01:21:56.		-
Medeiros, C., Perez-Alcazar, J., Digiampietri, L., Pastorello Jr., G., Santanche, A., Torres, R., Madeira, E. & Bacarin, E., 2005. WOODSS and the Web: Annotating and reusing scientific workflows. SIGMOD Rec., Vol. 34(3), pp. 18-23 . Base: Scopus; Recuperado: 2009.05.08 10:58:00.		-
Troger, A. & Fernandes, A., 2004. A language for comprehensively supporting the in vitro experimental process in silico. Bioinformatics and Bioengineering, 2004. BIBE 2004. Proceedings. Fourth IEEE Symposium on, pp. 47-56 . Base: IEEE Xplore; Recuperado: 2009.05.08 11:07:00.		E

Curcin, V. and Ghanem, M. and Guo, Y., 2010. The design and implementation of a workflow analysis tool. <i>Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.</i> 368, pp. 4193—4208. Base: Scopus; Recuperado: 2010.09.29 10:03:00	E	E
Gil, Y. and Gonzalez-Calero, P.A. and Deelman, E., 2007. On the black art of designing computational workflows. <i>Proc. Workshop Workflows Support Large-scale Sci.</i> 2007, pp. 53—62. Base: -; Recuperado: 2010.09.29 12:58:23.	I	-

ANEXO C – DADOS EXTRAÍDOS PARA ANÁLISE NA QUASI-REVISÃO

Osterweil et al. 2010; Osterweil et al. 2008; Boose et. al. 2007; Ellison et. al. 2006;	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- A confiabilidade dos datasets (proveniência). O artigo argumenta que os recentes esforços no desenvolvimento de ferramentas e sistemas não apresentam um bom suporte a proveniência dos dados gerados (datasets). O que torna fraca a confiabilidade desses, portanto a solução seria capturar não só os dados, mas também o processo experimental que os geraram.
Tecnologia de construção	Não foi possível identificar.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem completa lida com Composição, Execução e Análise. Mas o foco do artigo é a fase de Composição.
Instrumentos	Little-JIL: É uma linguagem visual usada na representação do PDG. De fato, esse é a linguagem de workflow usada na abordagem. Data-Flow Graph (DFG): Este é o grafo que captura o fluxo de atividades do processo experimental. Isto não liga uma atividade a um recurso computacional (como uma ferramenta). É um grafo de alto nível. Dataset-Derivation graph: Este é o grafo que captura os datasets criados e utilizados em uma execução, e quando esses foram usados e criados. Process-derivation graph: Este grafo que também representa um fluxo de atividades, mas esse é mais formal que o DFG. Esse representa agentes (recursos humanos ou computacionais) que uma atividade utiliza. Sistema FLAVERS: Esse é o sistema que verifica os estados finitos. Este é utilizado para verificar o modelo do processo experimental no grafo PDG.
Mecanismos	Analytic web: Esse é o mecanismo que captura o processo experimental e, também, suporta a proveniência dos datasets e resultados gerados. Este é composto por três diferentes grafos: DFG; DDG; e, PDG. SciWalker: Este é o sistema utilizado pelos cientistas para modelar e definir o processo de experimentação. Além disso, este sistema pode executar o fluxo de trabalho criado e pode registrar as atividades e conjunto de dados utilizados na execução.
Artefatos	DFG diagram: Este artefato é gerado na composição e contém os fluxos das atividades do processo de experimentação. PDG diagram: Este artefato contém todos os caminhos de execução do processo experimental. Ele também define os recursos necessários para a execução. Esse diagrama é gerado na execução e é único para o experimento. DDG diagrams: Este é o artefato que contém a rastreabilidade dos conjuntos de dados e das atividades desenvolvidas em uma execução. Este é gerado para cada execução.
Atividades	A abordagem tem três atividades gerais: Composição, execução, e, Análise. Na composição, o DFG e PDG são criados pelo biólogo, usando o SciWalker. Na execução, o PDG é executado pelo SciWalker e o DDG será gerado e armazenado. Por último, o DDG, PDG e DFG pode ser re-executados e analisados.
Tipo de domínio	A abordagem foi aplicada em um domínio – Ecológico. Essa abordagem foi aplicada no domínio de Biologia. O artigo apresenta um caso de aplicação descrito na seção - “AN ANALYTIC WEB FOR ECOSYSTEM CARBON FLUX”.
Número de projetos	Somente um caso é descrito no artigo. Contudo, o artigo menciona outros projetos no qual a abordagem foi aplicada.
Dados para a questão secundária	
Especificação da técnica de verificação	A abordagem consiste na verificação do estado-finito do modelo PDG. Esta tem como foco a análise das seqüências esperadas de atividades que realmente ocorre.
Descrição da técnica de verificação	
Instrumentos	Sistema FLAVERS e técnica de análise: É um sistema para verificar diagramas de estados finitos baseado no diagrama de DPG. Este é usado para verificar o modelo do processo experimental representado como um grafo PDG.

Procedimentos	O procedimento consiste na aplicação de técnicas de verificação, após a criação do PDG, para identificar os possíveis caminhos que nunca irão ocorrer.
Tipo de domínio	Somente é aplicado no modelo PDG.

Bosin et. al. 2007; Amigoni et al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- A disponibilidade de um ambiente virtual cooperativo para a ciência que integre as várias aplicações: Alega-se que há uma falta de apoio à integração de aplicações e ferramentas difundidas entre organizações distintas, grupos ou locais. Estas aplicações são usadas na execução dos experimentos. Isto pode tornar a experimentação mais difícil.
Tecnologia de construção	Paradigma: A abordagem tem um paradigma para a composição dos experimentos científicos que utiliza tarefas cooperativas, o e-processo. Esse é baseado em SOA – Service Oriented Architecture and e-service. O protótipo da abordagem é construído utilizando o conceito de multi-camada .
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição e execução. O artigo centra-se na composição. A abordagem lida com o conceito de níveis de abstração.
Instrumentos	Experiment Master Pool (EMP): Este é um tipo de e-nó, que funciona como um controlador dentro do ambiente cooperativo. Linguagem XPDL: Ele será o idioma do <i>workflow</i> que será executado em uma <i>engine</i> baseada no WfMC e OMG. Esta <i>engine</i> não está implementada (na época do artigo).
Mecanismos	O Cooperative Framework: É um <i>framework</i> que é utilizada pelo cientista para compor os experimentos. Então, eles são capazes de explorar os vários serviços (e-nodes) disponíveis no pool de serviços, entre os diferentes grupos de pesquisa. Se for necessário, a colaboração entre os vários serviços espalhados por diferentes e-nós, eles podem usar o <i>Smart Scientific Space</i> que permite a interação entre os <i>Pool of Services</i> .
Artefatos	Workflow que representa o experimento cooperativo: Este artefato representa o experimento cooperativo que está definido no <i>Cooperative framework</i> . No futuro, será definido em linguagem XPDL. Um experimento cooperativo: Este é o principal artefato gerado pela abordagem e são definidos em forma de serviços necessários, os papéis envolvidos e os <i>workflows</i> que contém a ordem de execução.
Atividades	Modelagem de experimentos científicos: A abordagem geral para a composição é composta por 3 tarefas. A primeira, o cientista-chefe define o experimento a partir da hipótese. Depois disso, ele decompõe os experimentos em uma série de tarefas (serviços), e o cientista-chefe e o experiente usuário gerenciador negociam o uso dos recursos. Por último, o <i>workflow</i> , que é criado pela composição destes serviços, está disponível para uso.
Tipo de domínio	O artigo não especifica qualquer restrição sobre os tipos de domínio. Mas, a abordagem é baseada no conceito de serviço.
Número de projetos	O artigo não especifica o número de projetos em que a abordagem foi aplicada. Apenas um foi descrito e é usado como exemplo também.

Chard et. al. 2008	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Complexidade do desenvolvimento de workflows: Os workflows baseados em serviços Web requerem competências técnicas elevadas dos desenvolvedores (pesquisadores). A abordagem propõe uma solução para reduzir a complexidade desta tarefa.
Tecnologia de construção	Paradigma: Tem um paradigma para a composição de experimentos científicos que usa Service-Oriented (que é baseada em arquitetura orientada a serviços - SOA). A abordagem utiliza a integração de componentes para representar a solução.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição e execução. Mas, o artigo centra-se na composição. A abordagem centra-se, especialmente, na concepção do nível concreto de <i>workflows</i> .
Instrumentos	Linguagem Sculf: Ela é usada como linguagem de definição para os <i>workflows</i> na Taverna. XML Splitter: Processador do Taverna que lida com entradas e saídas complexas. Java bean shell: Processador do Taverna utilizado para interpretar o código Java e manipular insumos, criar laços, e outras funções. WSRF: É uma abordagem para o modelo de serviços web que será utilizado na composição do fluxo de trabalho concreto - "Web Service Resource Framework".
Mecanismos	Introduce: É um conjunto de ferramentas para desenvolvimento e implementação de Web

	Services. GRAVI: É um plug-in para Introduce, ele permite que “embrulhe” (wrapping) a aplicação. Taverna: É a <i>engine</i> de execução do <i>workflow</i> concreto. Globus Toolkit 4: É uma Implementação da especificação WSRF e é utilizada como recipiente para os serviços criados para o <i>workflow</i> .
Artefatos	Serviços Web criados no GRAVI plug-in: Cada serviço web representa um serviço que é usado por um componente processador no Taverna. Workflow no nível concreto: É definido na <i>workbench</i> do Taverna, estes fluxos são definidos em linguagem Sculf.
Atividades	Existem três atividades gerais. Em primeiro lugar, os serviços são criados em GRAVI, onde os pedidos são embalados em Web Services. Em segundo lugar, o <i>workflow</i> concreto é gerado no Taverna através da inserção dos serviços já definidos. Em terceiro lugar, a execução é realizada na <i>engine</i> de execução do Taverna.
Tipo de domínio	O artigo não especifica qualquer restrição sobre os tipos de domínio. Mas, a abordagem é baseada na arquitetura orientada a serviços e o exemplo de uso é de domínio da bioinformática.
Número de projetos	O número de projetos, onde a abordagem foi aplicada, não é especificado. O artigo apresenta apenas um exemplo do “fluxo de trabalho Transposon”.

Gil et. al. 2007; Gil et al., 2007b	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Escalabilidade da concepção de workflows: O artigo argumenta que os workflows estão aumentando em complexidade e isto torna essencial que novas abordagens para lidar com esse problema surjam. Isto é justificado pelo fato desses workflows serem difíceis de serem mantidos e modificados.
Tecnologia de construção	A abordagem usa a integração de componentes para representar a solução. A abordagem usa o conceito de descrição semântica para auxiliary a composição dos workflows. Isto usa ontologias para descrição dos elementos <i>File, Component, Nodes and Links</i> .
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição e execução . Mas, o artigo enfoca a composição.
Instrumentos	Composition Analysis Tool (CAT): Ferramenta que auxilia os cientistas na tarefa de compor o modelo de <i>workflow</i> científico. É baseado em definições apresentadas em ontologias. OWL-DL and Jena: Esta é a linguagem em que o modelo de <i>workflow</i> é especificado. Contém as ontologias para os seguintes itens: Arquivo; Component; Nó e, Link. Directed Acyclic Graph (DAX): Esse formato utilizado pelo sistema Pegasus.
Mecanismos	Pegasus: Este sistema é responsável por mapear os <i>workflows</i> para os recursos necessários na infra-estrutura. Wings: É um sistema para criação de workflow, que utiliza representação semântica e técnicas de planejamento. Assim permite a concepção de templates e instancias do workflow, que serão executadas no Pegasus. Condor DAGMan: Este é o mecanismo de execução utilizados na solução proposta. Ambientes Grid: Este ambiente é a infra-estrutura onde as atividades dos workflows serão executadas.
Artefatos	Workflows templates definidos em OWL-DL: Este trabalho é um objeto semântico definido, que contem a estrutura de alto nível do experimento científico. Este é independente dos dados e é definida em OWL-DL. Instâncias de Workflows em OWL-DL: Esse <i>workflow</i> é um objeto semântico definido, isto é, independente de recursos de execução. Este <i>workflow</i> contém os dados que serão utilizados na execução. Esse <i>workflow</i> deve ser tipo DAG. Workflows executáveis: Esse <i>workflow</i> é dependente de dados e recursos de execução. Esse <i>workflow</i> será executado. Esse é especificado no formato DAX.
Atividades	A abordagem é composta por três tarefas gerais. Primeiro, o template de <i>workflow</i> é criado, e apenas a estrutura de alto nível está definida. Em segundo lugar, a instância do <i>workflow</i> é criada quando alguns conjuntos de dados são ligados ao template de <i>workflow</i> fluxo de trabalho. Por último, o <i>workflow</i> executável é criado e executado no Pegasus no ambiente de Grid.
Tipo de domínio	O artigo não especifica qualquer tipo de restrição sobre o uso da abordagem. A única restrição é quanto à natureza do <i>workflow</i> , porque os <i>workflows</i> devem ser tipo DAG (<i>directed acyclic graph</i>). O domínio, onde a abordagem foi usada, é a earthquake science .
Número de projetos	O artigo não especifica o número de projetos em que a abordagem foi utilizada. Apenas um exemplo é apresentado.

Kaestle et. al. 1999	
Dados para a questão primária	
Tipo de fonte	Artigo teórico. Ela não tem nenhum caso real, apenas um exemplo.

literária	
Problemas lidados	P.1- Colaboração entre times distribuídos geograficamente: Atualmente, experimentos científicos têm sido utilizados por grupos de cientistas que estão distribuídos geograficamente. Este cenário exige uma nova divisão ativa de informação, suporte a publicação e busca. P.2- O acesso aos dados criados: Há uma preocupação com o acesso aos conjuntos de dados e os tipos de informação que um cientista está interessado.
Tecnologia de construção	A abordagem utiliza o conceito de documentos Hyper-media para estruturar a solução. Esta abordagem utiliza a integração de componentes para representar a solução proposta.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição, execução e análise . Esta abordagem não lida com o conceito de nível abstrato.
Instrumentos	eXtensible Markup Language (XML): É utilizada como base para a linguagem SEML. Document Type Definition (DTD): É utilizado para definição de <i>tags</i> e relacionamentos entre essas. OpenGIS Object Model standard: It is the format that the OASIS objects are modeled. Multicast Dissemination Protocol (MDP): Esse é o protocolo utilizado na comunicação e disseminação da informação. Virtual private networking technology: Esta tecnologia é utilizada para prover confiabilidade e garantir o controle de acesso aos dados.
Mecanismos	Projeto ESP2Net: Este projeto tem como objetivo permitir a colaboração entre os diferentes grupos de cientistas, de modo a acelerar a realização da experimentação. Este projeto utiliza a linguagem SEML como forma de realizar a colaboração. Experimento Markup Language (SEML): É uma linguagem usada para descrever os experimentos científicos e para que possa ser distribuído entre os diferentes grupos. É uma linguagem baseada em XML. OASIS: É o ambiente utilizado para a definição dos experimentos científicos, e também para proporcionar o acesso aos conjuntos de dados e serviços. Conquest: É o sistema usado pela OASIS para realizar o serviço de consulta distribuída. Semantic multicast framework: É um <i>framework</i> que é responsável pela propagação e filtragem dos documentos e informações entre os grupos distintos. Repositórios XML na Web: Local no qual os documentos são armazenados, de modo que permite a descoberta de mais informações e permite consultas mais detalhadas. SEML navegador e Agente lado-cliente: É um agente responsável por fornecer tratamento para tipos de dados, fornecendo ferramentas diferentes e para a gestão dos interesses dos usuários sobre a execução e replicação. Java Activation Framework: É usado para coordenar a utilização dos dados em ferramenta específica.
Artefatos	Documentos Web especificados para a linguagem SEML: Estes documentos são o único item produzido. Eles contêm não só as informações sobre o proprietário, mas também os experimentos, ferramentas, serviços e conjunto de dados utilizados e produzidos. Perfis para os cientistas: Este perfil permite o acesso ao conjunto de dados e também define o tipo de informação que um cientista está interessado.
Atividades	Não há descrição de um processo ou procedimento de concepção. Mas, há uma abordagem geral para a concepção dos documentos em SEML, a sua divulgação e busca.
Tipo de domínio	O artigo não especifica qualquer tipo de restrição sobre o tipo de domínio que poderia ser usado. Esta parte é sobre o tipo de estruturas que formam a linguagem a SEML.
Número de projetos	O artigo não especifica o número de projetos, onde a abordagem foi aplicada. Mostra-se apenas um exemplo no artigo. "2.2.1 SEML Example".

Li et. al. 2006	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo. Esta abordagem é motivada por um trabalho anterior. "[11] S.C. Cheung, D.K.W. Chiu and S. Till, "A Data-driven Methodology to Extending Workflows across Organizations over the Internet," Proc. 36th Hawaii International Conference on System Sciences (HICSS36), Jan. 2003."
Problemas lidados	P.1- Dificuldade de atender a composição com as exigências do experimento: O artigo sustenta que a dificuldade em representar os requisitos de um experimento simulado vem aumentando. Além disso, a representação tradicional baseada apenas em <i>workflows</i> não é apropriada. A abordagem propõe que é necessário identificar e capturar os outros fluxos que atendam os requisitos do experimento, como: dados, controle de exceção, semânticas e segurança.
Tecnologia de construção	A abordagem utiliza o conceito de web services para a construção de seu ambiente. Aqui, os aplicativos estão disponíveis como serviços web.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição e execução . Mas neste artigo o foco principal é a composição.
Instrumentos	Web Service: É usado para embrulhar as aplicações. Assim, o <i>workflow</i> científico vai usar o

	<p>serviço como meio para evocar a aplicação durante a execução.</p> <p>OWL: É usado para a representação do fluxo semântico, onde as dependências dos conceitos estão definidas. Esta linguagem é utilizada para representar o esquema de dependências como ontologia. Mas como lidar com ontologias heterogêneas não há mais comentários no texto.</p> <p>Diretório UDDI: Este é o repositório onde os serviços serão armazenados, assim os participantes e as organizações podem utilizar esses serviços em seus experimentos.</p> <p>Editor de fluxo e visualização: É o módulo que ajuda a projetar os fluxos e as vistas do processo. Ele é usado dentro do ADOME-WFMS.</p> <p>Gerenciador de Interação: É o módulo encarregado de todos os fluxos com o parceiro. Possui subsistema que lida com a instanciação dos <i>views</i> de processos, gestão e acompanhamento, por último, o mapeamento das mensagens.</p>
Mecanismos	<p>Tecnologia de Fluxo: É a tecnologia por trás da abordagem proposta. Aqui, os requisitos dos experimentos são definidos como os fluxos de diferentes tipos, como dados, controle de exceção, semânticas e segurança. Esta abordagem utiliza também a ontologia, como o fluxo semântico, para descrever a relação entre os outros fluxos.</p> <p>ADOME-WFMS: Este sistema é um sistema web-enabled de gestão de <i>workflow</i> que permite aos cientistas ou organizações criar e publicar suas visões do processo e serviços. É composto por um Editor de fluxo e visualização o diretório UDDI e Gerenciador de Interação.</p>
Artefatos	<p>Flows: It is the flows of messages identified for meet the requirements of the experiment to be defined. In fact, could be of the following types: semantic flows, control flows, data flows, security flows, and exception flows.</p> <p>Process view: It is a view of a structurally correct subset of a process, which it can be composed by distinct Flows, like control, semantic, and so on.</p> <p>Fluxos: É o fluxo de mensagens identificadas para satisfazer as exigências do experimento a ser definido. Na verdade, pode ser um dos seguintes tipos: fluxos semânticos, fluxos de controle, fluxos de dados, os fluxos de segurança, e os fluxos de exceção.</p> <p>Visão de Processos: É uma visão de um subconjunto estruturalmente correto de um processo, que pode ser composto por fluxos distintos, como o controle, as semânticas e assim por diante.</p>
Atividades	<p>Esta abordagem tem duas fases. Na primeira fase, as organizações criam visões para personalizar e restringir o acesso dos seus serviços. Na segunda fase, com uma abordagem de composição de serviços, os parceiros podem criar seus próprios aplicativos, invocando como apropriado.</p>
Tipo de domínio	<p>A abordagem trabalha com o conceito de Web Semântica. Esse se concentra em experimentos científicos distribuídos entre as organizações que utilizam os serviços da Web descrito em forma de fluxo. O artigo não especifica qualquer tipo de restrição quanto ao tipo de domínio. O estudo de caso é apresentado no domínio da Biologia.</p>
Número de projetos	<p>Apenas um projeto é descrito no artigo. Mas, o número de projetos, onde a abordagem foi aplicada, não é especificado.</p>

Liu et. al. 2006 e Liu et. al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo teórico. Apresenta a abordagem proposta, mas não está claro se o exemplo é real. Este artigo está relacionado com outro selecionado: "On design, verification, and dynamic modification of the problem-based scientific workflow model".
Problemas lidados	<p>P.1- Poucos modelos são adequados para a complexa modelagem científica: Workflows científicos têm características como: processo científico não pode ser totalmente formalizado a priori; workflows científicos podem mudar durante a execução; o cientista tem um papel importante. Este cenário leva a necessidade de estudar mais esse domínio e assim propor um modelo baseado em problemas que é mais adequado à modelagem científica.</p> <p>P.2- Ambiente que permite a automatização do processo científico: O processo científico modelado com conceito de resolução de problemas deve ser desenvolvido, de modo que poderá permitir a partilha colaborativa de recursos.</p>
Tecnologia de construção	Não foi possível identificar.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com composição e execução .
Instrumentos	<p>Ambiente Grid: São os recursos computacionais que podem ser usados durante a execução do modelo de <i>workflow</i> científico.</p> <p>Modelo de <i>Workflow Scientific</i>: Este é o modelo que representa o <i>workflow</i> científico. Este contém as unidades básicas (problemas) que representam o <i>workflow</i>, mas também o controle, dados e relações de contexto.</p> <p>Descrições de Classes: Estas são as especificações dos recursos computacionais (como ferramentas e sistemas disponíveis no ambiente de grade) e recursos humanos (como os cientistas).</p>
Mecanismos	<p>Modelo de <i>workflow</i> científico baseado em problemas: Essa é a perspectiva da modelagem de <i>workflow</i> científico. Aqui, o <i>workflow</i> científico que representa o processo científico é decomposto em problemas, que é uma unidade básica do modelo.</p> <p>Engine de Execução: <i>Engine</i> que irá executar o modelo de <i>workflow</i> científico, de acordo</p>

	com os recursos de rede (disponível em ambiente de grade), definidos nas descrições de Classes.
Artefatos	Experimento no modelo de workflow científico baseado em problemas: É o experimento (o processo científico), representado pelo modelo de <i>workflow</i> científico. Assim, o experimento é decomposto em uma série de problemas e a relação entre eles. Descrições dos recursos e os cientistas dos experimentos: Esta é uma descrição dos recursos (computacionais e humanos) necessários para a execução do experimento (no formato do modelo de <i>workflow</i> científico).
Atividades	A abordagem é dividida em duas atividades gerais . Na primeira atividade, é modelado o modelo de <i>workflow</i> científico, no qual é definido um conjunto de problemas que compõem os experimentos e as suas relações (entrada, saída e dados de contexto). O conjunto de recursos necessários para o modelo de <i>workflow</i> científico também é definido em uma série de Descrições de Classes . Na segunda atividade, há a vinculação dos recursos para os problemas e para a execução do modelo de <i>workflow</i> científico por uma <i>Engine</i> de execução.
Tipo de domínio	O exemplo é do domínio de previsão de clima . Mas não está claro se a abordagem pode ser usada em outros domínios. Também a abordagem centra-se na utilização de recursos Grid. O exemplo é do domínio de previsão de clima . Mas não está claro se a abordagem pode ser utilizada em outros domínios. "6. Case study".
Número de projetos	Este artigo não especifica o número de projetos no qual foi utilizado.
Dados para a questão secundária	
Especificação da técnica de verificação	A abordagem utilize análise formal para identificar problemas no experimento PBSWM.
Descrição da técnica de verificação	
Instrumentos	Algoritmos para verificar a PBSWM: conjunto de algoritmos que se baseiam nos teoremas 1 a 4, e pode ser usado para verificar a corretude (solidez) do experimento no PBSWM. Lista de casos que as alterações possam afetar a exatidão do experimento PBSWM: As modificações no experimento PBSWM devem ser seguidas de um dos casos, para detectar problemas no novo modelo.
Procedimentos	O procedimento tem duas fases . Primeiro, um conjunto de algoritmos são definidos com base nos teoremas de 1 a 4 e são aplicados no modelos de experimento PBSWM. A segunda fase só é utilizada quando ocorrer uma alteração durante a execução do experimento em modelo PBSWM.
Tipo de domínio	Essa técnica só pode ser aplicada sobre os experimentos definidos como modelo PBSWM.

Medeiros et. al. 2005	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Documentação dos experimentos: O artigo defende que os <i>workflows</i> são um bom paradigma de documentação. Porque os <i>workflows</i> são capazes de manter o rastro de evolução e anotação. O artigo propõe um <i>framework</i> de apoio a esta tarefa.
Tecnologia de construção	A abordagem utilize o conceito de Web semantic para embrulhar (wraps) as aplicações que são utilizadas no <i>workflow</i> .
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com composição e execução . A abordagem lida com níveis abstrato e concreto.
Instrumentos	WS-BPEL: É a linguagem que representa o fluxo de trabalho. Normas Semânticas: Estas normas são utilizadas para o apoio do mecanismo DCC.
Mecanismos	WOODSS framework: é <i>framework</i> para a criação e execução dos experimentos científicos que utiliza os <i>workflows</i> . Foi estendido para a web e anotação de experimentação. O WOODSS tem dois modos. No primeiro modo, o cientista implementa seu modelo em algum sistema (como o GIS), depois dela, WOODSS captura e gera o <i>workflow</i> . Esses são armazenados em um repositório. Na segunda modalidade, os cientistas podem editar, consultar e anotar o <i>workflow</i> que é obtido a partir do repositório de <i>workflow</i> . Digital Content Components (DCC): Este é um mecanismo que envolve um conteúdo digital. Isto é baseado no conceito de web semântica e ontologias. Este conceito ajuda ao reuso pelos cientistas.
Artefatos	Dados e tipos de atividade: Estes são os blocos de construção que serão utilizados na construção do <i>workflow</i> para o WOODSS. Estas são definidas em termos de interfaces. Workflows abstratos definidos no WOODSS: Trata-se de uma representação de alto nível do experimento. O <i>workflow</i> abstrato contém as transições e as dependências entre os tipos de atividades. Workflows concretos definidos no WOODSS: Estes são o <i>workflow</i> executável que é conseguido através da associação das atividades, que estão presentes no <i>workflow</i> abstrato com os recursos. Eles são especificados em WS-BPEL.

	DCC unidades reutilizáveis: Estas são as unidades criadas pelo WOODSS e são registradas em um repositório. Assim, a DCC pode ser reutilizada para cientistas criarem outros <i>workflows</i> . Estas unidades DCC reutilizáveis podem ser de dois tipos, um relacionado aos componentes de processo abstrato e outra está relacionada com componentes do processo executável.
Atividades	A abordagem é composta por duas fases . Primeira fase: o cientista define um conjunto de atividades e tipos de dados que serão utilizados na construção do <i>workflow</i> abstrato. Durante essa fase, o <i>workflow</i> abstrato é mapeado para o conjunto de agentes (recursos de infra-estrutura) para a criação do <i>workflow</i> concreto. Segunda fase: O cientista pode procurar uma unidade DCC reutilizáveis e compor seus novos experimentos.
Tipo de domínio	A abordagem é aplicada no domínio agro-ambiental . No entanto, o estudo de caso (seção 5.2), tem como exemplo o domínio da Bioinformática . A abordagem foi estendida para suportar outros domínios.
Número de projetos	O artigo não especifica o número, mas descreve dois estudos de casos (seção 5).

Sorde et. al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Complexidade da concepção do <i>workflow</i> para ambiente em grade: Com o uso crescente de ambientes em grade, a complexidade da concepção aumentou. Assim, as novas formas de representar o fluxo surgiram, utilizando não só, seqüencial e paralelo, mas também decisão e os padrões de iteração.
Tecnologia de construção	Não foi possível identificar.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição e execução . Mas o artigo centra-se na Composição (com verificação). Centra-se também no nível concreto, porque ele só captura o fluxo dos programas na forma de <i>Directed Graph</i> .
Instrumentos	Workflow DAG grafo: É o modelo base de <i>workflow</i> que permite a especificação de <i>workflows</i> . Este modelo é DAG, por isso não é possível representar iteração e loops. Grafo Petri-net: Este é um modelo de formato que é usado na verificação da correção do <i>workflow</i> . O modelo de <i>workflow</i> é transformado para este formato de modelo. Sun Grid Engine: Este é o ambiente de grade utilizado no mecanismo de execução. Repositório centralizado: Local onde os aplicativos (classes) são registrados. Eles são acessados durante a composição.
Mecanismos	Mecanismo de Execução (GWMS): O mecanismo não é descrito com detalhes no artigo. Mas, este foi construído em cima do Sun Grid e é capaz de mapear os <i>workflows</i> para a Petri-net, assim o mecanismo pode verificar qualquer erro. Este é responsável pela execução dos <i>workflows</i> . Modelo de <i>workflow</i> non-Dag: É um modelo base de <i>workflow</i> que permite a especificação de grafos de <i>workflow</i> não-DAG. Assim, é possível representar iteração, loops, escolha e padrões paralelos.
Artefatos	Grafo de <i>workflow</i> Non-Dag: É o <i>workflow</i> que representa os experimentos. Trata-se de um <i>workflow</i> no modelo de <i>workflow</i> não-DAG. Esse fluxo é utilizado no nível de usuário (para os cientistas). Os nós da tarefa são no formato XML e são registrados em um repositório. Grafo WF-net: É o resultado da transformação de um grafo de <i>workflow</i> não-DAG em um grafo no formato Petri-net. Ambos são relacionados e representam os mesmos experimentos. O grafo WF-net não é diretamente utilizado pelos cientistas.
Atividades	A abordagem tem duas tarefas gerais . Primeiro, o cientista cria um <i>workflow</i> , o grafo <i>workflow</i> não-DAG, utilizando a notação modelo de <i>workflow</i> não-DAG. Depois, o <i>workflow</i> é mapeado em WF-net grafo, com base no grafo Petri-net, onde técnicas de análise Petri-net para as verificações podem ser aplicadas.
Tipo de domínio	O artigo propõe um modelo que pode ser utilizado para os campos da ciência, como bioinformática, meteorologia e assim por diante. Ele não especifica qualquer tipo de restrição sobre o domínio, mas tem o exemplo do <i>workflow</i> ENAN que é no domínio da bioinformática. A abordagem é usada para compor os <i>workflows</i> que são baseados na tecnologia de Grid .
Número de projetos	O artigo não especifica o número de projetos em que a abordagem foi utilizada.
Dados para a questão secundária	
Especificação da técnica de verificação	A técnica de verificação é baseada em análise de rede de Petri. O grafo de <i>workflow</i> non-DAG é transformado em um grafo WF-net (baseado em grafo Petri-net).
Descrição da técnica de verificação	
Instrumentos	Regras de mapeamento de grafos de <i>workflow</i> non-DAG para grafo Petri-net (WF-net): este instrumento que é usado para gerar o gráfico WF-net (com base no gráfico Petri-net) a

	partir de um grafo de workflow non-DAG. Algoritmos para verificação do Petri-net (WF-net): este é um instrumento que é utilizado para verificar os experimentos mapeados no gráfico WF-net (com base no gráfico de Petri net).
Procedimentos	O processo tem duas etapas. Primeiro, o grafo de <i>workflow</i> não-DAG é mapeado para uma WF-net (com base no gráfico Petri Net), seguindo regras de mapeamento definidas. Depois disso, um algoritmo de verificação é aplicado no grafo WF-net, assim os problemas podem ser detectados.
Tipo de domínio	A verificação técnica é aplicada no modelo baseado em grafo de rede Petri.

Troger and Fernandes 2004	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Os experimentos in silico têm falta de isomorfismo com a prática in vitro: A abordagem propõe uma forma de apoiar o biólogo na modelagem de experimentos in silico de aproximando dos métodos de investigação que eles estão familiarizados. P.2- A falta de apoio para relacionar a obtenção de provas para uma hipótese explícita e um processo de validação: O artigo afirma que as ferramentas de bioinformática não suportam a relação entre a fase de coleta de provas do experimento, a hipótese definida e processo de validação. Esta tarefa é geralmente feito off-line e não são diretamente vinculadas com a coleta de provas. P.3- A falta de rastreabilidade das execuções do experimento e sua evolução: O artigo afirma que as ferramentas não oferecem suporte à rastreabilidade dos conjuntos de dados utilizados, a execução de diferentes experimentos. Afirmou também que ferramentas não suportam a evolução do experimento, ao longo do tempo. Isto parece estar relacionado com o aspecto de proveniência do experimento.
Tecnologia de construção	De fato, a abordagem propõe uma infra-estrutura e linguagem que age como uma camada em alto nível. Esta camada se mantém acima da engine de execução, compliando os workflows traduzidos para a infra-estrutura.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a composição, execução e análise . A abordagem propõe uma nova linguagem para a especificação do <i>workflow</i> , uma infra-estrutura para compilá-lo para um <i>engine</i> de execução (não fornecidos pela abordagem). A análise é apoiada pela rastreabilidade provida.
Instrumentos	Modelo conceitual para ISXL: É um modelo onde a relação lógica entre as partes de um experimento é declarada. Diagramas de transição de estado: é a forma em que a semântica operacional da ISXL estão definidos. Engine PML / PWI: É um mecanismo de execução que foi usado para a prova de aplicação do conceito. Xindice: Este é um banco de dados XML nativo que é usado como repositório para a infra-estrutura. XML: É o formato que os workflows são compilados pela infra-estrutura e armazenados.
Mecanismos	ISXL framework lingüístico: Linguagem de <i>workflow</i> que fornece suporte explícito para o recolhimento de evidência, com base em dados e integração de processos, e que liga o encontro com a hipótese de provas e do processo de validação. ISXL ambiente persistente: É um ambiente onde os experimentos são definidos e compilados e promulgada por um mecanismo de execução. Este é composto também por um repositório.
Artefatos	Modelo dos experimentos descritos na linguagem ISXL: É o modelo do experimento descrito em linguagem ISXL. Aqui, os métodos, materiais, especificação do protocolo (hipótese e validação). Mais tarde, ele será compilado e ligado por um mecanismo de execução. Metadados gerados pela execução: Esses metadados são relacionados aos dados produzidos durante a execução de um experimento. Estes são mantidos em repositório.
Atividades	Não existe nenhum procedimento ou processo descrito no artigo. Mas, foi possível identificar que a abordagem tem duas fases . A primeira está relacionada à definição do experimento na linguagem ISXL. Depois disso, este modelo é compilado e executado em um motor. O modelo e metadados gerados por esta execução são mantidos.
Tipo de domínio	A abordagem parece ser criada para experimentos in silico em Bioinformática .
Número de projetos	Este artigo não especifica o número de projetos em que a abordagem foi utilizada. O artigo apresenta apenas um exemplo - "3. Exemplo motivador".

Verdi et. al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas	P.1- Faltam suporte a validação e correção do processo de experimentação: Os

lizados	<p>experimentos geralmente permanecem na mente cientista. Muito embora cada sistema de workflow científico tenha a sua própria linguagem e, adicionalmente, tem-se o fato de que cada cientista tem seu próprio método para codificar o <i>workflow</i> concreto. Este cenário leva a uma necessidade de representar a experiência em um modelo conceitual.</p> <p>P.2- A falta de documentação leva à dificuldade na reprodutibilidade: O experimento pode ter variabilidade nos seus procedimentos e raramente é devidamente documentado. Esse cenário leva à dificuldade para a reprodutibilidade dos experimentos.</p> <p>P.3- Método para orientar a concepção do trabalho abstrato: O sistema de gestão científica fornecendo a funcionalidade de representação de workflows em vários níveis de abstração. No entanto, falta a eles suportarem um modelo de alto nível, além disso, nenhum fornece orientações sobre como criar um modelo de workflow abstrato.</p>
Tecnologia de construção	Não existe uma tecnologia de construção descrita no artigo. Mas a abordagem é parte de uma pesquisa maior, cujo foco é a construção de uma estrutura de apoio ao processo experimental no contexto do experimento RMN.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem centra-se na composição . A abordagem lida com a concepção (com atividades de validação) de <i>workflow</i> científico abstrato.
Instrumentos	<p>Modelo de contexto: Esse modelo é usado para representar uma visão hierárquica do processo experimental, descrevendo os principais processos e seus sub-processos, que são partes do processo de experimentação. Este processo representa o experimento como uma árvore de processos.</p> <p>Fornecedor-Produtor-Input-Output-Consumidor (SIPOC) modelo: Este modelo representa a entrada e saída, fornecedor e consumidor de um determinado processo. Isto também representa o fluxo em que os dados transitam no processo.</p> <p>Modelo de controle de fluxo: É um modelo que representa a orquestração do processo (atividade do processo de experiência).</p>
Mecanismos	<p>Workshops: técnica de identificação de requisitos utilizada em todas as fases da abordagem para identificar os processos e modelos.</p> <p>Walkthrough: É utilizada pelo especialista de domínio (cientista envolvido no levantamento) para rever o modelo. Eles procuram defeitos ou entre os modelos.</p>
Artefatos	<p>Um diagrama de modelo de contexto: É um diagrama que representa o processo e sub-processos do processo experimental (super-workfluxo), sob a identificação. Este diagrama é criado na fase Construir Context Model.</p> <p>Um conjunto de processos representados por modelos SIPOC: Para cada processo de alto nível identificado pelo modelo de contexto, é criado um modelo SIPOC. Esses modelos contêm para o processo: a entrada e o fornecedor, o produto e o consumidor. Esses são criados na fase de Construção SIPOC Model.</p> <p>Um conjunto de workflows representados por modelos de fluxo de dados: Essas são criadas no modelo de fluxo de dados Construct, onde um workflow representado por um modelo de fluxo de dados é criada para cada processo e sub-processo.</p>
Atividades	A abordagem contempla três fases , onde cada uma é composta por duas etapas : uma concepção e validação. Primeiro, a fase Construir modelo de contexto que identificam o processo e sub-processo do processo de experiência e oferece uma visão estrutural dos mesmos. Por segundo, a fase de construção modelo SIPOC que identificam as entradas e saídas (de dados) e como ele flui em um processo de experimentação. Por terceiro, construir modelo de fluxo de controle que capta o fluxo e a ordenação das etapas do processo de experimentação. A abordagem utiliza técnica identificação de requisitos para identificar o workflow (processá-lo e sub-processo), durante cada reunião. Além disso, ele poderia usar para cada fase, na etapa de validação, o Walkthrough (revisão técnica), com os cientistas envolvidos no levantamento, para identificar defeitos ou entre os modelos. Existem três papéis na abordagem: um especialista em modelagem de processos de negócios; especialista de domínio experimento e, facilitador participante. Esta abordagem utilizado o Microsoft Visio como um editor de modelador visual e é usada para capturar todos os modelos.
Tipo de domínio	O artigo não especifica qualquer tipo de restrição. Mas, a abordagem foi desenvolvida para um projecto no domínio da Bioinformática e o exemplo está no domínio da bioinformática (RMN) .
Número de projetos	Apenas se mostra como um estudo de caso. O artigo não especifica o número de projetos em que a abordagem foi aplicada.
Dados para a questão secundária	
Especificação da técnica de verificação	A técnica de verificação é baseada em atividades de engenharia de software e técnicas. A abordagem utiliza a atividade de validação para a garantia da qualidade dos modelos e, por isso, os cientistas (revisores) realizam um walkthrough .
Descrição da técnica de verificação	
Instrumentos	Walkthrough: esta é a técnica de revisão utilizada para identificar os problemas dentro e entre os modelos.
Procedimentos	O processo de garantia da qualidade baseia-se na validação do modelo criado. A principal técnica utilizada é walkthrough do modelo, e que consiste na revisão dos documentos pelos cientistas envolvidos na especificação.
Tipo de domínio	Não há descrição ou restrição descrita no artigo.

Yang et. al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- A complexidade da nova colaboração em pesquisa científica com base em workflows: A pesquisa científica tornou-se colaborativa e envolve mais grupos distribuídos. Eles compartilham dados, arquivos e conhecimento. Devido a isso o fluxo de informação torna-se uma chave para a segurança.
Tecnologia de construção	Esta abordagem é baseada no conceito de modelo formal, máquina do Estado hierárquica . A abordagem estende esse conceito para propor a solução.
Descrição do processo/abordagem	
Fase do ciclo de vida	A abordagem lida com a fase de composição (com verificação).
Instrumentos	Não foi possível extrair qualquer informação.
Mecanismos	Máquina do Estado hierárquica para workflow científico: É um modelo criado para representar o workflow científico. Este modelo tem como base a Máquina-Estado hierárquica.
Artefatos	Modelo de workflo composto por atores com uma HSMSW associados: um modelo composto por elementos ator. Para cada ator do workflow é criado um HSMSW associados.
Atividades	Não existe um procedimento detalhado descrito no artigo. Mas, a abordagem de modelagem é baseada na composição de um modelo composto por elementos ator. Para cada ator atômico (como acontece com as atividades) do workflow é criado um HSMSW associado.
Tipo de domínio	Nenhum tipo de domínio não foi especificado no artigo.
Número de projetos	O artigo tem apenas um exemplo. Um exemplo motivador.
Dados para a questão secundária	
Especificação da técnica de verificação	Ela propõe uma técnica para verificar a HSMSW que representa o workflow científico. Ele usa algoritmo para verificar formalmente o workflow.
Descrição da técnica de verificação	
Instrumentos	Algoritmo para a construção de máquinas de estado finito: Este algoritmo é utilizado para verificar o HSMSW. Técnica de controle de fluxo de informação: É utilizado para detectar problemas no fluxo dos objetos, e também detectar problemas de entrada e saída.
Procedimentos	O procedimento consiste basicamente na aplicação de um algoritmo para traduzir um ator para uma máquina de estado finito, depois é aplicada uma técnica de controle de fluxo de informações para verificar a correção do modelo.
Tipo de domínio	Só para o modelo baseado em HSMSW. Esta técnica é para verificar formalmente as propriedades de fluxos de trabalho científico.

Jablonski et. al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Complexidade em lidar com o gerenciamento do processo e dos dados. A complexidade do gerenciamento dos processos e dados aumentou devido às aplicações científicas mais complexas e grande volume de dados manipulados.
Tecnologia de construção	Utiliza a integração de componentes (frameworks) para apresentar sua solução.
Descrição do processo/abordagem	
Fase do ciclo de vida	Composição e Execução.
Instrumentos	iPM: Ferramenta para suportar a modelagem. IE: Ferramenta que faz a ligação com os recursos para execução.
Mecanismos	Perspective Oriented Process Modeling Framework: serve para criar as Perspectivas utilizadas na modelagem dos workflows. O processo experimental pode ser dividido em diversas perspectivas como dados, funcional, organizacional e operacional. Process Driven Architecture framework: Serve para criar o ambiente onde os dados reais serão imputados e também a infra-estrutura computacional para execução do workflow (computadores, clusters). Data Logistic with Ontologies framework: Serve para conectar as diferentes aplicações que possam ter seus dados em formatos distintos.
Artefatos	Workflow Perspective models: São os modelos que compõem o workflow, são perspectivas do mesmo workflow.
Atividades	Não é muito bem descrito. Mas primeiro é realizado uma modelagem no IPM, depois é

	instanciado pelo IPE. O Dalton faz a ligação entre os recursos, fazendo as conversões necessárias.
Tipo de domínio	Não há restrições explícitas sobre domínios. Foi aplicado no domínio de ecologia.
Número de projetos	Não há descrição do número de projetos, somente um exemplo é descrito na Seção 4.

Kalyanam et. al. 2007	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Diminuir a complexidade da composição. São duas limitações que buscam combater: Evitar manipular código e integrar as aplicações.
Tecnologia de construção	Utiliza a integração de componentes para apresentar sua solução.
Descrição do processo/abordagem	
Fase do ciclo de vida	Composição e Execução.
Instrumentos	WSDL Web Service Description Language: Liguagem de descrição dos serviços web utilizados como componentes do workflow. BPEL Business Process Execution Language: Linguagem de execução do workflow.
Mecanismos	JOpera: Ferramenta que prove um ambiente para execução e modelagem.
Artefatos	Workflow process: Descrição dos serviços web encadeados como um processo experimental.
Atividades	Não são descritas em detalhes. Existem 2 fases: Composição e Execução. Primeiro é construída a atividade (service web) para ser utilizada na concepção do workflow abstrato (alto nível). Posteriormente, essa atividade é ligada a um recurso computacional pelo Jopera.
Tipo de domínio	Não há descrição à restrição quanto a domínio. O domínio no qual foi aplicado na referência é no de previsão climática.
Número de projetos	Não é descrita a quantidade exata de projetos em que a solução foi aplicada. Um exemplo é descrito no domínio

Lacroix et. al. 2009	
Dados para a questão primária	
Tipo de fonte literária	Artigo descritivo.
Problemas lidados	P.1- Lidar com abstração na Composição. Conseguir combinar atividades do workflow que são de simulação com as atividades de avaliação dos parâmetros a serem utilizadas. Segundo os autores houve aumento na complexidade da composição por conta do grande número de atividades desses tipos no workflow.
Tecnologia de construção	Descrição semântica do domínio.
Descrição do processo/abordagem	
Fase do ciclo de vida	Composição.
Instrumentos	Não foi identificado nenhum instrumento.
Mecanismos	ProtocolDB: Ferramenta proposta pela abordagem no qual pode-se modelar o workflow e depois executá-lo.
Artefatos	Workflow design: Modelo em alto nível do experimento, composto por elementos descritos na ontologia.
Atividades	São duas fases. Projeto do workflow: Fase na qual o experimento é expresso como workflow em alto nível de conceitos do domínio. Feito no ProtocolDB. Análise do workflow: Fase na qual o experimento é expresso como workflow executável conectado aos recursos computacionais.
Tipo de domínio	Não há descrição de restrição de domínio. O exemplo é do domínio Medicina. Entretanto, ele faz uma verificação com experimentos in vitro para avaliar o experimento in silico (workflow) criado.
Número de projetos	Não há descrição do número de projetos que a abordagem foi aplicada.

ANEXO D – TAMPLATE PARA ESPECIFICAÇÃO DE WORKFLOW ABSTRATO

O *template* para especificação de workflow abstrato é composto pelos seguintes tópicos:

CONTROLE DE VERSÃO

Esta seção contém

Versão	Autor	Alteração	Data
0.1	JOÃO	CORREÇÃO FORMULÁRIO DE ATIVIDADE "CADASTRAR DADOS" PG. 28	20/11/2008

1. INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO DO DOMÍNIO

Pequena descrição do domínio do experimento científico e de seu conjunto de atividades. Não deve ser longo, ultrapassando duas páginas.

1.2 LEGENDA E NOTAÇÃO GRÁFICA

Texto padrão descrevendo a notação gráfica e formulários padrão.

A fim de facilitar o entendimento dos modelos apresentados, são representados, em uma legenda (na Figura 1), os principais elementos utilizados no diagrama de atividades. Os nós de atividade e ação representam, no modelo UML, os procedimentos da pesquisa como atividades do workflow abstrato. Enquanto o nó atividade pode conter a sub-atividades ou ações internamente, o nó ação representa uma etapa atômica do processo experimental, não podendo ser sub-dividida. Ambos os nós, atividade e ação, podem conter diversos nós de objeto associados. Os nós de objeto representam, nesta abordagem, os insumos (quando o artefato é utilizado na execução da atividade ou ação) ou produtos (quando o artefato é produzido durante a execução da atividade ou ação). Existem ainda os nós de controle que têm por objetivo realizar o sincronismo, divisão e decisão nos fluxos de dados e controle. O nó de controle do tipo Decisão deve apresentar as condições de guarda (entre colchetes []), que representam os possíveis resultados da decisão. O nó do tipo Merge imediatamente passa para um nó de saída todos os fluxos de dados ou de controle, possui a mesma notação do nó Decisão, por isso para facilitar a visualização foi atribuída à cor cinza ao nó do tipo Merge. A Divisão de fluxo representa a divisão do

fluxo de dados ou de controle em diversos fluxos de saída ao mesmo tempo, há uma replicação dos controles e dados oriundos dos fluxos de entrada.

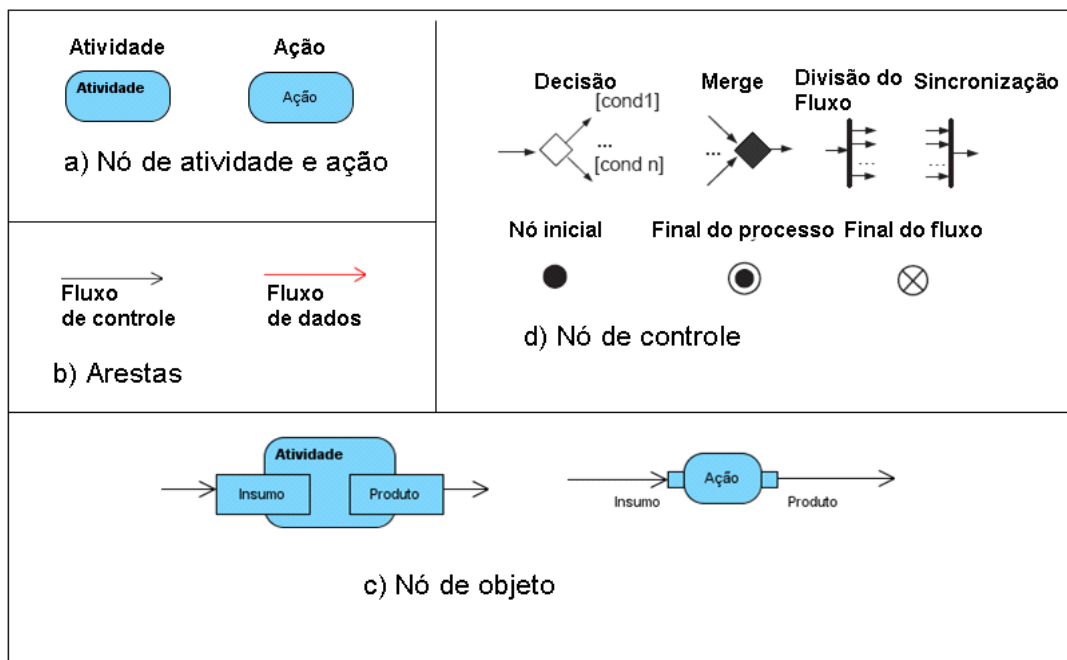


Figura 1. Legenda com elementos do diagrama de atividades (UML 2.0) utilizados na modelagem dos workflows abstratos.

Para enriquecer a modelagem, foi criado um padrão de representação gráfico e estereótipos para os elementos utilizados na modelagem. A Figura 2 apresenta o padrão relacionado aos elementos Atividade ou Ação. O estereótipo é marcado pelo símbolo “<< >>” com o nome no meio. Nesta abordagem, existem quatro estereótipos definidos: Manual (Figura 2 – a); Semi-Automatizada (Figura 2 – b); Automatizada (Figura 2 – c); Sub-Workflow (Figura 2 – d); e, Workflow (Figura 2 – d). Uma ação pode ser Manual, Semi-Automatizada ou Automatizada.

O estereótipo <<Manual>> é atribuído a uma ação na qual não há ou é pouca a utilização de ferramentas que auxiliem na execução, sendo essencial a participação do pesquisador. O estereótipo <<Semi-Automatizada>> é atribuído a uma ação, na qual é necessária a utilização de ferramentas que auxiliem na execução, sendo fundamental a participação do pesquisador. O estereótipo <<Automatizada>> é atribuído a uma ação, na qual existe a utilização de ferramentas que auxiliam na execução, sendo quase nenhuma a participação do pesquisador. Já o estereótipo <<Sub-Workflow>> somente se aplica a elementos do tipo ação que representam atividades, pois estes podem ser compostos por outros elementos atividades ou ações, além de fluxo de dados e controle internos e nós de dados e controle. Assim, definindo-se Sub-workflow

é etapa do processo experimental encapsulado, contendo outras atividades ou ações encadeadas de maneira organizada. O estereótipo <<Workflow>> somente se aplica a elementos do tipo atividade, na modelagem.

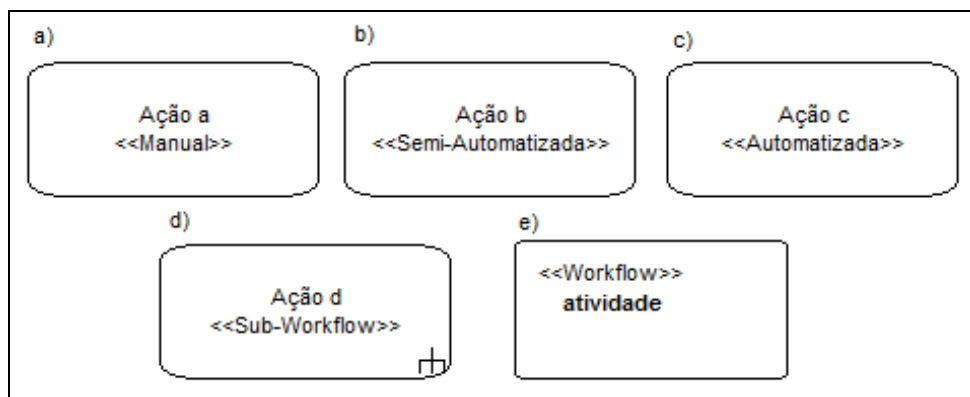


Figura 2. Notação gráfica e padrão de representação das atividades e ações.

A fim de representar visualmente os artefatos e as ferramentas, dois modelos auxiliares são utilizados, cujos elementos são classes do diagrama de classes da UML. Os estereótipos adotados estão representados na figura 3. Na Figura 3-a se utiliza o estereótipo <<Artefato>> para um artefato consumido/produzido no experimento científico. Na figura 3-b se utiliza o estereótipo <<Ferramenta>> para uma ferramenta utilizada na execução de uma atividade do experimento científico.

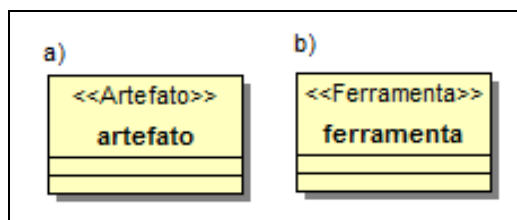


Figura 3. Notação gráfica e padrão de representação das ferramentas e dos artefatos.

A seguir estão apresentados os três formulários utilizados neste documento, contendo os seus campos e as suas descrições. A tabela 1 apresenta o formulário de Atividade; a tabela 2 apresenta o formulário de Artefato e a tabela 3 apresenta o formulário de Ferramenta.

Tabela 1. Formulário de Atividade e seus campos.

Atividade	[NOME DA ATIVIDADE]
Descrição	Descreva as etapas de execução da atividade de forma detalhada, também deve descrever o objetivo da atividade.

Tipo de atividade	Informe o tipo de atividade. Os possíveis tipos de atividades são: <ul style="list-style-type: none"> Manual: uma atividade na qual não há ou existe pouca utilização de ferramentas que auxiliem em sua execução, sendo essencial a participação do pesquisador/cientista. Semi-Automatizada: uma atividade ou ação na qual há uma interação do pesquisador e apoio de algum recurso computacional necessário. Automatizada: uma atividade ou ação que não requer a participação essencial do cientista, pois existe recurso computacional responsável por apoiar sua execução completamente.
Obrigatoriedade	Informe se essa atividade sobre a sua obrigatoriedade. As possíveis classificações são: obrigatória ou opcional. Caso a atividade seja classificada como opcional, cite as situações em que a sua execução é desejada.
Ferramenta	Liste os <i>softwares</i> , sistemas ou simuladores utilizados durante a execução da atividade.
Insumo	Liste os artefatos consumidos pela atividade. Esses artefatos representam a “matéria-prima” da atividade, ou seja, os dados de entrada necessários para executar a atividade.
Produto	Liste os artefatos produzidos pela atividade. Essa produção pode ser de um artefato novo ou a transformação de um artefato existente. Esses artefatos são os dados de saída produzidos pela atividade. Registre a quantidade gerada para cada produto.
Papel	Liste os papéis desempenhados por um ou mais indivíduos para a realização da atividade. Responsáveis pela execução da atividade (e.g., geólogo).
Pré-condição	Liste as pré-condições que devem ser atendidas para iniciar a execução da atividade (e.g., atividade X executada, recurso financeiro disponível e alocado, etc.).
Pós-condição	Liste as pós-condições que são geradas ao final da execução da atividade (e.g., dado gravado no repositório, recurso computacional liberado).
Pré-atividade	Liste as atividades que devem ter sido concluídas para que seja possível a execução dessa atividade.
Sub-atividade	Liste as atividades que compõem esta atividade (se existirem). Neste caso, cada sub-atividade listada deverá ter uma ficha, equivalente a essa, preenchida.
Capacidade de paralelismo	Informe se a atividade pode ser paralelizada ou distribuída. Caso positivo, cite as situações em que essa atividade pode ser paralelizada.
Risco associado	Liste os riscos associados à atividade. Os riscos são fontes de perigo, possibilidade de perda ou infortúnio. Como exemplo: entrada de dados manual, transformação de dados ou entradas.
Frequência de utilização	Classifique a frequência de uso da atividade. Pode ser utilizada uma classificação qualitativa (Alta, Média ou Baixa). Também pode ser especificado um valor numérico.
Custo computacional	Informe o tempo médio despendido na execução da atividade. Caso haja uma fórmula que calcule o tempo de execução, explicita-a.
Outros comentários	Se necessário, utilize este espaço para fazer outros comentários que julgar pertinente.

Tabela 2. Formulário de Artefato e seus campos.

Artefato	[NOME DO ARTEFATO]			
Descrição	Descreva resumidamente o artefato.			
Origem	A classificação quanto à origem é: interna ou externa. Interna se é produzido por uma das atividades do <i>workflow</i> científico. Externa se é produzido por alguma atividade que não faz parte do <i>workflow</i> científico que está sendo modelado.			
Utilização	Informe na tabela abaixo as atividades na qual esse artefato é gerado/usado. Se o artefato é usado pela atividade, esse é um Insumo. Se o artefato é gerado pela atividade esse é um Produto. Para cada caso em que o artefato é usado/gerado por uma atividade, descreva sua condição de obrigatoriedade. Faça isso para cada atividade com a qual esse artefato se relaciona. As classificações possíveis são: <ul style="list-style-type: none"> Obrigatório: Artefato é um insumo obrigatório de uma atividade, isto é, sempre é consumido; Artefato é produto obrigatório de uma atividade, isto é, sempre é produzido. Opcional: Artefato é um insumo opcional de uma atividade, isto é, pode ser consumido, dependendo de uma determinada condição; Artefato é produto opcional de uma atividade, isto é, pode ser produzido, dependendo de uma determinada condição. 			
	[Informe o formato do artefato (digital, digitalizável ou físico). Digital se é arquivo de computador. Digitalizável se é um documento em papel, que pode ser transformado em um arquivo de computador (através de um scanner). Físico se é uma amostra de material (ex.: rocha, areia, etc.). Caso o Formato do artefato seja digital, pode-se listar a extensão digital do formato.			
	Atividade	Insumo/ Produto	Condição de obrigatoriedade	Formato
Nome da Atividade	Insumo ou Produto	Descrição da obrigatoriedade: Obrigatório; Opcional.	Físico, digitalizável, Digital.	Extensão do artefato

Tipos de extensão digital	A lista de todas as extensões digitais possíveis para o artefato.
Temporalidade	Artefato é um insumo/produto de uma atividade cujo tempo de vida é limitado, isto é, após a execução do experimento científico é descartado. Ex.: O artefato A é Insumo (entrada) da atividade Atv1 e Atv2, mas artefato A é insumo Obrigatório de Atv1 e Opcional de Atv2.
Metadado	Descreva informações relacionadas ao artefato, quando pertinentes. Informações como campo, nº de colunas, disposição dos campos (e.g., colunas de 1 a 4 representam o id, colunas de 5 a 10 representam o peso).
Ferramenta	Se o artefato for um arquivo de computador, diga que ferramenta(s) lida com este arquivo.
Sinônimos	Liste aqui os sinônimos deste artefato.
Outros comentários	Se necessário, utilize este espaço para fazer outros comentários que julgar pertinente.

Tabela 3. Formulário de Ferramenta e seus campos.

Ferramenta	[NOME DA FERRAMENTA]
Descrição	Descreva a ferramenta de forma detalhada.
Tipo de aplicação	Informe o tipo de aplicação. Alguns tipos de ferramentas são: - Batch: Arquivos utilizados pra automatizar tarefas; - Serviço: Aplicação que provê funcionalidades para o uso por outra; - Interface: Aplicação na qual há uma interação com o usuário através de telas, formulários ou comandos.
Versão	Cite a versão da ferramenta utilizada.
Sistema Operacional	Liste os sistemas operacionais nos quais a ferramenta é executada.
Extensões digitais suportados	Liste os formatos que a ferramenta suporta, consome ou manipula.
Local de Execução	Liste os locais de execução que a ferramenta suporta. Alguns tipos são:- • Local – se é executado no mesmo local que o <i>Workflow</i> ; • Remota – se é invocado e é executado em infraestrutura diferente do <i>Workflow</i> ; • Cluster – se é executado em ambientes de alto desempenho; • Grade (Grid) – se é executado em ambientes computacionais dispersos para alto desempenho.
Forma de disparo	Liste a forma de disparo da ferramenta. Esta é a forma como a ferramenta será ativada. Por exemplo: Linha de comando "H:\Movimento\CMD09.EXE".
Outros comentários	Se necessário, utilize este espaço para fazer outros comentários que julgar pertinente.

1.3 OBJETIVO DO DOCUMENTO

Texto padrão para preenchimento com informações específicas do experimento.

Este documento apresenta como objetivo documentar o workflow abstrato referente ao experimento "NOME DO EXPERIMENTO". Este contém uma especificação do workflow científico, contendo um conjunto de formulários, e a sua modelagem, em diagramas de atividades da UML 2.0. O documento é estruturado da seguinte forma:

- Na seção 2 é apresentada uma lista de termos relacionados ao domínio do experimento "NOME DO EXPERIMENTO".
- Na seção 3 é apresentada a lista dos papéis envolvidos no "NOME DO EXPERIMENTO", contendo não só o conhecimento exigido, mas também o seu nível.

- Na seção 4 é apresentada a modelagem do processo “NOME DO EXPERIMENTO”, o resultado é um conjunto de modelos de workflow abstrato correspondente, representados em diagramas de atividades (baseado na UML). Também, na seção 4.2, são apresentados os formulários preenchidos, contemplando as ferramentas, as atividades e os artefatos relacionados aos workflows “NOME DO EXPERIMENTO”.

2 LISTA DE TERMOS

Lista com a definição de termos e conceitos do domínio do experimento científico que são utilizados nos formulários e atividades. Devem-se listar termos, nomes de métodos, nome de conceitos, dentre outros itens.

3 LISTA DE PAPEIS ENVOLVIDOS

Apresenta-se a lista dos papéis envolvidos no “NOME DO EXPERIMENTO”, contendo não só o conhecimento exigido, mas também o seu nível.

4 MODELOS DE WORKFLOW ABSTRATO & FORMULÁRIOS

4.1 MODELOS DE WORKFLOW ABSTRATO

Seção no qual se apresentam os modelos de workflow. Pode-se descrever brevemente o objetivo de cada workflow representado.

4.2 FORMULÁRIOS

4.2.1 ATIVIDADES

Seção no qual estão descritos todos os formulários de atividade.

4.2.2 ARTEFATOS

Seção no qual estão descritos todos os formulários de artefatos

4.2.3 FERRAMENTAS

Seção no qual estão descritos todos os formulários de Ferramentas.

4.3 DIAGRAMA DE FERRAMENTAS & DIAGRAMA DE ARTEFATOS

Seção na qual se descreve os diagramas de Artefatos e Ferramentas.

ANEXOS

Seção no qual se pode colocar qualquer anexo importante para o entendimento do documento.

ANEXO E – METAMODELO PARA EXPERIMENTAÇÃO UTILIZANDO WORKFLOW CIENTÍFICO

Metamodelo para experimentação com base em *workflow* científico

Os elementos como *Workflow*, *Sub-workflow*, Atividade, Artefato e Ferramenta são utilizados em modelos de *workflow* científico, principalmente, os que se encontram em nível concreto, pois são conceitos normalmente representados nas linguagens de modelagem dos diversos SGWfCs. Porém, somente esses elementos básicos não permitem uma representação mais abrangente dos *workflows* científicos, perdendo-se a oportunidade de agregar mais informações aos modelos, visto que não são capazes de representar informações importantes para os procedimentos de um experimento *in silico* ou *in virtuo*.

Por exemplo, atividades manuais realizadas pelos pesquisadores não são usualmente capturadas ou representadas em um SGWfC, como o *Kepler* (ALTINTAS ET.ET AL. 2004) e ou *Vistrails* (CALLAHAN ET.ET AL. 2006). Entretanto, essas são atividades importantes dos experimentos científicos, pois envolvem o julgamento do pesquisador e, em geral, resultam em novos conhecimentos científicos. Logo, as atividades manuais deveriam estar documentadas em algum local, contudo nestes SGWfCs, estas informações são armazenadas em documentos genéricos ou notas e nem sempre são anexadas ao *workflow* científico em utilização. A forma como essa documentação é realizada, acaba por tornar-se ineficiente e propensa a erros (VERDI ET AL., 2007).

Considerando esse problema, especializações dos elementos básicos do *workflow* foram definidas, juntamente com a criação de novos, com o intuito de agregar mais semântica ao modelo. Conseqüentemente, espera-se auxiliar ao pesquisador na representação de suas atividades de pesquisa, registrando detalhes não representados diretamente em *workflow* concretos, como no caso de atividades manuais, mas que são importantes quando se trata de *workflow* abstrato.

Adicionalmente, atributos foram definidos e adicionados aos elementos já existentes e também aos novos para representar as informações que os caracterizam detalhadamente. Como resultado, um metamodelo foi desenvolvido para representar todos esses elementos, seus atributos e seus relacionamentos. A Figura E.1 apresenta o metamodelo proposto contendo: os elementos de modelos do *workflow* científico abstrato, o relacionamento entre estes e os atributos que os caracterizam.

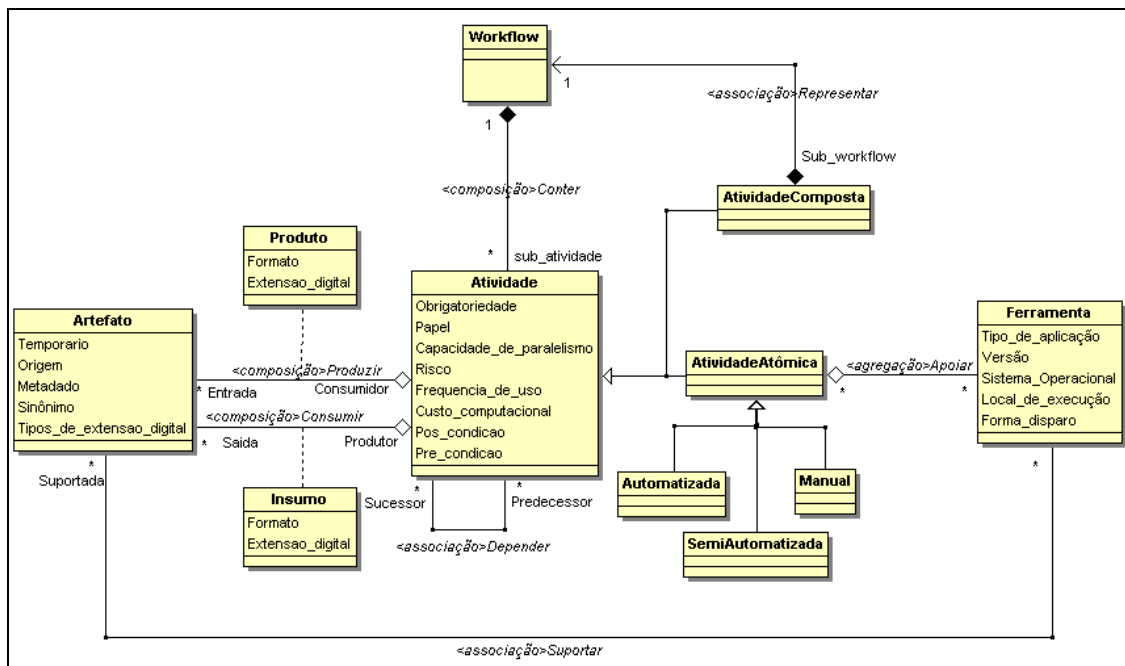


Figura E.1. Metamodelo estendido para experimentos científicos baseados em *workflow* científico.

O metamodelo é representado na notação de Diagrama de Classes, UML 2 (OMG, 2009). Na Figura E.1, um elemento do domínio de *workflow* científico será representado por uma classe. Para fins de representação, as características dos elementos de experimentação baseada em *workflow* científico foram representadas como atributos das classes. Na Figura E.1, além dos elementos do domínio de *workflow* científico, estão definidos os relacionamentos entre estes elementos, sendo representados através das associações do diagrama de classes. Todas as associações são relações de uso ou de composição de agregação entre os conceitos.

Durante a definição do metamodelo, o elemento *Sub-workflow* sofreu alteração para se adequar a nossa visão sobre o domínio de *workflow*. No entendimento desta abordagem, existia uma sobreposição com o elemento Atividade composta definido na extensão, pois, de fato, ambos são caracterizados por serem elementos contendo fluxos de atividades internas. Desta maneira, não houve necessidade de definir no metamodelo dois elementos com semântica equivalente. Assim, o conceito de *Sub-workflow* passou a ser considerado como um papel no relacionamento entre Atividade Composta e *Workflow*.

O metamodelo atende a uma característica dos modelos de *workflow* científico que é a capacidade de representar iterações e repetições. Em geral, os modelos de *workflows* científicos são classificados em: *DAG (Directed Aciclyc Graph)*, incapazes de representar iterações, e *non-DAG*, que são capazes. O metamodelo da Figura E.1 não restringe o uso de nenhum dos modelos, pois um modelo de *workflow DAG* pode

ser representado através de estruturas de non-*DAG*, desde que não apresente sequência de atividades contendo iterações. Além disso, essa característica em geral está associada às restrições da infra-estrutura computacional e do SGWfC, e não do experimento *in silico*.

A seguir são apresentados os detalhes dos elementos do metamodelo, explorando em especial Atividade, Ferramenta e Artefato. Também são discutidos os relacionamentos definidos, incluído as especializações, composições e associações.

Elemento Atividade

O elemento Atividade contém um conjunto de atributos como definidos na Figura E.1, que capturam características relacionadas, por exemplo, às condições de uso, à responsabilidade sobre execução, dentre outros. A seguir são explicados os seus atributos, sendo eles:

- **Obrigatoriedade:** O experimento científico pode ter variações no conjunto de atividades que serão executadas, porém algumas são obrigatórias. De fato, o experimento científico é um conjunto de diferentes fluxos de atividades, que variam entre si em alguns pontos específicos. Para exemplificar, considere que um pesquisador deseja refinar um resultado obtido numa rodada de execução do experimento científico, para avaliar um comportamento dos resultados e na rodada seguinte, o pesquisador não deseja executar esta atividade, pois ele já compreendeu o comportamento. Tendo em vista a possibilidade de escolha entre atividades a serem executadas, com o intuito de configurar o experimento científico, torna-se importante registrar a informação sobre a obrigatoriedade de uma atividade dentro do conjunto de atividades. Nesta abordagem foi definida uma classificação em **obrigatória** ou **opcional**. Durante a realização do experimento científico, se o pesquisador puder escolher não executar uma atividade, essa é caracterizada como opcional. Entretanto, se o pesquisador for obrigado a realizar a atividade, esta é caracterizada como obrigatória.
- **Pré-condição:** Uma atividade pode necessitar que uma situação, contexto ou condição seja verdadeira, para que seja permitido o seu início. A UML define o termo pré-condição como uma restrição que deve ser satisfeita quando a execução é iniciada (OMG, 2009). Com essa informação, durante a concepção do *workflow* concreto, regras ou meios para validar as condições precisam ser definidos e tratados, assim quando houver a execução desta atividade seja garantido que seu início só é efetuado se todas as pré-condições forem atendidas.

- **Pós-condição:** Uma atividade, para ter sucesso ao término de sua execução, pode necessitar que uma determinada situação, contexto ou condição seja verdadeiro. A UML define o termo pós-condição como uma restrição que deve ser satisfeita quando ao termino da execução (OMG, 2009). Da mesma forma que uma pré-condição, a informação capturada com o atributo pós-condição serve para que se criem regras ou meios para validação as condições necessárias ao término da execução da atividade.
- **Papel:** Uma atividade é realizada sob a responsabilidade de um indivíduo, que desempenha um papel no experimento. Por exemplo, considere uma atividade de análise de genoma defeituoso, um possível papel associado a essa atividade seria um geneticista. Destaca-se que um papel pode ser desempenhado por um ou mais indivíduos e um mesmo indivíduo pode desempenhar mais de um papel em todo o experimento científico. Uma atividade pode possuir mais de um papel necessário para sua execução, e neste caso existe um papel que é determinado como primário, ou seja, o responsável pela atividade, e os demais papéis são secundários, que são papéis de apoio a execução.
- **Capacidade de paralelismo:** Algumas atividades permitem que sua execução seja paralelizada e / ou distribuída, isto é, os passos são divididos e executados separadamente através de recursos computacionais, como por exemplo, ambientes em grade ou *clusters*. Essas atividades podem ser realizadas por diversos indivíduos ou recursos computacionais ao mesmo tempo. Quando há paralelismo, possibilita-se que a atividade tenha um aumento no seu desempenho. Entretanto, se a atividade não puder ser executada em paralelo ou de forma distribuída, então há uma restrição quanto a sua execução. Desta forma, torna-se importante que esse atributo capture quais são as situações nas quais a paralelização e/ou distribuição é realizada.
- **Risco:** A Atividade quando sujeita a fontes de perigo, possibilidade de perda ou infortúnio, está vulnerável a falhas. Essas situações são denominadas risco e estão associadas à execução do *workflow* científico e suas atividades. Por exemplo, algumas situações podem acarretar riscos ao experimento científico, como a entrada manual de dados, transformação de dados e entradas, ou fatores do ambiente onde o experimento científico executado. Com a captura das situações e dos riscos que podem ocorrer, o pesquisador ao criar seu *workflow* concreto deve considerar meios para mitigar sua ocorrência.

- **Freqüência de utilização:** A atividade pode ser utilizada mais de uma vez no mesmo experimento, pois o fluxo de execução pode conter decisões ou repetições que obrigam a sua re-execução diversas vezes. Nesta abordagem, isso se caracteriza como freqüência de utilização da atividade. A freqüência de utilização pode ser classificada como: alta, média ou baixa. Também pode ser especificado um valor numérico que representa a quantidade de vezes que a atividade é executada.
- **Custo computacional:** A atividade pode necessitar de apoio de recursos computacionais para que seja realizado. Quando isto acontece torna-se interessante registrar qual o seu custo computacional e o tempo de execução. Caso haja uma fórmula que calcule o tempo de execução, esta deve ser explicitada. Nesta abordagem, também se adota uma classificação qualitativa para o custo computacional: alta, média e baixa. Com o custo computacional das atividades registradas, pode-se estimar o tempo de execução do experimento.

Elemento Artefato, Produto e Insumo

Equivalente ao elemento atividade, também foram definidos atributos que representam características do elemento artefato. A seguir são apresentados os atributos definidos, sendo eles:

- **Origem:** Um artefato pode ser produzido internamente ao experimento, como consequência da execução de uma atividade interna, ou pode ser oriundo de atividades externas ao experimento científico. Na representação por *workflow* científico, um artefato é classificado como externo quando produzido fora do conjunto de atividades do *workflow*. Quando um artefato é produzido internamente é classificado como interno. Esta informação serve para caracterizar o conjunto de artefatos que foram criados durante a execução de um *workflow* científico, permitindo, por exemplo, determinar a responsabilidade de uma determinada informação utilizada.
- **Metadado:** Um artefato pode ser estruturado internamente e composto por um conjunto de informações. Quando isto acontece, é interessante que essa estrutura seja capturada, pois representa a forma como o conhecimento está organizado. Isto permite que ao se criar o *workflow* científico concreto, possa se determinar onde as informações estão presentes nos artefatos.
- **Sinônimo:** Um artefato pode possuir mais de um nome pelo qual é conhecido, quando isto ocorre é importante que seja registrado.

- **Temporário:** Um artefato pode possuir associado a ele um tempo de vida determinado ou finito. Existem artefatos que permanecem armazenados mesmo ao final da execução da atividade ou *workflow*. Quando isto ocorre são considerados permanentes. Entretanto, existem artefatos que só existem temporariamente e acabam descartados ou destruídos ao final da execução da atividade ou *workflow*.
- **Tipos de extensão digital:** Os tipos de extensão digital de um artefato quando produto/insumo de uma atividade são registrados neste atributo. Essa informação permite que posteriormente seja associado o produto/insumo às ferramentas que a suportam, caracterizando possíveis restrições no *workflow* científico concreto.

No metamodelo da Figura E.1, existem dois elementos que surgiram a partir da definição da associação entre elemento Artefato e elemento Atividades e seu uso durante a execução, sendo eles os elementos Produto e Insumo. De fato, esses elementos representam instâncias de Artefato para uma relação de uso com Atividade no modelo de *workflow* científico. A seguir, são descritos os atributos definidos para ambos os elementos:

- **Formato:** Durante o uso, um artefato instanciado pode estar no formato digital, físico ou digitalizável. Um produto/insumo é classificado como digital se é arquivo de computador ou dado computacional, em um formato específico. Um produto/insumo é digitalizável se é um documento em papel, que pode ser transformado em um arquivo de computador, por exemplo, através de um scanner. Um produto/insumo é físico se é uma amostra de material, por exemplo, rocha ou areia. Esta informação permite determinar se haverá algum esforço na transformação de uma instância do artefato para formato digital ou se ele é utilizado em meio físico. A transformação, isto é, a digitalização pode adicionar mais esforço à execução do *workflow* científico.
- **Extensão digital:** se o produto/insumo está no formato digital, este possui extensão digital associada. Pode haver mais de um possível formato para esse Produto / Insumo e esta relaciona-se à Instância do Artefato em uso numa Atividade.

Elemento Ferramenta

Os atributos definidos para o elemento Ferramenta os caracterizam nos modelos de *workflow* científico abstrato, mas também tem como propósito fornecer

informações a serem utilizadas na concepção do *workflow* científico concreto. As informações contidas nesses atributos são requisitos restritivos para o uso desta ferramenta e devem ser considerados na concepção do *workflow* concreto. Por exemplo, se ferramentas utilizadas em uma atividade só podem ser executadas localmente, há uma restrição do uso de tecnologias como *web services* (W3C, 2010) ou chamadas remotas, que deve ser levada em consideração na concepção do *workflow* concreto no SGWfC . Os atributos do elemento Ferramenta são descritos a seguir:

- **Tipo de aplicação:** Ferramentas possuem características próprias quanto ao seu uso, podendo ser: acessadas através de uma interface com usuários; ou acessada por outra aplicação ou sistema como um serviço; ou executada automaticamente através de comandos ou chamadas remotas. Assim, foi definido que ferramentas podem ser classificadas em três tipos básicos: *Batch* – aplicação que utiliza arquivos pra automatizar sua execução; *Serviço* – aplicação que provê funcionalidades para o uso por outra; *Interface* – aplicação na qual há uma interação com o usuário através de telas, formulários ou comandos.
- **Versão:** Ferramentas podem evoluir com o passar do tempo e por necessidade podem adicionar, modificar ou retirar funcionalidades, e conseqüentemente isso pode acarretar versões novas. Em alguns casos, uma atividade pode necessitar do apoio de uma versão específica do *software*. Outro exemplo é um artefato somente suportado por versão da ferramenta.
- **Sistema Operacional:** Ferramentas, em geral, são executadas sobre determinado sistema operacional. Esta informação é importante, pois pode constituir em restrição para execução do *workflow* científico do experimento e sua concepção.
- **Local de execução:** Ferramentas podem ser executadas em locais distintos. Foi definido que o local de execução deve ser classificado como: *Local* – se é executado no mesmo local que o *Workflow*; *Remota* – se é invocado e é executado em infra-estrutura diferente do *Workflow*; *Cluster* – se é executado em ambientes de alto desempenho; *Grade (Grid)* – se é executado em ambientes computacionais dispersos para alto desempenho.
- **Forma de disparo:** Ferramentas estão armazenadas em algum local e, em geral, podem ser invocadas por um caminho (*path*) ou método de invocação. O atributo forma de disparo tem como objetivo capturar a

maneira como a ferramenta é disparada/invocada durante a execução do *workflow*.

Extensão de elemento Atividade

No metamodelo, novos elementos foram criados como especializações do elemento Atividade. Na Figura E.2 está representada a hierarquia desses elementos, sendo o elemento Atividade o pai da hierarquia, e abaixo dele existem mais dois níveis de especializações, o 1º e 2º nível. O primeiro (1º) nível está relacionado à característica da atomicidade da atividade, ou seja, a possibilidade de conter sub-atividades encadeadas como fluxo interno de execução. Portanto, uma atividade pode ser “atividade atômica”, denominada *AtividadeAtômica* no metamodelo (Figura E.1), ou “atividade composta”, denominada *AtividadeComposta* no metamodelo.

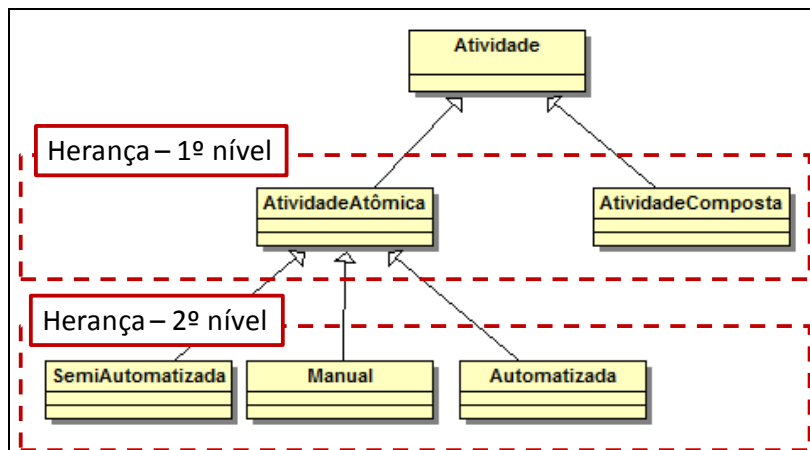


Figura E.2. Extrato do Metamodelo proposto – Herança do elemento Atividade.

O segundo (2º) nível hierárquico é a especialização do elemento atividade atômica, estando relacionado ao apoio computacional que esta necessita para ser realizada. As três especializações do elemento atividade atômica são “manual”, “semi-automatizada” e “automatizada”. Uma atividade é “manual”, se não há ou existe pouca utilização de ferramentas que auxiliem em sua execução, sendo essencial a participação do pesquisador/cientista. Uma atividade é “semi-automatizada”, se há necessidade de uma interação do pesquisador/cientista e de apoio de algum recurso computacional. Uma atividade é “automatizada”, se não requer uma participação essencial do pesquisador/cientista, pois há um recurso computacional que é responsável por apoiá-la completamente.

Relacionamento entre elementos no metamodelo

O metamodelo (Figura E.1) propõe um conjunto de relacionamentos (associações no diagrama de classes da UML 2) que representam interações entre os

elementos do domínio de experimentação baseada em *workflow* científico. Os relacionamentos propostos no metamodelo são descritos a seguir nesta seção.

Relacionamento 1: As atividades do experimento científico são ordenadas em seqüências de execução. Logo, é normal que uma atividade possa ter atividade predecessora e/ou atividade sucessora nesta seqüência. No metamodelo, para capturar esta relação entre atividades, foi definido um relacionamento do tipo auto-associação entre elemento Atividade. A Figura E.3 representa esse relacionamento, denominado como <associação>Depender. Nota-se que essa associação tem cardinalidade * em ambas as extremidades, isto significa que a atividade pode ter zero ou mais predecessoras e zero ou mais sucessoras.

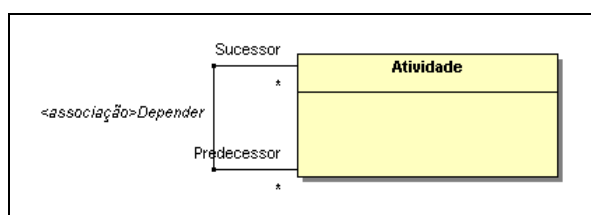


Figura E.3. Relacionamento de precedência e sucessão entre elemento Atividade – auto-associação.

Relacionamento 2: A definição de *workflow* científico propõe que internamente neste existe um conjunto de atividades encadeadas a formar um ou mais fluxos de execução, logo as atividades pertencem aquele *workflow* científico. No metamodelo definindo, essa relação entre o elemento *Workflow* e Atividade foi definida como um relacionamento (associação) do tipo composição, denominada <composição>Conter. Assim, uma atividade só existe do ponto de vista do experimento científico se está interligada ao *workflow* científico. A Figura A.4 apresenta como o relacionamento entre *Workflow* e Atividade foi representada no metamodelo. Percebe-se que dentro desta associação uma atividade tem um papel de sub-atividade para um *workflow*.

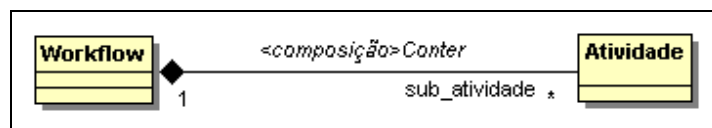


Figura E.4. Relacionamento de composição entre *Workflow* e Atividade.

Relacionamento 3: Durante sua execução, a atividade pode consumir, modificar ou gerar dados e/ou informações, sendo esses considerados aqui como artefatos. Quando um artefato é consumido no início da execução da atividade, é denominado insumo. Quando gerado durante a execução, denomina-se como produto.

No metamodelo, esses relacionamentos são explicitados como relações de agregações entre os elementos Atividade e Artefato, e classes associativas representando os elementos Produto e Insumo. A associação de agregação é justificada pelo fato que um artefato não está relacionado exclusivamente com uma única atividade, pois pode ser insumo de uma atividade X e produto de uma atividade Y.

A Figura A.5 é o extrato do metamodelo no qual estão representadas as duas relações (associações) que definem a interação entre Artefato e Atividade. Uma relação denominada <composição>Produzir, para representar a geração de um artefato por uma atividade, enquanto outra relação denominada <composição>Consumir para representar a utilização de um artefato por uma atividade. E duas classes associativas, Produto ligado a relação <composição>Produzir e Insumo ligado a relação <composição>Consumir.

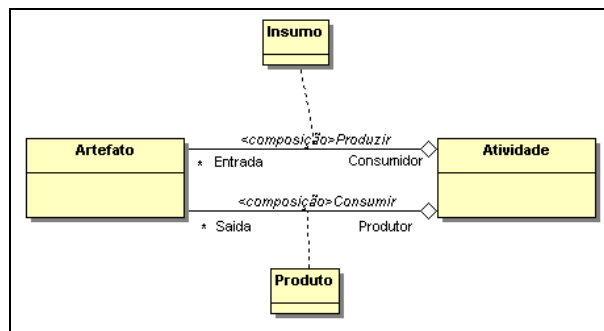


Figura E.5. Relacionamento de agregação entre elementos Atividade e Artefato.

Relacionamento 4: Uma atividade pode ser atômica ou composta. Neste metamodelo, quando a atividade é composta, ela representa um *workflow*. Esse relacionamento foi definido com base na característica em comum entre ambas, pois, assim como o *workflow*, a atividade composta contém atividades atômicas encadeadas representando um fluxo de execução. Entretanto, seria redundante se o metamodelo contivesse tanto elementos Atividade Composta e *Workflow* como composição de atividades. Portanto, foi definido que uma atividade composta, dentro do fluxo de execução de um *workflow*, tem o papel de *Sub-workflow*. A Figura A.6 é o extrato do metamodelo exibindo o relacionamento entre elementos *Workflow* e Atividade Composta. Este relacionamento é uma associação denominada como <associação>Representar. Essa associação significa que a instância de atividade composta associa-se unicamente a instância de *workflow*. De fato, somente é necessário que o elemento Atividade Composta tenha conhecimento do elemento

Workflow que ele representa, e não o contrário, uma vez que pode haver diversas atividades compostas distintas que representam o mesmo *workflow*.

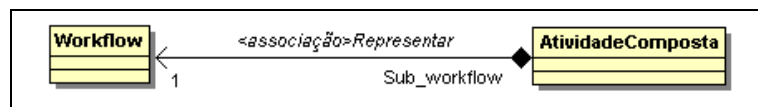


Figura E.6. Relacionamento unidirecional entre Atividade Composta e *Workflow*.

Relacionamento 5: Certas atividades têm suas execuções apoiadas por recursos computacionais, ou seja, o cientista ou pesquisador ao realizar uma atividade utiliza uma ou mais ferramentas. Então, pode se inferir que existe um relacionamento entre o elemento Atividade e Ferramenta. Entretanto, quando uma atividade é composta, relacionar ferramentas a essa atividade pode se tornar confuso e propenso a erros. Por exemplo, uma atividade composta representa um *workflow*, assim ela contém um conjunto de atividades encadeadas. Entretanto, esse conjunto é composto por atividades atômicas manuais, semi-automatizadas e automatizadas, isto é, algumas atividades podem não se relacionar a ferramentas ou podem se relacionar a diversas ferramentas diferentes. Logo, não há como relacionar uma única ferramenta a atividade composta (que representa o *workflow*). Além disso, listar todas as ferramentas na atividade composta é redundante, pois, essas informações já estão presentes nas atividades atômicas que a compõe (que representa o *workflow*).

A Figura E.7 é o extrato do metamodelo que representa esse relacionamento, nela há uma associação do tipo agregação, denominada <agregação> Apoiar, entre atividade atômica e ferramenta, no qual uma atividade atômica se relaciona com zero ou mais (*) ferramentas, enquanto uma ferramenta está relacionada com zero ou mais atividades. Essa cardinalidade se explica pelo fato de haver, comumente, a possibilidade de se escolher mais de uma ferramenta para apoiar a execução de uma atividade e ao mesmo tempo, uma ferramenta ser utilizada em diversas atividades distribuídas pelo experimento.

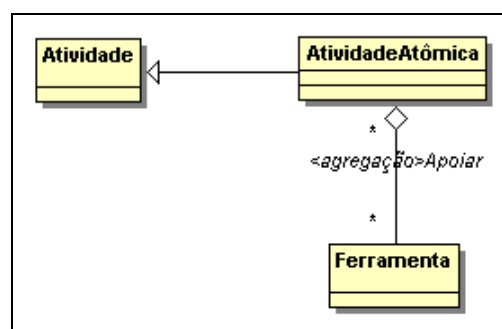


Figura E.7. Relacionamento entre elementos Atividade Atômica e Ferramenta.

Relacionamento 6: Devido à dependência de recursos computacionais nos experimentos *in virtuo* e *in silico*, os Artefatos podem ser manipulados através de Ferramentas durante a execução de atividades do *workflow*. Então, a fim de capturar esse relacionamento, foi definida no metamodelo uma associação entre Artefato e Ferramenta, denominado <associação> Suportar. O elemento Artefato associação tem o papel de “Suportada”, isto é, a ferramenta a suporta durante a execução. Essa associação tem cardinalidade * em ambas as extremidades, o que significa que os múltiplos Artefatos se relacionam com múltiplas Ferramentas, o que acontece na realidade é que um artefato pode ser lido por mais de uma ferramenta (por exemplo, arquivo “.doc”, lido por *Microsoft Word* e *BrOffice*) e uma ferramenta pode ler mais de um artefato. A Figura E.8 é o extrato do metamodelo representando a associação entre Artefato e Ferramenta.

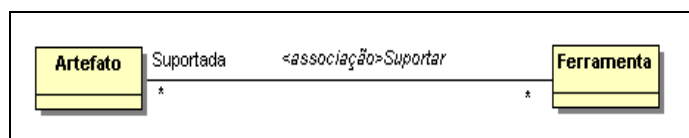


Figura E.8. Relacionamento entre elementos Artefato e Ferramenta.

Restrições sobre elementos da experimentação baseada em *workflow* científico

Após definir os elementos e os relacionamentos entre eles, tornou-se necessário definir algumas restrições para que o metamodelo possa representar corretamente o domínio de experimentação baseada em *workflow* científico. Essas restrições são regras que devem ser obedecidas. Então, para formalizar essas restrições no metamodelo, foi utilizada a linguagem OCL - *Object Constraint Language* (OMG, 2010a). Essa linguagem declarativa tem por objetivo descrever regras que se aplicam a modelos UML, tais como o diagrama de classes utilizado no metamodelo (Figura E.1). A seguir são apresentadas as restrições definidas em OCL.

Dependência cíclica de Artefatos na Atividade:

O relacionamento entre os elementos Atividade e Artefato possui uma restrição, pois, como foi representado no metamodelo, existe a possibilidade que um artefato seja ao mesmo tempo insumo e produto de uma atividade. Isto semanticamente não faz sentido, afinal, uma atividade não pode consumir uma informação que vai ser gerada ao final/durante sua execução. Além disso, espera-se que ao restringir esta dependência cíclica, não ocorram *deadlock* por esse motivo. Um *deadlock* é o travamento da execução do *workflow* por necessidade de insumos ou pela geração de produtos que nunca ocorrerá. A Figura E.9 apresenta um exemplo do

problema descrito, no caso “Atividade” possui dois insumos “Insumo1” e “Insumo2” e um produto “Produto”. Apesar do “Insumo1” ser oriundo de outra atividade do experimento, o “Insumo2” é oriundo da execução da própria “Atividade”, o que acarreta em deadlock na execução, pois nunca se iniciará.

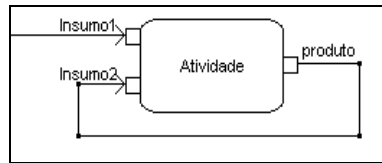


Figura E.9. Exemplo de problema evitado com a restrição dependência cíclica de Artefatos na Atividade.

Definição: “Uma instância da classe Atividade não pode ter em seu conjunto de insumos artefatos que foram produzidos por ela própria.”

Expressão OCL:

context Atividade **inv** DependenciaCiclicaArtefatos:

self.saida -> intersection (self.entrada) -> isEmpty()

Dependência procedência e sucessão entre atividades:

Uma atividade se relaciona com outras atividades, através da <associação>Depender, assim podendo conter tanto atividades predecessoras quanto sucessoras. Contudo, o metamodelo não restringe o caso no qual atividade se relaciona com ela própria, sendo predecessor e sucessor dela mesma. Quando isto ocorre, é caracterizado um problema no *workflow* científico que não deveria acontecer, já que a atividade pode ficar presa em *deadlock*, isto é, sempre ao final de uma execução, evocando-se novamente. A Figura E.10 apresenta um possível exemplo de problema do problema descrito. Neste exemplo tem-se um fluxo de execução, composto por duas atividades, “Atividade a” e “Atividade b”. Após a realização da “Atividade a”, a sua sucessora no modelo é “Atividade b”, contudo, ao se realizar essa atividade, a próxima atividade é a própria “Atividade b”. Modelagens desse levam a iterações executadas infinitamente, acarretando deadlock na execução.

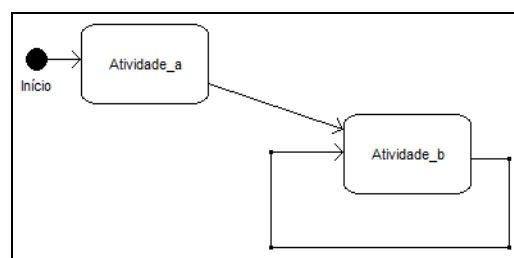


Figura E.10. Exemplo de problema evitado com a restrição dependência procedência e sucessão entre Atividades.

Definição: “Uma instância da classe atividade não pode ter como sucessor e predecessor ela mesma.”

Expressão OCL:

context Atividade **inv** DependenciaCiclicaAtividades:

self.sucessor -> intersection (self.predecessor) -> isEmpty()

Dependência cíclica entre diferentes níveis de workflow:

No metamodelo proposto é possível que um *workflow* científico seja composto por atividades atômicas e atividades compostas. A diferença entre uma atividade atômica e uma atividade composta está no fato que a atividade composta possui fluxo de execução interno, ou seja, é composta por outras atividades, enquanto a atividade atômica é indivisível. Na verdade, uma atividade composta dentro de um *workflow* é a representação de outro *workflow*, assumindo o papel de *Sub-workflow*. Esse mecanismo permite que o *workflow* científico seja partido em conjunto de atividades e hierarquizado em níveis. Entretanto, o metamodelo não restringe uma atividade atômica de preceder uma atividade composta que representa o *workflow*, cuja essa mesma atividade atômica faz parte. Essa relação se caracteriza como uma dependência cíclica entre diferentes níveis, sendo um possível problema no modelo.

A Figura E.11 apresenta um exemplo para o problema descrito, neste temos um “*Workflow*”, que é composto por: atividade atômica “Atividade a”; atividade atômica “Atividade b”; e atividade composta “AtividadeComposta”, que representa o próprio “*Workflow*”. Após a realização da “Atividade a”, duas outras atividades são realizadas “Atividade b” e “AtividadeComposta”, porém, como esse último é a representação de “*Workflow*”, todo fluxo de execução é reiniciado. Isto pode acarretar *loop* infinito, conseqüentemente, o *workflow* ficando em *deadlock*.

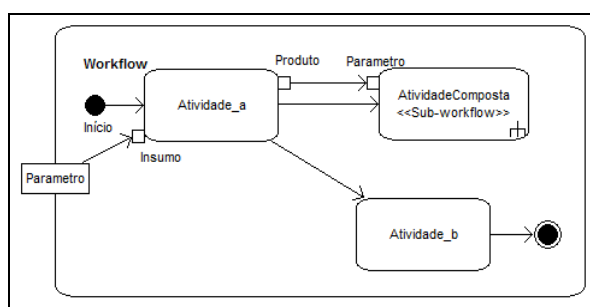


Figura E.11.. Exemplo de problema evitado com a restrição dependência cíclica entre diferentes níveis de *Workflow*.

Definição: “Uma instância da classe atividade não pode ter como sucessor uma atividade composta que representa um *workflow* que a contém.”

Expressão OCL:

context Atividade **inv** DependenciaHierarquica:

```
self.predecessor -> intersection (self.predecessor.workflow.sub_atividade) ->  
isEmpty()
```