



Programa Interdisciplinar de Pós-Graduação em

Computação Aplicada

Mestrado Acadêmico

Maurício de Oliveira Saraiva

COMMON: Uma Ferramenta para Composição
Semiautomática de Modelos de Processos de Negócio

São Leopoldo, 2016

Maurício de Oliveira Saraiva

**COMMON: Uma Ferramenta para Composição Semiautomática de Modelos de
Processos De Negócio**

Dissertação apresentada como requisito parcial
para a obtenção do título de Mestre, pelo
Programa Interdisciplinar de Pós-Graduação em
Computação Aplicada da Universidade do Vale
do Rio dos Sinos – UNISINOS

Orientador: Dr. Kleinner Silva Farias de Oliveira

São Leopoldo

2016

S243c Saraiva, Maurício de Oliveira
COMMON : uma ferramenta para composição semiautomática de modelos de processos de negócio / Maurício de Oliveira Saraiva. – 2016.

118 f. : il.

Dissertação (Mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, São Leopoldo, RS, 2016.

“Orientador: Dr. Kleinner Silva Farias de Oliveira”

1. Gerenciamento de processos de negócio. 2. Modelos de processos de negócio. 3. Composição de modelos de processos de negócio. 4. BPMN. I. Título.

CDU 004:658

Catalogação na Fonte:

Mariana Dornelles Vargas – CRB 10/2145

Maurício de Oliveira Saraiva

**COMMON: Uma Ferramenta para Composição Semiautomática de Modelos de
Processos de Negócio**

Dissertação apresentada à Universidade do Vale do Rio dos Sinos – Unisinos, como requisito parcial para obtenção do título de Mestre em Computação Aplicada.

Aprovado em 28 de julho de 2016

BANCA EXAMINADORA

Prof. Dr. Kleinner Silva Farias de Oliveira – UNISINOS

Prof. Dr. Toacy Cavalcante de Oliveira - UFRJ

Prof. Dr. Jorge Luis Victória Barbosa - UNISINOS

Prof. Dr. Kleinner Silva Farias de Oliveira (Orientador)

Visto e permitida a impressão

São Leopoldo, ____ de setembro de 2016

Prof. Dr. Sandro José Rigo
Coordenador PPG em Computação Aplicada

*Dedico este trabalho a todos que acreditam que
a construção do conhecimento se dá pela interação,
troca de saberes e compartilhamento de experiências.*

AGRADECIMENTOS

Agradeço a todos que de uma forma ou outra contribuíram para o desenvolvimento deste trabalho, entre eles, professores, colegas de mestrado, colegas de trabalho, esposa e, em especial, ao meu orientador Prof. Dr. Kleinner Silva Farias de Oliveira pela motivação, orientações e ensinamentos e à Universidade Corporativa dos Correios por possibilitar a minha participação no programa Bolsas de Estudo de Pós-Graduação – Eixo Estratégico.

RESUMO

A evolução da tecnologia tem permitido um suporte cada vez melhor à disciplina de gerenciamento de processos de negócio (BPM). A automação de práticas, a partir de ferramentas de modelagem, possibilita construir modelos visando apoiar as organizações no desenho e gerenciamento de seus processos. Em BPM, a transformação e melhoria contínua de processos são fundamentais para o gerenciamento da organização e os modelos de processos de negócio devem acompanhar essa evolução. Para realizar a composição desses modelos, seja pela transformação de um ou mais processos ou pela fusão de organizações, analistas usam as técnicas disponíveis na literatura. Porém, frequentemente elas são consideradas imprecisas ao não assegurar a correta integração dos modelos. Consequentemente, compor modelos de processos de negócio se torna uma tarefa difícil, custosa e propensa a erros, uma vez que as técnicas existentes criam modelos compostos difíceis de entendimento baseados em grafos e notação genérica que não correspondem à realidade do mercado de trabalho. Portanto, este trabalho propõe uma técnica de composição de modelos de processos de negócio que reúne, em uma única fonte, o conjunto das técnicas mais utilizadas na literatura, bem como dá suporte à composição de modelos direcionada para uma notação largamente conhecida e utilizada pelas organizações. Além disso, propõe a implementação de uma ferramenta, baseada nessa técnica, que foi criada para definir a similaridade e apoiar a integração de modelos de forma semiautomática. Por fim, estudos experimentais foram realizados com analistas que resolveram seis cenários de composição de modelos, tanto de forma manual como com o apoio dessa ferramenta, cujo resultado mostrou que a técnica proposta, de fato, melhora a precisão e reduz o esforço de composição, proporcionando redução de custos e aumento de produtividade.

Palavras-Chave: Gerenciamento de Processos de Negócio. Modelos de Processos de Negócio. Composição de Modelos de Processos de Negócio. BPMN.

ABSTRACT

The evolution of technology has allowed an increasingly better support to the discipline of business process management (BPM). The automation of practices from modeling tools, allows building models to support organizations in the design and management of its processes. In BPM, the transformation and continuous improvement of methods are fundamental to the management of the organization and models of business processes should follow this evolution. To perform the merge of these models, by the transformation of one or more processes, or by merging of organizations, software analysts use the techniques available in the literature. However, frequently they are considered imprecise, by not ensuring the correct integration of models. Consequently, merging business processes models becomes a difficult, error prone task, since existing techniques create composite models difficult to understand based on graphs and generic notation that does not match to the industrial reality. Therefore, this study proposes a merging technique of business process models that combines in a single source, some of the most used techniques available in the literature, and supports merging of models directed to a notation widely known and used by organizations. Furthermore, it also proposes the implementation of a tool, based on this technique, which was created to define the similarity and support the integration of models in a semi-automatic way. Lastly, experimental studies were performed with analysts who solved six scenarios of merging models, both manually and with the help of this tool. The result showed that the proposed technique, in fact, improves the accuracy and reduces composition effort, providing cost savings and increased productivity.

Keywords: Business Process Management. Business Process Models. Merging Business Process Models. BPMN.

LISTA DE FIGURAS

Figura 1: BPM e a conexão com os objetivos estratégicos	28
Figura 2: Ciclo de vida PDCA.....	28
Figura 3: Ilustração de modelagem BPMN	32
Figura 4: Ilustração de modelagem EPC	33
Figura 5: Ilustração de modelagem Petri Nets.....	34
Figura 6: Ilustração de modelagem <i>workflow</i>	35
Figura 7: Ilustração de modelagem IDEF.....	36
Figura 8: Ilustração de modelagem UML (diagrama de atividades)	37
Figura 9: Ilustração de modelagem VSM.....	38
Figura 10: Ilustração do conteúdo XML de um modelo BPMN	41
Figura 11: Dois modelos BPMN e seu respectivo mapeamento para o modelo composto.....	43
Figura 12: Composição Sequencial	47
Figura 13: Composição Paralela.....	47
Figura 14: Composição Condicional	48
Figura 15: Composição Iterativa	48
Figura 16: Conversão de modelo EPC em grafo de função	50
Figura 17: Integração de grafos de função	50
Figura 18: Conversão de grafo de função em EPC	51
Figura 19: Similaridade de Correspondência de Rótulos - métricas	52
Figura 20: Similaridade Estrutural – distância entre grafos	53
Figura 21: Similaridade Comportamental – vetor de espaços entre dois grafos	53
Figura 22: Composição de grafos configuráveis	55
Figura 23: INTPRO – Tela de composição de modelos EPC	56
Figura 24: Arquitetura da ferramenta BPMNDiffViz	57
Figura 25: Lista de correspondências entre dois modelos.....	58
Figura 26: Associação do fragmento F ₁ de modelo para matriz.....	59
Figura 27: Associação do fragmento F ₂ de modelo para matriz.....	59
Figura 28: INTPRO – Tela de composição de modelos EPC	60
Figura 29: Visão geral da técnica de composição	65
Figura 30: Comparação de elementos entre M _A e M _B	69
Figura 31: Comparação de elementos nos modelos M _A e M _B	70
Figura 32: Inclusão de elemento entre elementos	74
Figura 33: Inclusão de elemento entre elementos	74

Figura 34: Exclusão de elemento na integração	75
Figura 35: Processo de integração de modelos.....	77
Figura 36: Diagrama de componentes da ferramenta Common.....	79
Figura 37: <i>Tags XML</i> de modelos BPMN	81
Figura 38: Identificação de similaridades entre dois modelos BPMN	83
Figura 39: Funcionalidade – preparar a ferramenta.....	85
Figura 40: Funcionalidade – compor modelos	85
Figura 41: Exibição dos modelos M_A e M_B durante o processo de composição	87
Figura 42: Modelo M_C – resultado da integração de M_A e M_B	88
Figura 43: Diagrama de caso de uso dos experimentos 1 e 2.....	91
Figura 44: Experiência em desenvolvimento de software.....	95
Figura 45: Experiência em modelagem de software	96
Figura 46: Média de tempo de execução por participante.....	103
Figura 47: Média de tempo de execução por cenário.....	103
Figura 48: Taxa de erros por participante.....	105
Figura 49: Taxa de erros por cenário.....	105
Figura 50: Corretude por experimento	107

LISTA DE TABELAS

Tabela 1: Principais diferenças entre diagrama/mapa de processo e modelo de processos	29
Tabela 2: Principais características das notações mais conhecidas	31
Tabela 3: Elementos básicos BPMN	40
Tabela 4: Comparação dos trabalhos relacionados.....	62
Tabela 5: Resumo da similaridade entre elementos (VLS = 0,75).....	73
Tabela 6: Cenários do experimento 1	90
Tabela 7: Cenários do experimento 2.....	90
Tabela 8: Resumo das hipóteses.....	94
Tabela 9: Grau de escolaridade dos participantes	94
Tabela 10: Ocupação dos participantes	95
Tabela 11: Faixa etária dos participantes	95
Tabela 12: Caracterização dos modelos utilizados no experimento 1.....	97
Tabela 13: Resultados individuais por participante – experimento 1.....	98
Tabela 14: Resultados agrupados por cenário teste.....	99
Tabela 15: Caracterização dos modelos utilizados no experimento 2.....	100
Tabela 16: Resultados individuais por participante – experimento 2.....	101
Tabela 17: Resultados agrupados por cenário teste.....	102
Tabela 18: Resultado do esforço - teste Wilcoxon.....	104
Tabela 19: Resultado da taxa de erros - teste Wilcoxon.....	106
Tabela 20: Resultado da corretude - teste McNemar	107

LISTA DE ALGORITMOS

Algoritmo 1: Leitura e importação de modelo	81
Algoritmo 2: Identificação de similaridade e processo de composição	82
Algoritmo 3: Gravação do modelo composto.....	84

LISTA DE SIGLAS

ANSI	<i>American National Standards Institute</i>
API	<i>Application Programming Interface</i>
ARIS	<i>Architecture of Integrated Information Systems</i>
AES	<i>Asymmetric Event Structure</i>
BAM	<i>Business Activity Monitoring</i>
BPA	<i>Business Process Analysis</i>
BPDM	<i>Business Process Definition Metamodel</i>
BPEL	<i>Business Process Execution Language</i>
BPG	<i>Business Process Graph</i>
BPM	<i>Business Process Management</i>
BPMI	<i>Business Process Management Initiative</i>
BPMN	<i>Business Process Model and Notation</i>
BPMS	<i>Business Process Management Suite</i>
BRMS	<i>Business Rules Management Systems</i>
CORBA	<i>Common Object Request Broker Architecture</i>
CWM	<i>Common Warehouse Metamodel</i>
DFS	<i>Depth-First Search</i>
EA	<i>Enterprise Architecture</i>
EFS	<i>Escore Final de Similaridade</i>
EPC	<i>Event-Driven Process Chain</i>
EPML	<i>EPC Markup Language</i>
ER	<i>Enterprise Repository</i>
ERP	<i>Enterprise Resource Planning</i>
ESE	<i>Estratégia de Similaridade Semântica</i>
ESI	<i>Estratégia de Similaridade Sintática</i>
GPL	<i>General Public Licence</i>
IDEF	<i>Integrated Definition Language</i>
JVM	<i>Java Virtual Machine</i>
MDA	<i>Model Driven Architecture</i>
OMG	<i>Object Management Group</i>
PDCA	<i>Plan, Do, Check and Act</i>
PIPCA	Programa Interdisciplinar de Pós-Graduação em Computação Aplicada
PM	<i>Plano de Mudança</i>

PMT	<i>Process Mining Tools</i>
PN	Petri Nets
SOA	<i>Service-Oriented Architecture</i>
SVG	<i>Scalable Vectorial Graphics</i>
TI	Tecnologia da Informação
UML	<i>Unified Modeling Language</i>
UNISINOS	Universidade do Vale do Rio dos Sinos
VLS	Valor Limite de Similaridade
VSM	<i>Value Stream Mapping</i>
XML	<i>Extensible Markup Language</i>
XPDL	<i>XML Process Definition Language</i>

SUMÁRIO

1 INTRODUÇÃO	23
1.1 Problemática	24
1.2 Questões de pesquisa	24
1.3 Objetivos	25
1.4 Metodologia	25
1.5 Organização do trabalho	26
2 FUNDAMENTAÇÃO TEÓRICA.....	27
2.1 Gerenciamento de processos de negócio	27
2.2 Modelagem de processos de negócio	29
2.3 Notações de modelagem de processos de negócio	30
2.3.1 BPMN	31
2.3.2 EPC	32
2.3.3 Petri Nets.....	33
2.3.4 Workflow.....	34
2.3.5 IDEF	35
2.3.6 UML	36
2.3.7 VSM	37
2.4 Definição da notação de modelagem deste trabalho	38
2.5 Especificação BPMN.....	39
2.6 Composição de modelos de processos de negócio	41
3 TRABALHOS RELACIONADOS.....	45
3.1 Metodologia para a escolha dos trabalhos	45
3.2 Merging Worflows	46
3.3 Merging Event-driven Process Chains	49
3.4 Similarity of Business Process Models	51
3.5 Business Process Model Merging	53
3.6 Análise de Similaridade e Unificação de Processos de Negócio.....	55
3.7 BPMNDiffViz – A Tool for BPMN Models Comparison	57
3.8 An Automatized Approach for Merging Business Process Fragments	58
3.9 Comparação dos trabalhos relacionados.....	60
3.10 Oportunidades de pesquisa	62
4 TÉCNICA DE COMPOSIÇÃO DE MODELOS	65
4.1 Visão geral da técnica de composição	65
4.2 Definições de composição de modelos	66
4.2.1 Definição 1 – Conjunto de modelos	66
4.2.2 Definição 2 – Um modelo BPMN 2.0	67
4.2.3 Definição 3 – Modelo BPMN bem formado.....	67
4.2.4 Definição 4 – Comparação de elementos	68
4.3 Identificação de similaridades	68
4.3.1 Estratégia de Similaridade Semântica (ESE).....	69
4.3.2 Estratégia de Similaridade Sintática (ESI)	71
4.4 Integração de elementos	72
5 FERRAMENTA DE COMPOSIÇÃO DE MODELOS	77
5.1 Processo de integração de modelos.....	77
5.2 Arquitetura da ferramenta Common	78
5.3 Algoritmos da ferramenta Common	80
5.3.1 Leitura e interpretação de modelo	80
5.3.2 Identificação de similaridade e processo de composição.....	81
5.3.3 Exportação do modelo composto.....	84
5.4 Apresentação da ferramenta Common	84
5.4.1 Funcionalidades	84
5.4.2 Interface da ferramenta Common	85
6 AVALIAÇÃO DA TÉCNICA PROPOSTA	89

6.1 Método de pesquisa.....	89
6.2 Formulação das hipóteses	92
6.3 Perfil dos participantes.....	94
6.4 Experimento 1 – Composição de modelos sem uso de ferramenta de apoio	96
6.4.1 Tabulação dos resultados do experimento 1	98
6.5 Experimento 2 – Composição de modelos com uso da ferramenta Common	99
6.5.1 Tabulação dos resultados do experimento 2	100
6.6 Análise comparativa dos experimentos	102
6.6.1 QP-1: avaliação do esforço de composição.....	102
6.6.2 QP-2: avaliação da taxa de erros.....	104
6.6.3 QP-3: avaliação da corretude.....	106
6.6.4 Avaliação final	107
6.7 Ameaças à validade do estudo.....	108
7 CONCLUSÃO	111
7.1 Contribuições	111
7.2 Limitações e trabalhos futuros	112

1 INTRODUÇÃO

No contexto da administração de organizações, o gerenciamento de processos de negócio (BPM) tem se tornado cada vez mais frequente. Esse gerenciamento visa identificar, desenhar, documentar e controlar os processos visando alcançar os objetivos estratégicos de cada organização, através da transformação contínua de processos que vai além da melhoria, uma vez que essa transformação implica em repensar, inovar e atuar de diferentes maneiras, de acordo com a tendência e o rumo do mercado (DELGADO et al., 2013).

Entretanto, esse gerenciamento necessita de ferramentas que apoiem a execução e manutenção desses processos para que as atividades sejam executadas de forma mais ágil e eficaz. Nesse sentido, a Tecnologia da Informação (TI) atua como facilitador, pois fornece uma infraestrutura que, aliada ao negócio, permite automatizar processos visando entregar melhores produtos e serviços para os clientes (BPM-CBOK, 2013).

Entre as atividades envolvidas na automação de processos pela TI, a modelagem BPM ocupa um importante papel, uma vez que compreende a representação dos processos de maneira completa e precisa, especificando seu funcionamento, interfaces, componentes e responsabilidades através de ícones universalmente conhecidos que formam os fluxos de um processo (GENON et al., 2010). Dessa forma, a modelagem proporciona uma visão dos processos que é entendível tanto pelos gestores como pelos arquitetos e analistas responsáveis pelo desenvolvimento de software na empresa (BPM-CBOK, 2013).

Em se tratando de modelagem BPM, é comum encontrar centenas ou até milhares de modelos de processos de negócio nas organizações (DIJKMAN et al., 2011). Muitos desses modelos são diferentes versões de um mesmo processo ou fragmentos de processos que servem de apoio para a construção de outros processos maiores (YAN et al., 2010; ZEMNI et al. 2016). Em virtude disso, torna-se importante a questão de composição de modelos (LA ROSA et al., 2010), seja para implantar um novo processo a partir de dois ou mais processos existentes, no caso de reestruturação de setores, ou em fusões de companhias, quando duas companhias se unem para formar uma terceira, visando crescimento e redução de custos com a eliminação de setores duplicados (BJURSELL, 2011; BARROS, 2001).

Na prática, analistas fazem uso de ferramentas que utilizam notações conhecidas tanto na academia como na indústria para desenhar modelos de processos de negócio como, por exemplo, a *Business Process Model and Notation* (BPMN), que atualmente é vista como padrão de mercado para modelagem BPM, em virtude de seu vasto uso (PILLAT, 2014). Contudo, apesar de existir diversas ferramentas de modelagem BPMN, tais como, *Eclipse BPMN2 Modeler ECLIPSE* (2015), *IBM Business Process Manager IBM* (2015), *Oracle Business Process Manager ORACLE* (2015), *Microsoft BPMN Visio Modeler VISIO* (2015) e *Bizagi Modeler BIZAGI* (2015), não se conhece alguma que seja capaz de auxiliar a execução de composição de modelos de processos de negócio. Consequentemente, analistas tendem a realizar a integração de modelos de processos de negócio de forma manual, sendo essa prática amplamente reconhecida pela literatura como difícil, onerosa e propensa a erros (LA ROSA et al., 2013; BRANCO, 2014).

De fato, a literatura atual reforça que as técnicas de composição de modelos de processos de negócio não são efetivas. Em La Rosa et al., 2013, os autores destacam que as técnicas atuais não proporcionam ganho de produtividade aos analistas por gerarem modelos compostos difíceis de entender. Em Zemni et al. (2016) e Dijkman et al. (2011), os autores destacam também que as técnicas tipicamente tratam os modelos como grafos, ou mesmo tratam os modelos de processos de negócio que são ricos semanticamente utilizando notação genérica.

Logo, as técnicas acabam não garantindo a eficácia das integrações. Em outro trabalho estudado, Gottschalk et al. (2008) reforçam também que as técnicas do estado da arte aplicam notações pouco utilizadas atualmente, não atendendo às necessidades das empresas que trabalham com uma notação largamente difundida como a BPMN.

Além disso, as ferramentas e algoritmos propostos pela academia realizam a integração de modelos de forma automática, compreendendo o comportamento de ambos os modelos de entrada como, por exemplo, nos trabalhos de Zemni et al. (2016), Santos (2014) e La Rosa et al. (2013). Essa característica vai de encontro às práticas das organizações, uma vez que não é possível identificar automaticamente quais elementos devem pertencer ao modelo composto, uma vez que o modelo produzido nem sempre deve conter todos os elementos de ambos os modelos de origem.

Portanto, diante do contexto da ausência de ferramentas de modelagem que realizem composição de modelos e da deficiência das técnicas de composição disponíveis na literatura, visualiza-se uma lacuna na área de composição de modelos de processos de negócio, em especial com o uso da notação BPMN 2.0.

1.1 Problemática

Conforme descrito anteriormente, a literatura atual na área de composição de modelos de processos de negócio reporta que as técnicas disponíveis na academia não são efetivas. Isto é, não produzem ganho de produtividade ao reduzir o esforço dos analistas nem diminuem a taxa de erros no processo de composição. Além disso, as ferramentas de modelagem BPMN disponíveis no mercado não dispõem de mecanismos de composição de modelos, o que torna a atividade difícil, onerosa e propensa a erros. Nesse sentido, listam-se a seguir as problemáticas que são abordadas neste trabalho:

- Os mecanismos citados nos estudos não proporcionam ganho de produtividade ao reduzir o esforço dos analistas, uma vez que geram modelos compostos de difícil entendimento, baseados em grafos e notações genéricas, que não correspondem a uma notação largamente utilizada atualmente no mercado.
- Os mecanismos que implementam técnicas contemporâneas de composição de modelos realizam a integração de maneira automática, incorporando o comportamento de ambos os modelos de entrada. Consequentemente, isso aumenta a propensão a erros, uma vez que esses mecanismos não permitem a interação do analista nem possuem uma representação semântica que proporcione produzir o modelo composto corretamente, baseado nas regras de negócio.
- As ferramentas de modelagem BPMN disponíveis no mercado não implementam uma funcionalidade de composição de modelos que apoie analistas na execução dessa atividade tornando, assim, a composição uma tarefa difícil, onerosa e propensa a erros.

1.2 Questões de pesquisa

Sabendo-se da limitação dos mecanismos e técnicas contemporâneas propostas na academia e, também, da ausência de ferramentas corporativas, disponíveis no mercado, que auxiliem na composição de modelos com uso da notação BPMN 2.0, identifica-se uma lacuna de pesquisa. Dessa forma, tomando por base a proposta de uma ferramenta de composição de modelos que realize a integração de forma semiautomática, com base nas melhores técnicas disponíveis na academia, surgem as seguintes questões de pesquisa deste trabalho:

- QP-1. O esforço para integrar modelos usando uma técnica semiautomática é menor que usando uma técnica manual?
- QP-2. A taxa de erros encontrada nos modelos integrados de forma semiautomática é menor que nos modelos integrados de forma manual?
- QP-3. O número de modelos integrados corretamente é maior usando uma técnica semiautomática que usando uma técnica manual?

1.3 Objetivos

O objetivo geral desta pesquisa é *propor uma técnica semiautomática que facilite a composição de modelos de processo de negócio, visando reduzir o esforço e a propensão a erros.* Objetivos específicos são:

- **Propor uma técnica de composição de modelos de processos de negócio.** Este objetivo se concentra em propor uma técnica de composição baseada na compilação das melhorias técnicas já documentadas na literatura, que compreendem atividades de identificação de similaridades entre os elementos dos modelos e integração dos mesmos em um modelo composto.
- **Propor uma ferramenta que implemente a técnica de composição.** Este objetivo visa pôr em prática a técnica proposta neste trabalho com o intuito de carregar, persistir, comparar, integrar e visualizar modelos de processos de negócio. Ao criar essa ferramenta, espera-se facilitar o trabalho de analistas que precisam compor modelos, visando reduzir o esforço e a propensão a erros, proporcionando economia de tempo e de custos.
- **Avaliar a técnica proposta.** Este objetivo visa realizar experimentos controlados de composição de modelos de processos de negócio para validar a eficácia da técnica e da ferramenta propostas. Esses experimentos compreendem a execução de cenários de teste realizados tanto de forma manual como com o apoio da ferramenta proposta.
- **Producir conhecimento empírico sobre composição de modelos de processos de negócio.** Este objetivo visa produzir conhecimento empírico sobre como os analistas investem esforço para compor modelos, bem como mensurar os erros produzidos nesses modelos, a partir dos experimentos que serão realizados.

1.4 Metodologia

Para desenvolver este trabalho, o primeiro passo consiste na realização de uma pesquisa a partir da revisão da literatura sobre identificação de similaridades e integração de modelos de processos de negócio. Essa revisão tem por objetivo analisar as técnicas disponíveis na academia, bem como identificar suas limitações com o intuito de elencar novas oportunidades de pesquisa. Segundo Marconi e Lakatos (2003), essa pesquisa é considerada exploratória quanto aos seus objetivos, uma vez que busca propiciar maior familiaridade com o tema, além de aprofundar seus conceitos preliminares.

O segundo passo consiste em estudos empíricos de integração de modelos de processos de negócio de forma manual. Esses estudos visam compreender os fatores que afetam a integração manual de modelos a partir da realização de experimentos com o intuito de mensurar

o esforço que analistas investem para integrá-los, assim como verificar a ocorrência de inconsistências nos modelos produzidos.

O terceiro passo consiste na apresentação de uma técnica de identificação de similaridades e integração de modelos de processos de negócio, baseada nas melhores técnicas e métricas disponíveis na literatura, bem como o desenvolvimento de uma ferramenta que implemente a técnica proposta.

Por fim, o último passo compreende a realização de novos experimentos de composição de modelos de processos de negócio com o apoio da ferramenta desenvolvida, visando analisar dados comparativos de esforço, taxa de erros e corretude entre a composição manual e a composição semiautomática de modelos. Segundo Gil (2002), esses experimentos caracterizam a pesquisa como quantitativa quanto ao método, visto que apresentam comparação de medidas e uso de técnicas estatísticas para avaliação dos resultados.

1.5 Organização do trabalho

Este trabalho está organizado em sete capítulos, conforme descrito a seguir. O Capítulo 2 relaciona os conceitos fundamentais sobre processos de negócio. O Capítulo 3 apresenta as principais abordagens que possuem relação com a proposta desta pesquisa, relacionando suas principais características, e descreve as oportunidades de pesquisa. O Capítulo 4 apresenta a técnica de composição proposta neste trabalho. O Capítulo 5 introduz a ferramenta de composição de modelos que implementa a técnica descrita no Capítulo 4. O Capítulo 6 descreve a avaliação da técnica e da ferramenta propostas. Por fim, a conclusão deste trabalho é apresentada no Capítulo 7, bem como as contribuições, as limitações e sugestões de trabalhos futuros desta pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Este Capítulo apresenta os conceitos fundamentais para o entendimento da proposta deste trabalho. A Seção 2.1 apresenta uma introdução sobre gerenciamento de processos de negócio, na qual destaca a importância do BPM para a gestão de processos nas organizações. Em seguida, a Seção 2.2 introduz os conceitos de modelagem de processos de negócio. A Seção 2.3 descreve as principais notações de modelagem de processos de negócio disponíveis no mercado. A Seção 2.4 justifica o uso da notação BPMN 2.0, utilizada neste trabalho. Após, os principais recursos e características da notação BPMN são apresentados na Seção 2.5. Por fim, a Seção 2.6 relaciona os conceitos de composição de modelos de processos de negócio.

2.1 Gerenciamento de processos de negócio

Agregar valor ao cliente é o objetivo principal da disciplina de gerenciamento de processos de negócio. Através do BPM, novas práticas de negócio são criadas e/ou transformadas, de forma a produzir processos mais eficazes e ágeis, para proporcionar maior retorno aos interessados (KO et al., 2009).

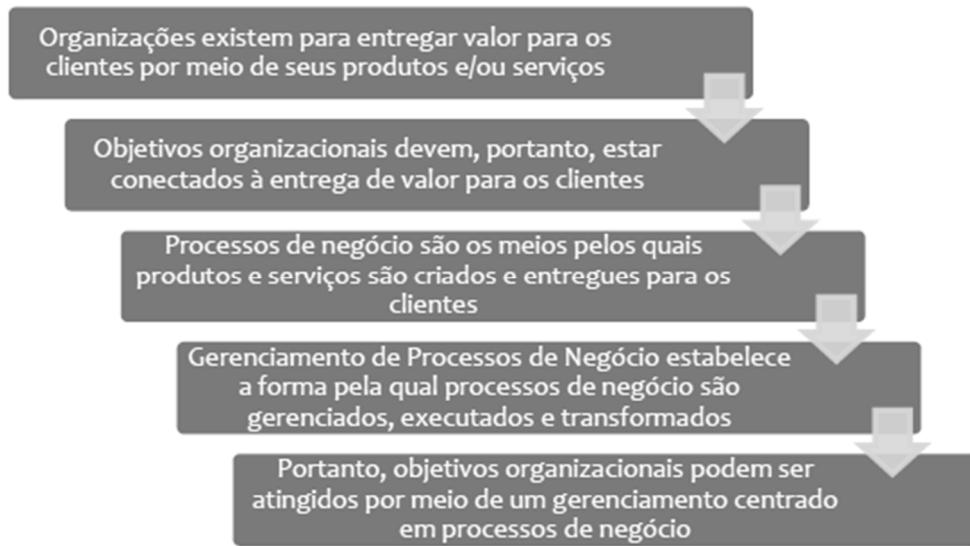
Não importa se a organização tem ou não fins lucrativos, seja pública ou privada, de micro, pequeno ou grande porte, o propósito principal de uma organização é gerar valor para o cliente por meio de seus produtos e/ou serviços. Esse é princípio que deveria direcionar todos os objetivos organizacionais (BPM-CBOK, 2013, p. 45).

Para Baltzan e Phillips (2012), é função do BPM integrar estratégias e objetivos da organização com as necessidades dos clientes, com foco nos processos ponta a ponta e a orquestração das atividades ao longo das funções de negócio. Esse gerenciamento engloba estratégias, objetivos, cultura, estruturas organizacionais, papéis, políticas, métodos e tecnologias para analisar, desenhar, implementar, transformar e estabelecer governança de processos.

Nesse sentido, o BPM atua como auxílio para estabelecer um conjunto de princípios e práticas que visam tornar as organizações mais eficientes e eficazes na execução de seus processos, através do melhor gerenciamento de recursos, sem, contudo, estabelecer estruturas de trabalho, metodologias ou ferramentas específicas, na qual cada organização deverá internalizar e implementar ações conforme sua realidade e necessidade. A Figura 1 descreve a conexão entre o BPM e os objetivos estratégicos da organização.

O BPM se faz valer das melhores práticas de gestão de processos de negócio. Porém, essas práticas exigem o comprometimento permanente e contínuo da organização, uma vez que um ciclo de atividades deve ser constantemente realizado, visando a melhoria sucessiva (BALDAN et al., 2014). Um exemplo de ciclo de vida típico é o PDCA (*Plan, Do, Check e Act*), em que as atividades são planejadas, executadas, verificadas e refinadas continuamente, conforme ilustrado na Figura 2.

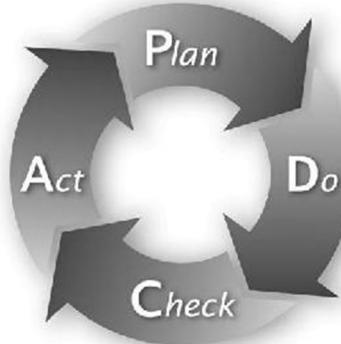
Figura 1: BPM e a conexão com os objetivos estratégicos



Fonte: BPM-CBOK (2013, p. 46)

A tecnologia realiza o papel de suporte ao BPM, contudo, não pode ser tratada como ponto principal (papel de liderança), uma vez que a automação é um processo desejável. Assim, o foco deve ser a aprimoração e a transformação contínua dos processos, visando alcançar os resultados esperados, através da otimização de recursos, da redução de custos, defeitos e desperdícios e do aumento da rentabilidade, produtividade, competitividade e satisfação do cliente (BPM-CBOK, 2013).

Figura 2: Ciclo de vida PDCA



Fonte: <http://www.luisfigaro.com/techblog/wp-content/uploads/2012/10/PDCA2.jpg>

Entretanto, apesar de não ser o foco do BPM, a tecnologia fornece uma infraestrutura importante ao gerenciamento de processos, pois provê suporte em diversas atividades, as quais incluem, mas não se limitam a: planejar, desenhar, modelar, documentar, analisar, executar e monitorar processos de negócio; e acompanhar, gerenciar e controlar mudanças de processos de negócio (HITPASS, 2014).

2.2 Modelagem de processos de negócio

Campos (2014) e Guerrini et al. (2014) citam que, no contexto do BPM, a modelagem de processos de negócio é uma atividade de fundamental importância para o gerenciamento de uma organização, uma vez que permite criar uma representação completa e precisa sobre o funcionamento de determinados processos. Essa representação busca aumentar a compreensão do negócio, facilitar a comunicação e identificar pontos de melhoria e oportunidades de crescimento, entre outras.

A modelagem de processos de negócio é uma representação abstrata de um negócio, normalmente apresentado em formato de diagramas, mapas ou modelos propriamente ditos. Possui diversas representações, que podem ser suficientemente simples em alguns casos e necessariamente completa em outros, conforme o nível de detalhamento e a necessidade do negócio, podendo assim servir para diferentes propósitos. BPM-CBOK (2013) conceitua diagrama, mapa e modelo conforme segue:

Um diagrama é uma representação mais enxuta de um processo de negócio, uma vez que ilustra o fluxo de trabalho de forma mais simplificada, sem detalhes, através da demonstração de aspectos gerais. Entretanto, mesmo sem exibir detalhes, pode ajudar na identificação e na compreensão das principais atividades de um processo.

Um mapa, por sua vez, possui uma representação mais detalhada que um diagrama, de modo a ilustrar os principais componentes de um processo com maior nível de precisão. Dessa forma, permite apresentar alguns elementos mais importantes como, por exemplo, atores, eventos e relacionamentos.

Um modelo de processos de negócio é mais completo que diagramas e mapas, e é apropriado para qualquer nível de captura de processos. Apresenta atores, atividades, eventos, mensagens, condições, artefatos e outros elementos dos processos que são representados por ícones de modo que seja possível compreender o fluxo do processo, os relacionamentos, as responsabilidades e os artefatos que são produzidos em determinados negócios. Seu principal objetivo compreende o auxílio na documentação, análise e desenho de processos de negócio. A Tabela 1 apresenta as principais diferenças entre diagrama/mapa de processos e modelo de processos.

Tabela 1: Principais diferenças entre diagrama/mapa de processo e modelo de processos

Diagrama ou mapa de processo	Modelo de processos
Notação ambígua	Convenção padronizada da notação
Baixa precisão	Não precisa quanto necessária
Menos detalhado	Mais detalhado
Ícones inventados ou vagamente definidos	Ícones definidos e padronizados
Limitado a retratar um momento específico da realidade	Pode crescer, evoluir e amadurecer
Pode ser criado com ferramentas simples de diagramação	Deve ser criado com ferramenta adequada ao objetivo
Difícil de conectar com outros modelos existentes	Possui ligações que mostram relacionamentos entre processos e diferentes níveis de processo
Apropriado para certas capturas rápidas de ideias	Apropriado para qualquer nível de captura de processos, análise e desenho

Fonte: Adaptado de BPM-CBOK (2013, p. 74)

A representação de um modelo ilustra o estado de um negócio, podendo ser o estado atual (AS-IS) ou um estado futuro (TO-BE), e os recursos que estão envolvidos no processo como, por exemplo, pessoas, entradas, saídas, insumos, informações, regras, etc. Ou seja, permite representar os processos modelados da forma tão precisa quanto necessária (KO et al., 2009).

Ohlsson et al. (2014) explicam que o estado AS-IS de um processo define a situação atual em que ele se encontra, assim, proporciona a compreensão de como o negócio está sendo realizado na organização. Já o estado TO-BE faz referência às mudanças que serão realizadas nos processos, visam melhorar a qualidade e a produtividade e eliminar problemas e defeitos, através do redesenho do processo. Assim, nesse estado devem ser eliminadas as tarefas que não são mais necessárias e incluir aquelas que poderão agregar mais valor ao negócio.

A modelagem de processos pode ser realizada por qualquer tipo de ferramenta ou notação, desde que atenda aos requisitos necessários do processo para o qual foi selecionada (FERNANDES; ABREU, 2014). Como nenhuma técnica é exclusiva às demais e por serem diferentes - apresentando vantagens e desvantagens conforme a circunstância do processo, todas podem ser utilizadas, uma vez que a disciplina de BPM não estabelece um padrão específico para a modelagem de processos de negócio.

2.3 Notações de modelagem de processos de negócio

Existem diversos padrões de notação de modelagem de processos de negócio disponíveis no mercado. Cada notação foi criada com um objetivo e visa atender determinadas necessidades organizacionais. Segundo Campos (2014, p. 41), notação significa: "...um sistema definido com padrões e regras de tal modo que seja possível registrar e expressar ideias de um determinado domínio de conhecimento".

Entre as notações mais conhecidas, destacam-se: *Business Process Model and Notation* (BPMN) OMG-BPMN (2013), *Workflow* (WF) Van der Aalst (2004), *Event-driven Process Chain* (EPC) Gottschalk (2008), Petri Nets (PN) Sun et al. (2006), *Unified Modeling Language* (UML) OMG-UML (2015), *Integrated Definition Language* (IDEF) Menzel (2006), *Value Stream Mapping* (VSM) Dal Forno et al. (2014), cujas principais características encontram-se na Tabela 2.

A seguir descrevem-se as principais características, bem como as vantagens, desvantagens e quando seu uso é aconselhável, além de um pequeno exemplo de cada uma das principais notações existentes no mercado.

Tabela 2: Principais características das notações mais conhecidas

Notação	Descrição
BPMN	Padrão criado pela <i>Object Management Group</i> (OMG), útil para apresentar um modelo para públicos-alvo diferentes como, por exemplo, desenvolvedores, arquitetos de software, gerentes e diretores.
EPC	Considera eventos como “gatilhos para” ou “resultados de” uma etapa do processo; útil para modelar conjuntos complexos de processos.
PN	Petri Nets, também conhecido como Redes de Petri, é uma alternativa para modelar processos de avaliação de comportamento e desempenho de sistemas computacionais, especialmente de manufatura.
WF	Originalmente aprovado como padrão ANSI, inclui um conjunto simples e limitado de símbolos não padronizados; facilita o entendimento rápido do fluxo de um processo.
IDEF	Padrão da <i>Federal Information Processing Standard</i> dos EUA que destaca entradas, saídas, mecanismos, controles de processo e relação dos níveis de detalhe do processo superior e inferior; ponto de partida para uma visão corporativa da organização.
UML	Mantida pela OMG, consiste em um conjunto-padrão de notações técnicas de diagramação orientado à descrição de requisitos de sistemas de informação.
VSM	Consiste em um conjunto intuitivo de símbolos que é usado para mostrar a eficiência de processos por meio do mapeamento de uso de recursos de elementos de tempo.

Fonte: Adaptado de BPM-CBOK (2013, p. 79)

2.3.1 BPMN

Business Process Model and Notation é uma notação de padrão aberto mantida pela OMG – organização que estabelece padrões para sistemas de informação. Apresenta um conjunto robusto de símbolos, cuja representação permite a modelagem de diferentes aspectos dos processos de negócio. Possui vasta aceitação, tanto no mercado como na academia, sendo que está presente em uma grande quantidade de ferramentas de modelagem (WHITE; MIERS, 2008).

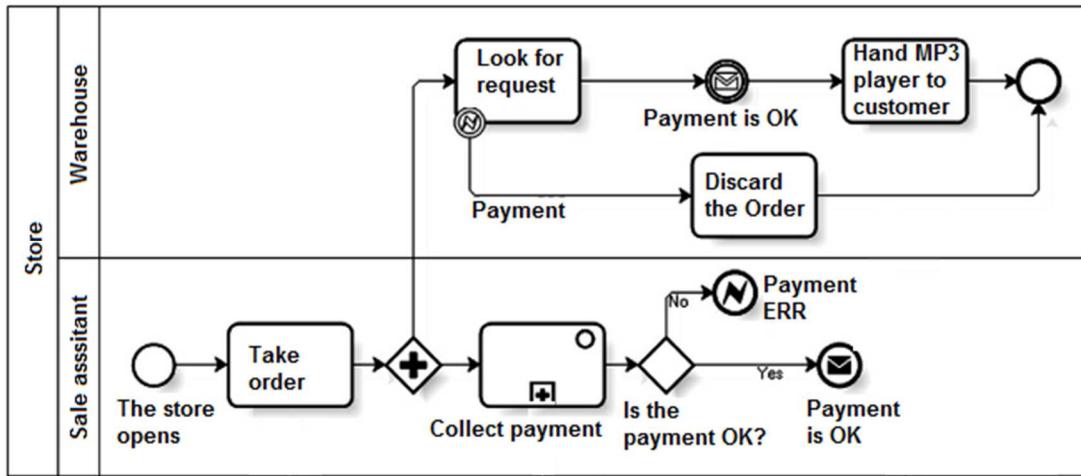
Através de sua representação padronizada, permite modelar desde processos simples até processos mais complexos, facilmente entendível desde gestores ocupantes de altos cargos de direção até analistas de processos responsáveis pela modelagem, haja vista que os símbolos descrevem eventos, mensagens, relacionamentos, responsabilidades, decisões e fluxos de atividades de forma clara e eficaz (OMG-BPMN, 2013).

Segundo Mili et al. (2010), seu uso é aconselhável para a apresentação de um modelo para públicos-alvo diferentes, devido sua facilidade de entendimento dispensar conhecimentos técnicos específicos, além de poder gerar aplicações em *Business Process Model System* (BPMS), bem como possibilitar a simulação de processos de negócio com um motor de processo.

As vantagens se caracterizam por sua boa usabilidade para modelar processos de negócio através de fácil e rápido aprendizado, por suportar ferramentas BPMS e por ter seu uso e entendimento difundido em muitas organizações. Entretanto, como desvantagens, destacam-se a dificuldade de integração com outras ferramentas devido ao fato de ser uma notação gráfica que não é destinada ao manuseio de diferentes visões, visto que atribui foco apenas ao processo (BPM-CBOK, 2013).

A Figura 3 ilustra um exemplo de modelo de processo de negócio modelado com a notação BPMN.

Figura 3: Ilustração de modelagem BPMN



Fonte: Adaptado de Chinosi e Trombetta (2012, p. 129)

2.3.2 EPC

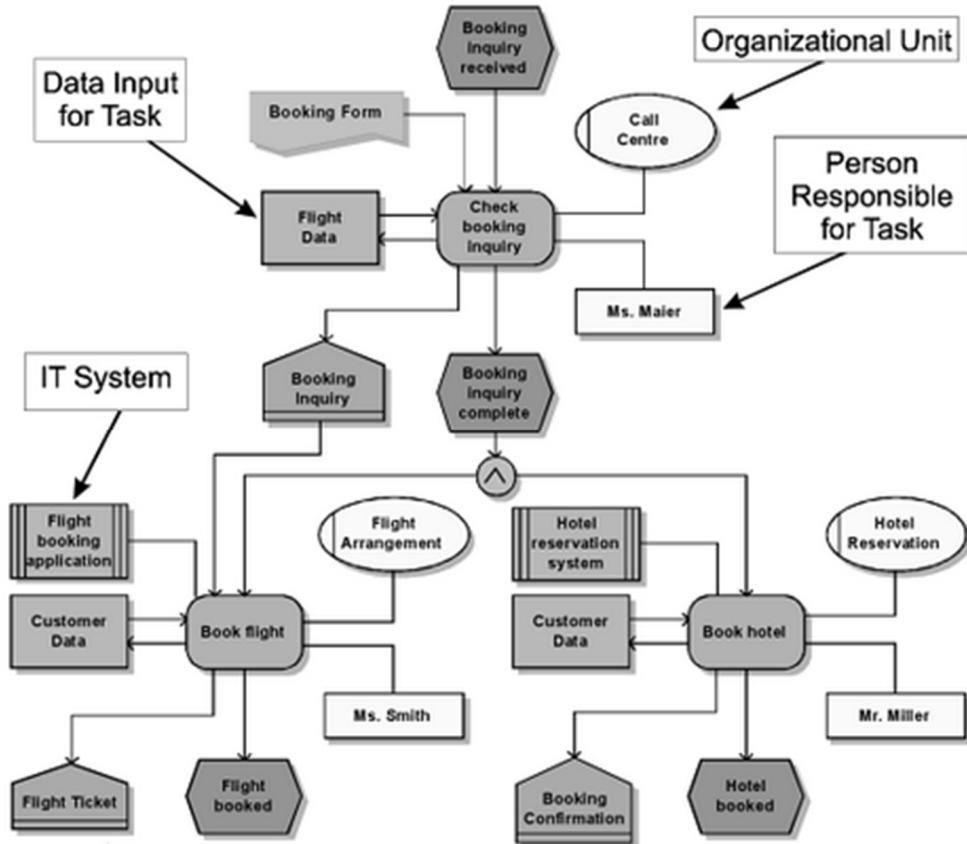
Event-driven Process Chain é uma notação que pertence ao framework *Architecture of Integrated Information Systems* (ARIS). Possui um conjunto de símbolos para modelagem de processos de negócio que descreve eventos de uma etapa do processo, chamada de função, tal que o fluxo seja evento-função-evento. Baseado em regras de decisão, utiliza-se dos operadores lógicos *E*, *OU* e *XOR* para a tomada de decisão dos fluxos de operação (DAVIS; BRABÄNDER, 2007).

EPC auxilia na modelagem, análise e redesenho de processos de negócio através de eventos, funções, regras e recursos, tanto para processos simples como para processos mais complexos. Além disso, a notação é de fácil entendimento e utilização, sendo assim uma notação bem aceita no mercado (MILI et al., 2010).

O autor complementa que seu uso é aconselhável para a construção e compreensão rápida de modelos, tanto simples como os complexos que possuem várias interfaces e submodelos de processos, e para modelagem de processos que visam a integração para sistemas de gestão empresarial *Enterprise Resourcing Planning* (ERP).

As vantagens apontam uso difundido em várias organizações, facilidade para identificar restrições de processo e capacidade de construção rápida de modelos. As desvantagens ficam por conta das limitações das ferramentas proprietárias da família ARIS e ao intenso cuidado na modelagem para evitar a criação de lacunas devido ao uso dos operadores lógicos (BPM-CBOK, 2013). A Figura 4 ilustra um exemplo de modelo de processo de negócios modelado com a notação EPC.

Figura 4: Ilustração de modelagem EPC



Fonte: Davis e Bradänder (2007, p. 31)

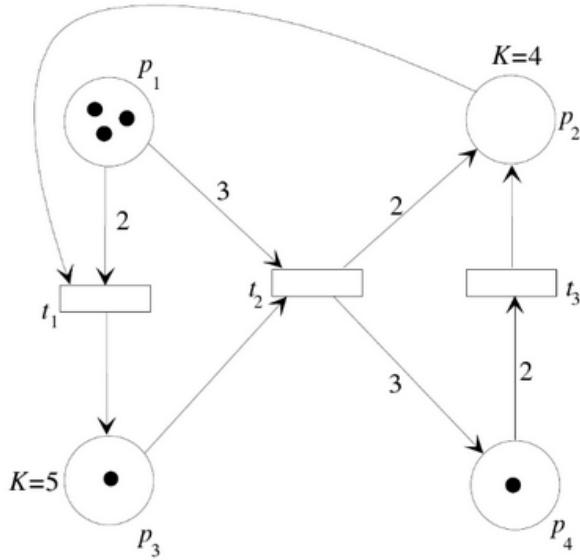
2.3.3 Petri Nets

Conforme Costa (2011), Petri Nets foi criado por Carl Petri em 1962 para ilustrar fluxos de atividades através de grafos direcionados que apresentam lugares e transições. As atividades (vértices) são ligadas por arestas (arcos) ponderadas, que estabelecem pesos associados, capazes de representar a concorrência de eventos.

Uma Petri Net é capaz de descrever as diversas etapas de um sistema através de um conjunto de símbolos que envolve controle, concorrência, conflitos, condições e sincronização. O fluxo de sequência de uma Rede de Petri indica os estados das atividades antes e após de executar cada evento, representado por um *token*, que percorre todo o modelo (MURATA, 1989).

Entre as principais vantagens, destacam-se o uso para modelagem hierárquica através de uma abordagem *top-down* e *bottom-up*, possibilidade de decomposição de sistemas e definições matemáticas bem definidas. Com relação às desvantagens, destacam-se a construção de modelos muito grandes para desenhar processos complexos e a dificuldade de desenvolver algoritmos para a edição das redes (COSTA, 2011). A Figura 5 ilustra um exemplo de modelo de processo de negócio modelado com a notação Petri Nets.

Figura 5: Ilustração de modelagem Petri Nets



Fonte: Costa (2011).

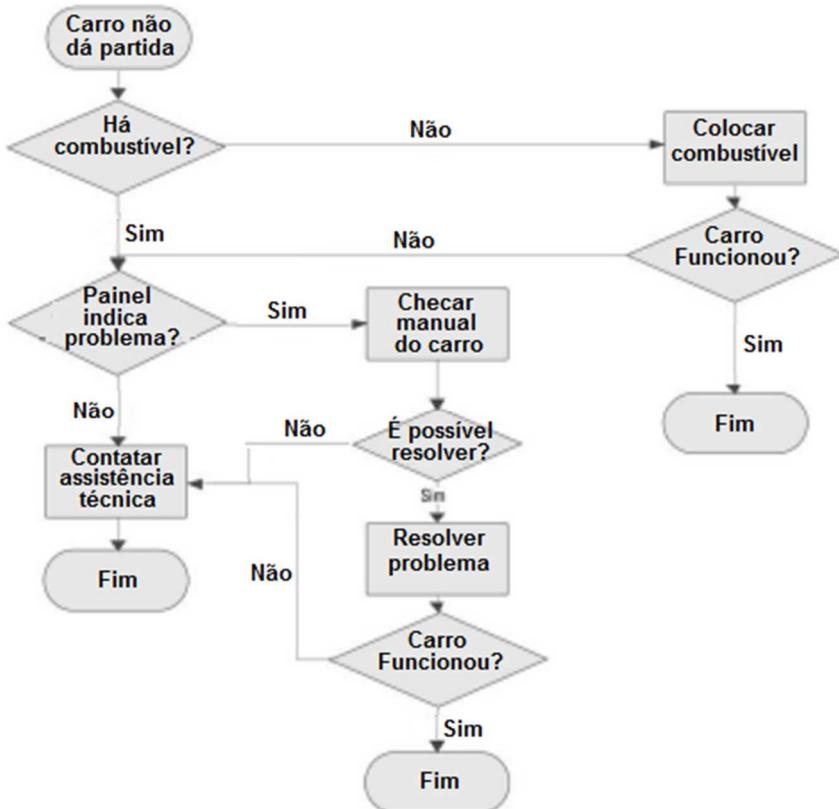
2.3.4 Workflow

Por ser uma das notações mais antigas, é considerada como precursora, pois serviu de inspiração para a criação de várias outras notações. Baldam et al. (2014) conceituam *workflow* como um diagrama que descreve, passo a passo, o fluxo de uma atividade baseado em decisões intermediárias que visam a execução de um processo. Lucinda (2010) completa que através de uma forma gráfica, um *workflow* descreve as etapas de um processo desde o início até a sua finalização, auxiliando assim na compreensão de forma rápida do seu funcionamento.

Apesar de ter sido aprovada como um padrão ANSI para representar fluxos de sistemas na década de 70, muitos símbolos diferentes têm sido utilizados para descrever entradas, fluxos, papéis, artefatos e saídas, fazendo com que essa notação não seja padronizada (BPM-CBOK, 2013). Seu uso é aconselhável para capturar fluxos de processos sem muitos detalhes, em especial no início de projetos, quando ainda não existe financiamento para aquisição de ferramentas com mais recursos disponíveis.

Segundo Baldam et al. (2014), as vantagens rodeiam a facilidade de aprendizado e entendimento, que focam na ilustração em alto nível do processo apresentado, na representação de fluxos de processos simples, comumente chamados de *caminhos felizes* e no suporte de inúmeras ferramentas de modelagem de baixo custo. Já as desvantagens se concentram na imprecisão para descrever processos complexos, ausência de um conjunto de atributos descritivos e falta de padronização dos elementos que podem fazer parte dos fluxogramas. A Figura 6 ilustra um exemplo de modelo de processo de negócio modelado com *workflow*.

Figura 6: Ilustração de modelagem workflow



Fonte: Adaptado de Baldam et al. (2014).

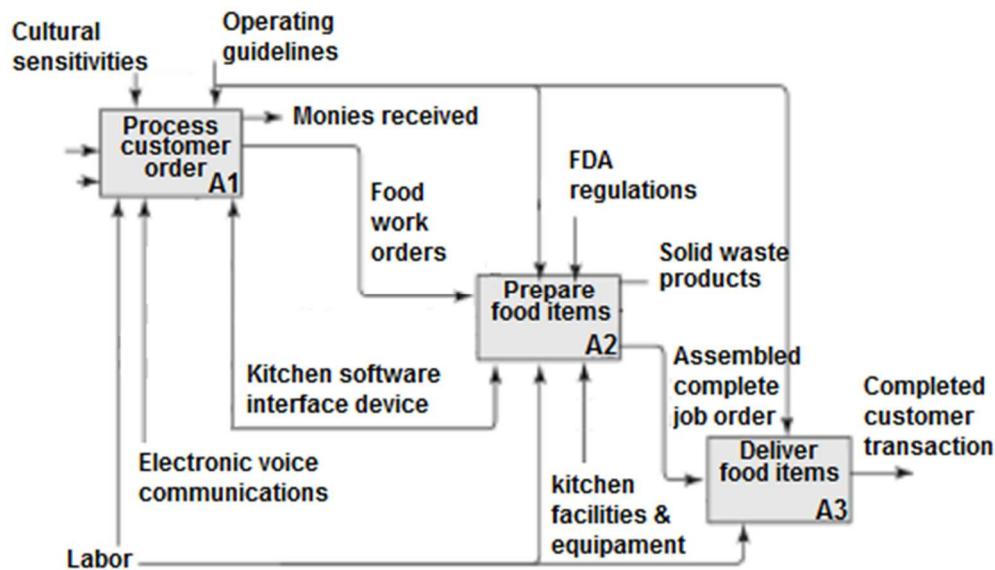
2.3.5 IDEF

Integration Definition for Function Modeling é uma notação de modelagem desenvolvida pela Força Aérea americana para descrever decisões, ações e atividades de uma organização, em especial em ambientes de manufatura (BPM-CBOK, 2013). Sua notação é composta por um conjunto simples de símbolos, que consiste de caixas e setas que indicam entradas, controles, mecanismos e saídas. A leitura e interpretação se faz de cima para baixo e da esquerda para a direita, representando, assim, que os níveis superiores definem os tópicos que são modelados.

Hernandez-Matias e Abreu (2006) explicam que a notação IDEF além de ser muito simples, foi largamente utilizada na indústria visando atender aos mais variados propósitos, tais como, auxílio em estudos de como implementar sistemas ERP, modelagem para implementar técnicas de manufatura e avaliação de técnicas de informática para serem introduzidas em pequenas e médias indústrias, entre outras.

Como vantagens, destacam-se a existência de uma representação precisa, a facilidade de seguir o fluxo lógico dos níveis dos tópicos e a vasta documentação que o Governo Federal americano disponibilizou (DRISCOLL, 2011). Em contrapartida, a visualização dos modelos é pouco amigável devido a confusão de caixas e setas que se disponibilizam nos casos mais complexos (BPM-CBOK, 2013). A Figura 7 ilustra um exemplo de modelo de processo de negócio modelado com IDEF.

Figura 7: Ilustração de modelagem IDEF



Fonte: Adaptado de Driscoll (2011, p. 48)

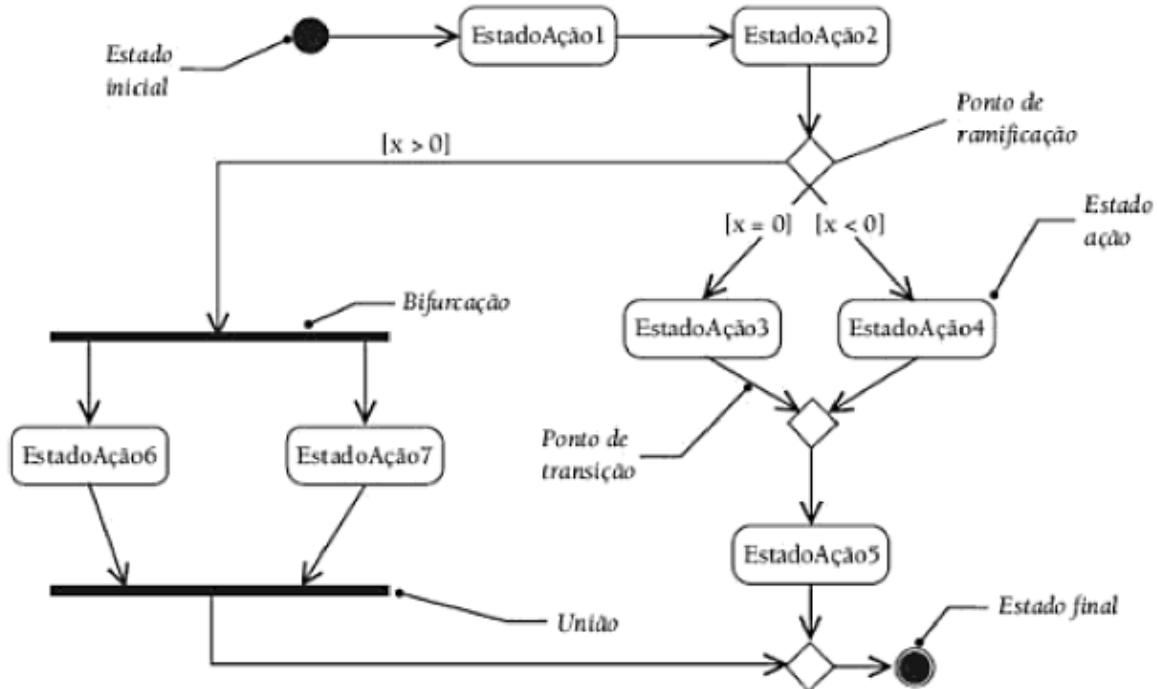
2.3.6 UML

Unified Modeling Language é uma linguagem de modelagem de sistemas de informação mantida pela OMG. Possui um conjunto padrão de técnicas e notações que servem para descrever, desenhar e documentar atividades relacionadas ao desenvolvimento de software e modelagem de classes, componentes, comunicação, atividades, pacotes entre outros (OMG-UML, 2015; KO et al., 2009).

Bezerra (2015) e Brocke e Rosemann (2013) citam que no contexto do BPM, a UML define o diagrama de atividades para realizar a modelagem de processos de negócios através de fluxos que estabelecem o controle de uma atividade para outra, dentro das etapas sequenciais de um processo, muito embora seu uso não seja muito frequente.

Apesar de o diagrama de atividades não ser de uso muito frequente, a UML possui como vantagem uma comunidade de usuários bem estabelecida, além de vasta disponibilidade de documentação. A desvantagem se destaca no sentido de que a UML é mais voltada para a área técnica de desenvolvimento de *software* e para modelagem de aplicações, deixando a modelagem de processos de negócio em segundo plano (BPM-CBOK, 2013). A Figura 8 ilustra um exemplo de modelo de processo de negócio modelado com UML.

Figura 8: Ilustração de modelagem UML (diagrama de atividades)



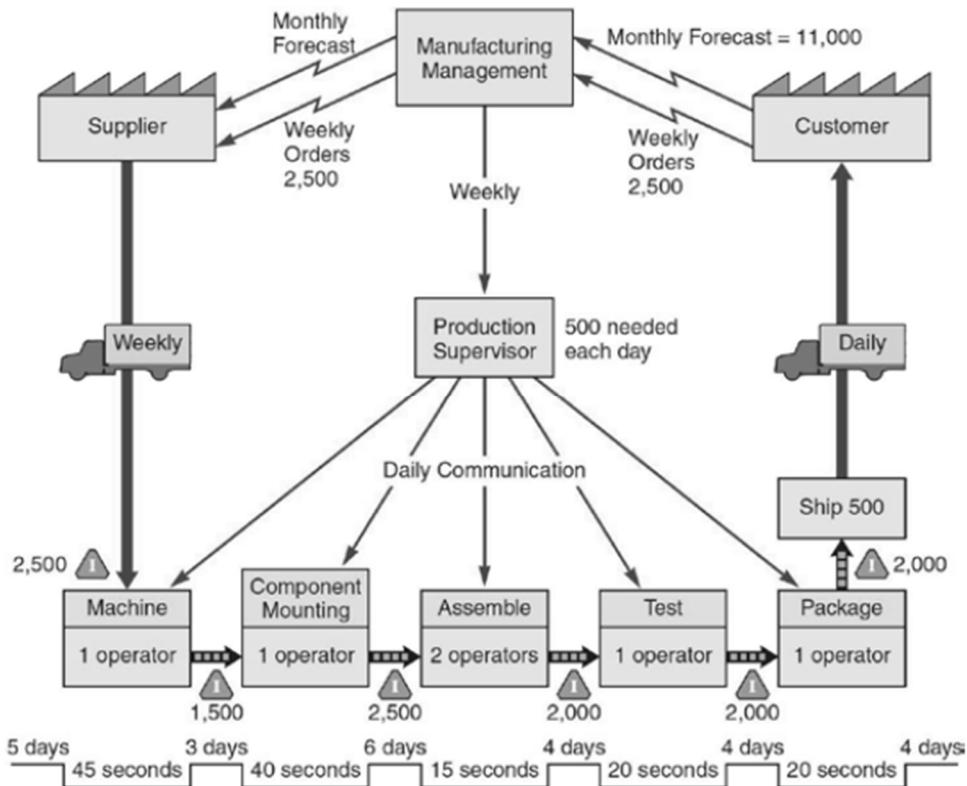
Fonte: Bezerra (2015)

2.3.7 VSM

Value Stream Mapping é uma notação que possui um conjunto simplificado de símbolos que permite mapear o fluxo de materiais/produtos e informações (valor). Esse mapeamento proporciona uma visão eficiente do processo ao adicionar custos de recursos e elementos de tempo em um modelo de processos (NASH; POLING, 2008).

Conforme Locher (2008), seu uso se concentra na análise e desenho para otimização de fluxos de processos através de múltiplas atividades associadas à manufatura, logística, serviços relacionados à indústria e cadeia de suprimentos. A vantagem dessa notação está na simplicidade de uso e entendimento, e as desvantagens apontam modelos planos e pouco complexos em nível de detalhes (BPM-CBOK, 2013). A Figura 9 ilustra um exemplo de modelo de processo de negócio modelado com VSM.

Figura 9: Ilustração de modelagem VSM



Fonte: BPM-CBOK (2013, p. 91)

2.4 Definição da notação de modelagem deste trabalho

A partir dos estudos realizados na seção anterior, que descreveram um resumo das características, funcionalidades, pontos positivos e negativos das principais notações de modelagem de processos de negócio, optou-se por utilizar BPMN como notação padrão neste trabalho, conforme as seguintes justificativas:

- É um padrão mantido pela OMG, organização aberta, sem fins lucrativos, que estabelece padrões para diversos sistemas de informação como, por exemplo, *Unified Modeling Language* (UML), *Common Object Request Broker Architecture* (CORBA), *Common Warehouse Metamodel* (CWM), *Model Drive Architecture* (MDA) e outros (OMG-BPMN, 2013).
- Possui nível considerável de uso e entendimento por diversas organizações e é largamente aceita no cenário mundial atualmente, tanto na indústria como na academia. Além disso, no Brasil, o Governo Federal elegeu BPMN como notação oficial para modelagem de processos de negócio, cuja formalização se encontra na arquitetura EPING - Padrões de Interoperabilidade de Governo Eletrônico (CAMPOS, 2014; EPING, 2015).
- Existem mais de 100 ferramentas conhecidas para modelagem de processos de negócio que utilizam a notação BPMN, sendo que muitas delas são gratuitas (HITPASS, 2014). Em complemento a isso, a especificação da notação define uma linguagem padrão que permite que um modelo BPMN possa ser intercambiável por diversas ferramentas de

modelagem que utilizam esse padrão, permitindo assim que as empresas possam trocar de ferramenta sem ter que redesenhar seus processos de negócio (OMG-BPMN, 2013).

- Tem sido largamente estudada na academia devido sua popularização no mercado, uma vez que as ferramentas BPMN dão suporte à automação das principais etapas de gerenciamento de processos de negócio como, por exemplo, análise, simulação, execução e monitoramento de processos (PILLAT et al., 2015).
- Possui um conjunto robusto e padronizado de símbolos que representam a modelagem de processos de negócio de forma clara, através de relacionamentos que definem a ordem de precedência de cada atividade, conforme o papel desempenhado por cada ator na execução do trabalho, sendo facilmente entendível por diversos níveis dentro da organização (BPM-CBOK, 2013).
- Apresenta a possibilidade de indicar eventos de início, intermediário e fim, além do fluxo de mensagens, comunicação e colaboração intranegócio, que são essenciais para definir, de forma completa, os diferentes aspectos de processos de negócio existentes nas organizações (OMG-BPMN, 2013).

2.5 Especificação BPMN

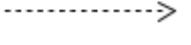
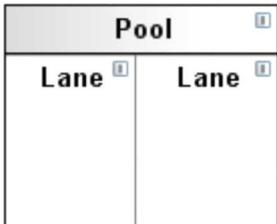
A primeira versão (1.0) foi disponibilizada em 2004, quando ainda pertencia à *Business Process Management Initiative* (BPMI), e teve por objetivo padronizar as diversas notações de modelagem de processos de negócio existentes no mercado, além de torná-la facilmente entendível nos diversos níveis das organizações, desde analistas até altos cargos de direção. Em 2005 a OMG incorporou a BPMI e em 2008 e 2009 foram lançadas as versões 1.1 e 1.2, respectivamente. A versão atual (2.0) está no mercado desde 2011 e tem sido largamente utilizada em diversas organizações (HITPASS, 2014).

A especificação BPMN contempla cinco categorias, compostas de subcategorias que descrevem os elementos da notação: (1) *Flow Objects* (*Events, Activities e Gateways*), (2) *Data* (*Data Objects, Data Inputs, Data Outputs e Data Stores*), (3) *Connecting Objects* (*Sequence Flows, Message Flows, Associations e Data Associations*), (4) *Swimlanes* (*Pools e Lanes*) e (5) *Artefacts* (*Group e Text Annotation*). A Tabela 3 descreve os elementos básicos da especificação, sendo que a relação completa dos elementos e seus respectivos aspectos pode ser encontrada em OMG-BPMN (2013).

Brocke e Rosemann (2015) explicam que a especificação BPMN não apenas definiu um conjunto padrão de símbolos para representar os aspectos de modelos de processos de negócio, como também estabeleceu uma definição padrão para intercâmbio, chamada *XML Process Definition Language* (XPDL).

Esse padrão comprehende um esquema XML, através de uma linguagem de definição, que armazena todos os aspectos de um modelo em um formato padrão. Isso permite que os modelos BPMN possam ser intercambiáveis entre as diversas ferramentas de modelagem que atendam essa especificação, facilitando assim a importação e a exportação de modelos desenhados em ferramentas diferentes ou, até, uma possível troca de ferramenta, sem ter que redesenhar os modelos já existentes.

Tabela 3: Elementos básicos BPMN

Elemento	Descrição	Notação
<i>Flow Object: Event</i>	Com o símbolo em formato de círculo, events são algo que aparecem durante o fluxo de um processo. Podem ser <i>start event</i> , <i>intermediate event</i> e <i>end event</i> , e se caracterizam por ações que iniciam um fluxo (<i>start</i>), que ocorrem durante a execução de um fluxo (<i>intermediate</i>) ou de finalização de um fluxo (<i>end</i>).	 Start Event  Intermediate Event  End Event
<i>Flow Object: Activity</i>	Em um modelo de processo de negócio, activity podem ser uma simples <i>activity</i> ou um <i>sub-process</i> (conjunto de <i>activity</i>) e significam um trabalho que a organização realiza.	 Task  Sub-Process
<i>Flow Object: Parallel Gateway, Exclusive Gateway e Inclusive Gateway</i>	São elementos que definem a sequência do fluxo de decisão através do controle de convergência ou divergência, os principais gateways são: <i>parallel</i> , <i>exclusive</i> e <i>inclusive</i> . <i>Parallel</i> indica que todas opções são executadas paralelamente; <i>exclusive</i> executa apenas uma das opções que são avaliadas; e <i>inclusive</i> pode executar um ou mais caminhos, dependendo dos critérios de avaliação.	 Parallel Gateway  Exclusive Gateway  Inclusive Gateway
<i>Data Association: Data Input, Data Output e Data Store</i>	<i>Data</i> são artefatos produzidos ou incorporados pelas <i>activity</i> que são realizadas no modelo de processo de negócio. As principais são <i>input</i> , <i>output</i> e <i>store</i> .	 Data Input  Data Output  Data Store
<i>Connecting object: Sequence Flow, Message Flow e Association</i>	<i>Connecting</i> definem o fluxo entre as atividades que são realizadas no modelo. Podem ser de <i>sequence</i> , <i>message</i> e <i>association</i> . <i>Sequence flow</i> definem a ordem em que as atividades são realizadas no processo; <i>Message flow</i> mostram a troca de mensagens entre participantes diferentes; e <i>Association</i> são usadas ligar informações e artefatos aos elementos.	  
<i>Swimlane: Pool e lane</i>	<i>Swimlane</i> ilustra um participante ou ator (externo) na representação de um modelo e <i>lane</i> ilustra a subpartição, como um perfil ou entidade interna da organização.	

Fonte: Adaptado de BPM-CBOK (2013, p. 26)

A Figura 10 ilustra parte do arquivo de um modelo desenhado em uma ferramenta que atende a especificação da WFMC-XPDL (2015). Esse modelo apresenta o cabeçalho que contém dados do esquema XML, informações sobre leiaute do modelo, alguns elementos como, por exemplo, *startEvent*, *task*, *sequenceFlow* e *endEvent* e suas respectivas propriedades, tais como, identificador (*id*), descrição (*name*), ligações de entrada (*incoming*) e saída (*outgoing*), entre outros.

Figura 10: Ilustração do conteúdo XML de um modelo BPMN

```

<?xml version="1.0" encoding="WINDOWS-125
<definitions xmlns="http://www.omg.org/spec/BPMN/2.0"
  <process id="PROCESS_1" isClosed="false"
    <extensionElements>
      <yao:pageFormat height="842.4" imageType="JPG" orientation="Portrait" pageWidth="1188" pageHeight="842.4" background="#FFFFFF" horizontal="true" vertical="true" style="border: none;"/>
    </extensionElements>
    <startEvent id="_2" isInterrupting="true">
      <extensionElements>
        <yaoqiang:label offset-x="1.0" offset-y="1.0" text="Start Event" style="font-size: 14px; font-weight: bold; color: black; border: 1px solid black; padding: 2px; border-radius: 5px;"/>
      </extensionElements>
      <outgoing>_24</outgoing>
    </startEvent>
    <task completionQuantity="1" id="_3">
      <incoming>_24</incoming>
      <outgoing>_21</outgoing>
    </task>
    <task completionQuantity="1" id="_4">
      <incoming>_21</incoming>
      <outgoing>_16</outgoing>
    </task>
    <sequenceFlow id="_21" sourceRef="_3" targetRef="_5">
      <endEvent id="_5" name="FIM">
        <incoming>_12</incoming>
        <incoming>_28</incoming>
        <incoming>_14</incoming>
      </endEvent>
    </sequenceFlow>
    ...
  </process>
</definitions>

```

Fonte: Elaborado pelo autor

2.6 Composição de modelos de processos de negócio

No contexto do gerenciamento de processos de negócio, a automação de processos possibilita analisar, simular, executar e monitorar rotinas e processos que visam aumentar a produtividade e reduzir os custos. Dessa forma, modelos são redesenhados continuamente em busca de melhorias e adaptações que se fazem necessárias, sendo que essa constante transformação produz inúmeros modelos, que especificam e detalham os processos de uma organização ao longo de sua existência (HALLERBACH et al., 2008; YAN et al., 2010).

Os modelos existentes servem de base para a remodelagem de processos, isto é, normalmente aproveita-se o modelo no estado atual (AS-IS) para modelar o estado futuro (TO-BE), seja para reestruturar um processo ou fundir dois ou mais processos em um único, que deve conter o novo fluxo do trabalho (SCHWEGMANN e LASKE, 2003).

Desenhar o estado futuro de um processo baseando-se no estado atual é uma tarefa simples e fácil, pois existem ferramentas que auxiliam nessa atividade (BALDAN et al., 2014). Contudo, modelar o estado futuro (ou um novo processo) a partir de dois modelos existentes, na qual as atividades correlatas devem ser identificadas e fundidas - através de uma rotina de

composição, pode se tornar uma tarefa difícil, onerosa e propensa a erros, explicam Zemni et al. (2016) e La Rosa et al. (2013).

Nesse sentido, a composição de modelos pode ser conceituada como sendo um conjunto de atividades que devem ser executadas para combinar dois modelos de entrada (M_A e M_B), visando produzir um modelo composto (M_C), cujo resultado representa a integração entre M_A e M_B (FARIAS et al., 2014). A partir dessa definição, entende-se que o aumento da produtividade dos analistas ocorrerá, de fato, quando as partes comuns dos modelos de entrada forem efetivamente identificadas, gerando o modelo M_C de forma rápida e livre de falhas.

A composição de modelos de processos de negócio se justifica quando ocorre a necessidade de comparar e integrar modelos que se referem a processos similares em empresas ou setores diferentes como, por exemplo, em processos de fusão ou aquisição de empresas. Rasmussen (1989) explica que a fusão de empresas ocorre quando duas companhias se fundem para gerar uma terceira. Já Barros (2001) esclarece que a aquisição de empresas ocorre quando uma companhia absorve e integra todos os ativos de outra.

Em ambos os casos, há a necessidade de comparar e integrar processos, que visam agrupar vantagens como redução de custos e aumento do ativo, explicam Clemente e Greespan (1998). Essa integração passa pela modelagem de processos, uma vez que os modelos representam a forma como as atividades são documentadas e realizadas nas empresas. Dois grandes exemplos de fusão realizados no Brasil são a união do Itaú com o Unibanco em 2008 que gerou a empresa Itaú-Unibanco e a Sadia com a Perdigão em 2009, que gerou a empresa BR-Foods, citam Vidal et al. (2014).

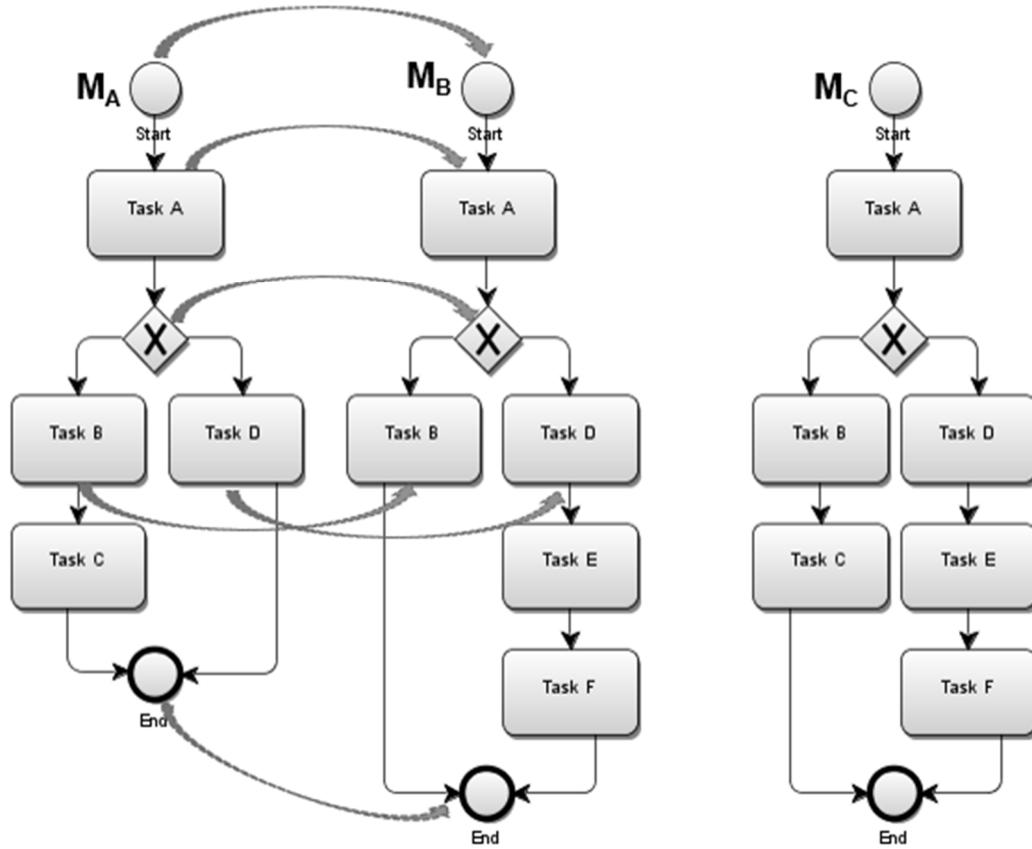
Para realizar a composição, os elementos (*event, activity, gateway, etc.*) e suas relações de interdependência, em cada modelo de entrada, devem ser comparados objetivando identificar possíveis similaridades. Os elementos que não são comuns são transportados para o modelo composto, enquanto que os elementos afins são sobrepostos. A Figura 11 apresenta um exemplo do processo de identificação de similaridades entre dois modelos, M_A e M_B , e o respectivo modelo M_C - resultado da composição.

Nesse exemplo, os elementos *Start*, *Task A*, *Task B*, *Task D* e *End* foram identificados como similares entre os dois modelos de entrada (M_A e M_B) e então transportados para o modelo composto (M_C), enquanto que os elementos *Task C*, *Task E* e *Task F* foram acrescentados ao modelo composto, ou seja, estavam em apenas um dos modelos de entrada.

Muito embora a composição demonstrada no exemplo supracitado seja simples, casos complexos também fazem parte das rotinas de integração de modelos como, por exemplo, *activity* com descrições diferentes que representam as mesmas atividades, *gateway* posicionado em locais diferentes nos modelos de entrada e situações em que determinados elementos não devem fazer parte do modelo composto, se assim estiver especificado nas regras de negócio do novo processo.

Na prática da composição de modelos não é incomum que os modelos integrados, a partir de dois modelos de entrada, apresentem inconsistências, ou seja, que os modelos sejam produzidos diferentemente do modelo desejado (LA ROSA et al., 2013; BRANCO, 2014). Rozman et al. (2008) explicam que inconsistências compreendem erros sintáticos e semânticos que prejudicam o entendimento dos modelos.

Figura 11: Dois modelos BPMN e seu respectivo mapeamento para o modelo composto



Fonte: Elaborado pelo autor

Os autores explicam ainda que erros sintáticos têm relação com a falta de entendimento ou descuido com relação à notação utilizada. Exemplos de erros sintáticos são fluxos de eventos intermediários mal posicionados ou a falta de etiqueta descritiva nos fluxos de decisão. Já erros semânticos são aqueles em que os elementos estão de acordo com a especificação, porém, possuem problemas de significância e dificultam o entendimento e a comprehensibilidade do modelo, ou seja, não estão de acordo com as regras de negócio estabelecidas nos requisitos funcionais. Exemplo de erro semântico é uma atividade que não possui ligação de saída com qualquer outro elemento dentro de um modelo. Branco (2014) cita que essas inconsistências podem causar sérios problemas como, por exemplo, o atraso no desenvolvimento de uma nova versão de um sistema.

A literatura atual mostra que a academia tem feito estudos de composição de modelos de processos de negócio e de identificação de similaridades entre modelos, conforme mostram os trabalhos de Zemni et al. (2016), Ivanov (2015), Armas-Cervantes et al. (2014), Santos (2014), Branco (2014), Zhu et al. (2014), La Rosa et al. (2013), Niemann et al. (2012), Becker e Laue (2012), Gerth et al. (2011), Dijkman et al. (2011), Fazal-Bagaie (2009), Brunet et al. (2006) e Sun et al. (2006), porém, as técnicas utilizadas são imprecisas e não proporcionam redução de tempo e de esforço dos analistas, uma vez que geram modelos complexos e de difícil entendimento, baseados em notações genéricas diferentes das utilizadas na indústria.

Além disso, essas técnicas apresentadas na academia não permitem a interação do analista durante a integração dos modelos. Essa interação se faz necessária devido ao fato que os modelos construídos devem se basear nas regras de negócio dos processos e não na

integração automática dos elementos, haja vista que nem sempre todos os elementos dos modelos devem ser transportados para o modelo composto.

Embora a academia tenha feito avanços em relação à composição de modelos, a indústria ainda deixa a desejar, uma vez que as ferramentas de modelagem de processos de negócio como, por exemplo, *Eclipse BPMN2 Modeler ECLIPSE* (2015), *IBM Business Process Manager IBM* (2015), *Oracle Business Process Manager ORACLE* (2015), *Microsoft BPMN Visio Modeler VISIO* (2015) e *Bizagi Modeler BIZAGI* (2015) não fornecem mecanismos que auxiliem a execução da integração de modelos, fazendo com que os analistas realizem a composição de forma manual, aumentando ainda mais a propensão de erros (BRANCO, 2014; ZEMNI et al., 2016).

Dessa maneira, percebe-se a ausência de uma ferramenta de composição semiautomática, que utilize as vantagens das melhores técnicas disponíveis na literatura, para apoiar o trabalho de composição de modelos que permita aos analistas interagir com ela, a partir da aprovação ou rejeição das similaridades apontadas, permitindo assim integrar modelos de modo eficaz e eficiente, aumentando a produtividade e diminuindo os custos ao construir modelos com redução do esforço e minimização de erros.

3 TRABALHOS RELACIONADOS

Este Capítulo apresenta os trabalhos selecionados que possuem relação com o objeto desta pesquisa descrito na Seção 1.3 ou que contribuíram de alguma forma para a sua construção e desenvolvimento. Possui o objetivo de relacionar alguns estudos sobre o estado da arte de composição de modelos de processos de negócio, destacando seus propósitos, características e técnicas utilizadas. Para isso, este Capítulo foi organizado da seguinte maneira. A Seção 3.1 apresenta a metodologia para a escolha dos trabalhos relacionados. As Seções de 3.2 a 3.7 descrevem os trabalhos relacionados selecionados e a forma como eles apresentam suas técnicas e ferramentas. A comparação dos trabalhos relacionados é realizada na Seção 3.8. Por fim, a Seção 3.9 relaciona as oportunidades de pesquisa deste trabalho.

3.1 Metodologia para a escolha dos trabalhos

O tema composição de modelos de processos de negócio envolve duas questões principais: a identificação de similaridade entre modelos (ou fragmentos de modelos) e a integração de elementos para formar um modelo composto. A identificação de similaridades se dá pelas diversas maneiras de comparar um modelo com outro e a integração compreende as formas como esses modelos são integrados para formar um modelo composto, fruto da composição dos modelos de origem.

Dessa forma, dois critérios de busca foram criados para realizar a pesquisa dos trabalhos relacionados: identificação de similaridades e integração de modelos. O primeiro critério permitiu encontrar trabalhos que possuem a capacidade de comparar modelos, seja por sua estrutura ou comportamento e o segundo critério possibilitou localizar trabalhos que realizam, de fato, a integração de modelos, mesmo que apenas teoricamente.

Para isso, foram consultadas as principais bibliotecas digitais de publicações de pesquisa como, por exemplo, ACM, Springer, Elsevier e IEEE. Além dessas bases, foram procurados, também, trabalhos de pesquisa em nível de Pós-Graduação em Universidades nacionais e estrangeiras. Diversos termos foram inseridos nas pesquisas como, por exemplo, *similarity business process*, *comparison business process*, *merging business process*, *collections business process*, *analisis business process*, *matching business process*, comparação de modelos, composição de modelos, integração de modelos, unificação de modelos, entre outros.

Diversos trabalhos foram encontrados com base nas pesquisas realizadas. No campo da identificação de similaridades, alguns trabalhos apresentam apenas identificação de similaridades como Ivanov et al. (2015), Armas-Cervantes (2014) e Bianchin et al. (2010) e outros trabalhos apresentam identificação de similaridades para melhorar gestão de repositórios de modelos como Dijkman et al. (2012) e Niemann et al. (2012). Já no campo de integração de modelos, alguns trabalhos apresentam técnicas de comparação de grafos como Zemni et al. (2016) e La Rosa et al. (2013). Outros apresentam propostas de ferramentas para realizar a composição automaticamente, tais como, Santos (2014), Dijkman et al. (2011) e Gottschalk et al. (2008).

De uma forma ou de outra, vários dos trabalhos pesquisados contribuíram para o desenvolvimento desta pesquisa, seja para análise e comparação de técnicas e ferramentas ou para a elaboração de conceitos utilizados na fundamental teórica. Desses trabalhos estudados, sete foram selecionados para compor um estudo mais detalhado, por representarem uma ligação direta com o objeto desta pesquisa, conforme descrito nas próximas Seções deste Capítulo.

3.2 Merging Worflows

Nesse trabalho, Sun et al. (2006) fazem uma introdução conceitual de composição de modelos de processos de negócio como um dos primeiros trabalhos publicados sobre esse tema na academia. Para isso, relacionam duas situações especiais que a integração de modelos pode ser aplicada: (1) reengenharia de processos de uma empresa, quando ocorre a reestruturação de setores; e (2) fusão de companhias, quando duas companhias se fundem para formar uma terceira.

Para exemplificar a teoria proposta, os autores escolheram Petri nets¹ como notação de modelagem padrão utilizada nesse trabalho. Os autores justificam a escolha dessa notação principalmente por possuir uma base formal que possibilita determinar a solidez de uma composição de modelos, além de ser facilmente entendível.

Um conjunto de definições foi elaborado para descrever a formação e a estrutura dos modelos, bem como compreender as regras gerais de composição. Essas definições e regras envolvem os elementos e suas inter-relações como, por exemplo, *tasks*, *AND-joins*, *AND-splits*, *OR-joins* e *OR-splits* e as comparações entre esses elementos a partir dos modelos de entrada.

A composição apresentada nesse trabalho estabelece um processo para combinar dois modelos com o objetivo de eliminar atividades redundantes, mantendo todos os itens necessários no modelo produzido. Além disso, novos elementos condicionais podem ser acrescentados ao modelo composto, dependendo dos fluxos dos modelos que são tratados.

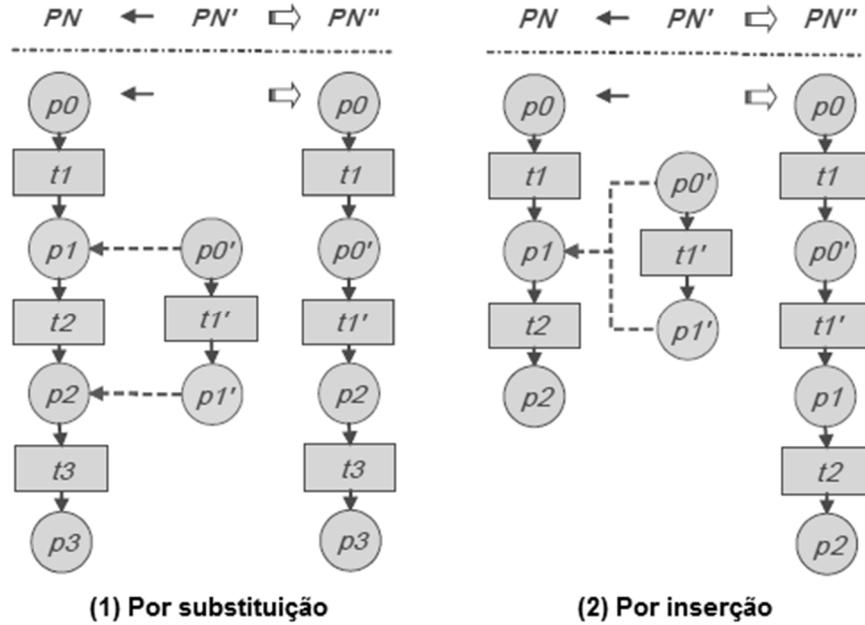
Quanto ao tipo de composição, dois conceitos principais são elaborados para a composição de modelos: composição sem perda de elementos e composição com perda de elementos. Sem perda de elementos significa que todos os elementos dos modelos de entrada são preservados e incluídos no modelo composto. Já a composição com perdas não garante que todos os elementos dos modelos de entrada serão preservados. Além disso, quatro categorias de composição de modelos são especificadas: Sequencial, Paralelo, Condisional e Iterativa. O detalhamento de cada categoria é apresentado na sequência.

A Composição Sequencial pode ser realizada de duas maneiras: com substituição de elementos e com inserção de elementos. A primeira estabelece que os elementos similares nos modelos de entrada serão inseridos no modelo produzido por substituição, ou seja, um elemento substitui o outro de modo que a quantidade de elementos do modelo composto será igual a quantidade de elementos do maior modelo de entrada – Figura 12 (1). Na segunda maneira, os elementos de ambos os modelos de entrada são acrescentados ao modelo composto, podendo torná-lo maior - Figura 12 (2).

A Composição Paralela é utilizada quando a ordem casual dos modelos de entrada não é relevante. Assim, elementos conectores *AND-split* e *AND-join* são inseridos no modelo composto para que os fluxos de ambos os modelos estejam representados, conforme ilustra a Figura 13. Nessa figura identifica-se a inclusão de *AND-split* após *t1* para executar paralelamente dois grupos de atividades (*p1*, *t2* e *p2*) e (*p1'*, *t2'* e *p2'*), e *AND-join* para realizar a união do fluxo em *t3*.

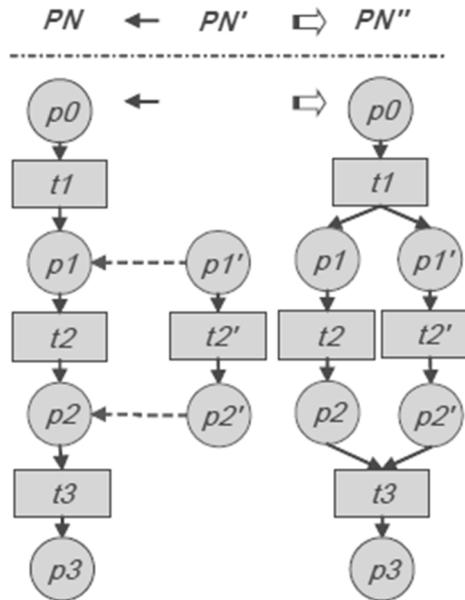
¹ <http://www.petrinets.info/>

Figura 12: Composição Sequencial



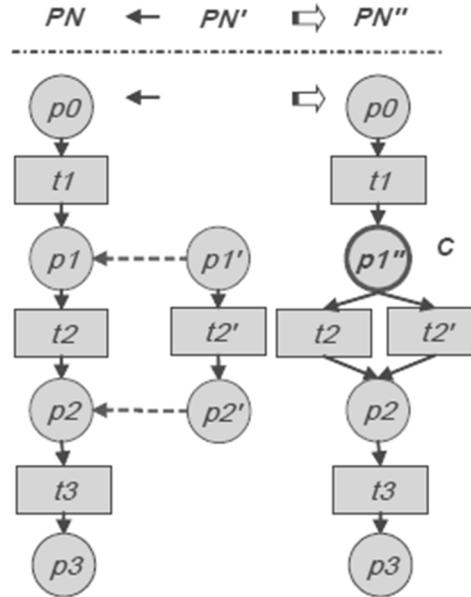
Fonte: Adaptado de Sun et al. (2006)

Figura 13: Composição Paralela



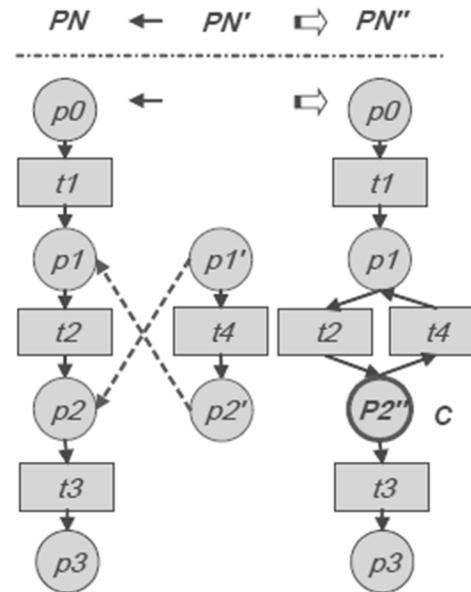
Fonte: Sun et al. (2006)

Na Composição Condicional, elementos de decisão *OR-split* e *OR-join* são incorporados ao modelo produzido para identificar critérios de seleção. Esses critérios definem fluxos alternativos para a execução de determinadas atividades, conforme pode ser visualizado na Figura 14. Nessa composição, o elemento *OR-split* verifica o critério de seleção C para decidir entre a execução da atividade $t2$ e $t2'$. Após a decisão e execução da atividade correspondente, o elemento *OR-join* unifica novamente o fluxo de execução do modelo.

Figura 14: Composição Condicional

Fonte: Sun et al. (2006)

O último tipo, definido como Composição Iterativa, atua quando um critério condicional no fluxo do processo estabelece o retorno do fluxo para uma atividade anterior, formando uma sequência iterativa, conforme ilustração apresentada na Figura 15. Nessa representação, o critério C , verificado sobre a atividade $P2''$, define uma condição para retornar para $t4$ ou seguir o fluxo para $t3$ assim que a condição é satisfeita.

Figura 15: Composição Iterativa

Fonte: Sun et al. (2006)

Essa abordagem não implementou uma ferramenta de composição para validar os conceitos introduzidos, tampouco realizou experimentos de integração de modelos, tronando-se assim apenas um referencial teórico sobre a temática, o que se pode considerar aceitável por ser um dos trabalhos pioneiros no ramo de composição de modelos.

Muito embora não tenha realizado experimentos ou apresentado uma ferramenta para compor modelos, elaborou conceitos fundamentais na área de composição que serviram de base para diversos outros trabalhos como, por exemplo, La Rosa et al. (2010), Gottschalk et al. (2008) e La Rosa et al. (2013) citados nesta pesquisa, além de ser mencionado por outros 69 artigos especializados, conforme pesquisa realizada no Google Scholar².

3.3 Merging Event-driven Process Chains

Gottschalk et al. (2008) apresentam uma proposta de composição automática de modelos de processos de negócio. Essa composição visa integrar dois modelos de modo que o comportamento descrito nos modelos de entrada continua preservado. Nesse trabalho, os autores propõem um *plug-in* que roda integrado ao framework ProM³ para produzir automaticamente um modelo composto que servirá de ponto de partida para a otimização do processo.

Event-driven Process Chain (EPC) foi a notação escolhida para a representação dos modelos. Segundo os autores, essa notação foi selecionada porque é facilmente entendida por analistas que modelam processos, uma vez que possui apenas três tipos de elementos (eventos, funções e conectores lógicos). Além disso, essa notação é suportada por ferramentas de modelagem conhecidas na indústria como, por exemplo, ARIS⁴ e Microsoft Visio⁵.

Um conjunto de definições foi elaborado para descrever como devem ser os modelos baseados na notação EPC, bem como as regras de integração. Essas definições estabelecem pré-requisitos que devem ser cumpridos para realizar a integração de modelos, baseada nessa notação, e como ocorre de fato a integração desses modelos para produzir o modelo composto.

O algoritmo proposto realiza a integração de modelos em três fases: na primeira fase os modelos EPC são transformados em grafos de funções, que representam o comportamento dos modelos de entrada; a segunda fase realiza a composição dos modelos representados por esses dois grafos; e, por fim, na terceira fase o modelo produzido em grafo de função é convertido novamente para a notação EPC.

Um grafo de função é uma redução de um modelo EPC em vértices e arestas que representam os elementos e o comportamento dos modelos. Tipos de divisão e junção também são utilizados para representar as operações dos fluxos dos modelos como, *AND* (\wedge), *OR* (\vee) e *XOR*, que identificam caminhos paralelos e estruturas condicionais. Um exemplo de conversão de modelo EPC para grafo de função é apresentado na Figura 16.

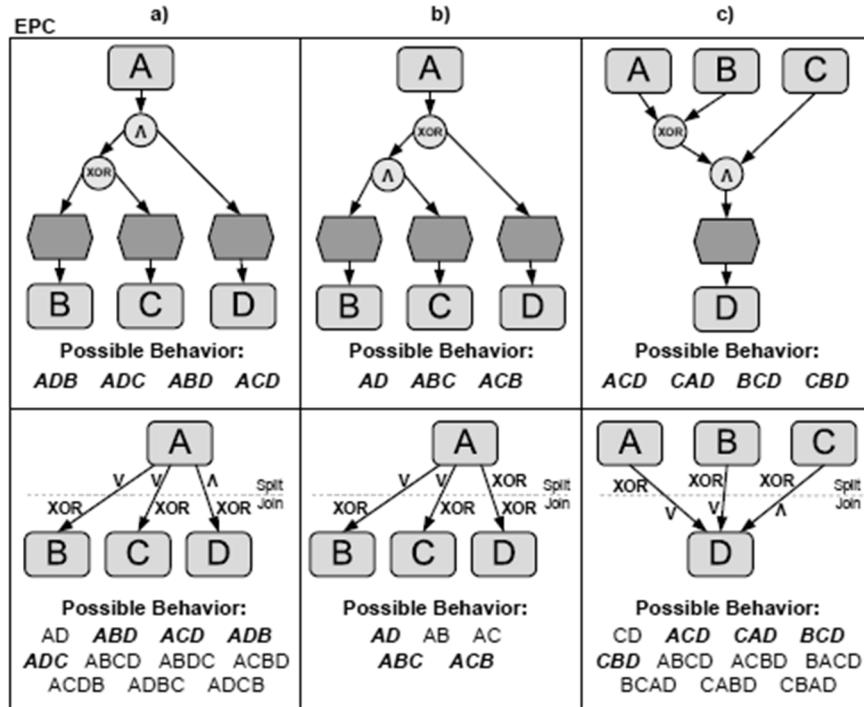
² <https://scholar.google.com.br>

³ <http://www.processmining.org/prom/start>

⁴ <http://www.ariscommunity.com/>

⁵ <https://products.office.com/pt-br/visio/microsoft-visio-top-features-diagram-software>

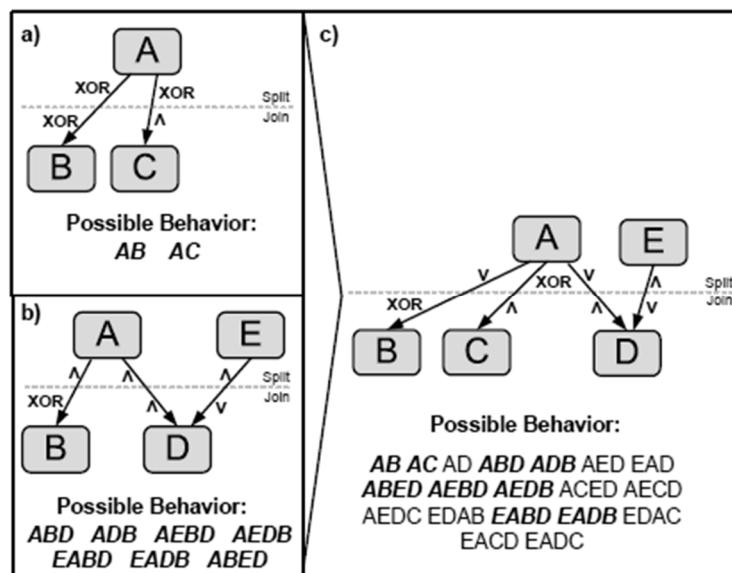
Figura 16: Conversão de modelo EPC em grafo de função



Fonte: Adaptado de Gottschalk et al. (2008)

Para integrar os dois modelos de entrada preservando o comportamento original, o algoritmo proposto combina os conjuntos de funções e de arcos para calcular a união ou separação dos conectores com base nos valores dos dois grafos de função. Para isso, é necessário analisar se as funções são sucedidas ou precedidas pelas mesmas funções em ambos os modelos, conforme ilustrado na Figura 17.

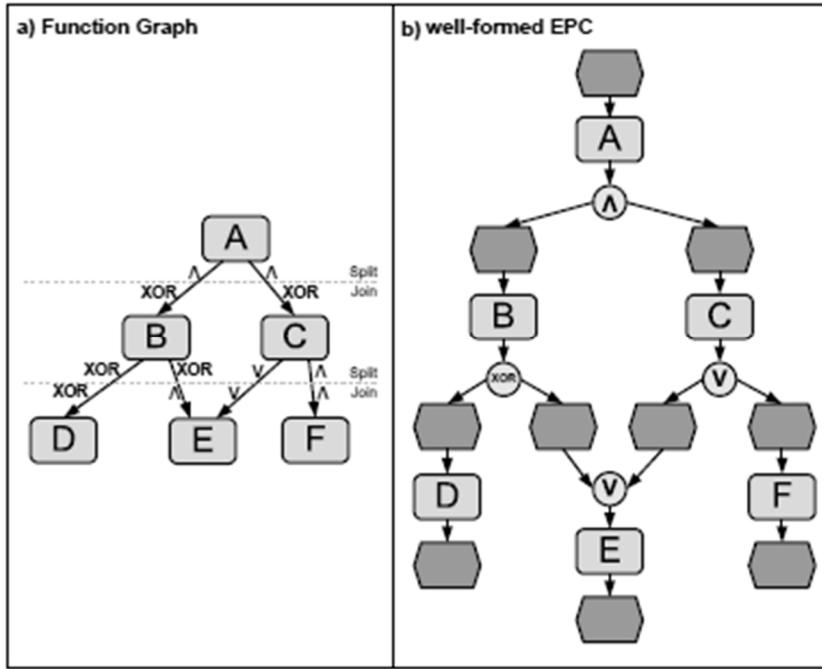
Figura 17: Integração de grafos de função



Fonte: Gottschalk et al. (2008)

A última fase transforma o grafo de função, resultado da integração dos dois grafos de origem, em um modelo EPC bem formado, conforme ilustra a Figura 18. Para isso, os valores dos conectores determinam o possível comportamento do modelo, a partir dos cálculos de separação e união dos tipos de arestas do grafo de função.

Figura 18: Conversão de grafo de função em EPC



Fonte: Adaptado de Gottschalk et al. (2008)

A ferramenta proposta por esse trabalho não faz identificação de similaridades de rótulos de elementos automaticamente. Essa associação é realizada manualmente pelo usuário através de duas listas que contêm a relação de elementos de cada modelo. O resultado final da composição produz um modelo que representa o comportamento de ambos os modelos originais, com acréscimo de conectores *AND*, *OR* e *XOR* para tornar isso possível.

3.4 Similarity of Business Process Models

Nesse trabalho, Dijkman et al. (2011) abordam técnicas para identificar a similaridade entre modelos de processos de negócio que estão localizados em um repositório de modelos. Essa identificação visa evitar que modelos sejam inseridos em duplicidade no repositório, pela inclusão de um novo modelo ou no caso de fusão de empresas, uma vez que grandes companhias possuem centenas ou até milhares de modelos. Para isso, três técnicas são apresentadas para realizar a comparação entre os modelos: similaridade de correspondência de elementos, similaridade estrutural e similaridade comportamental.

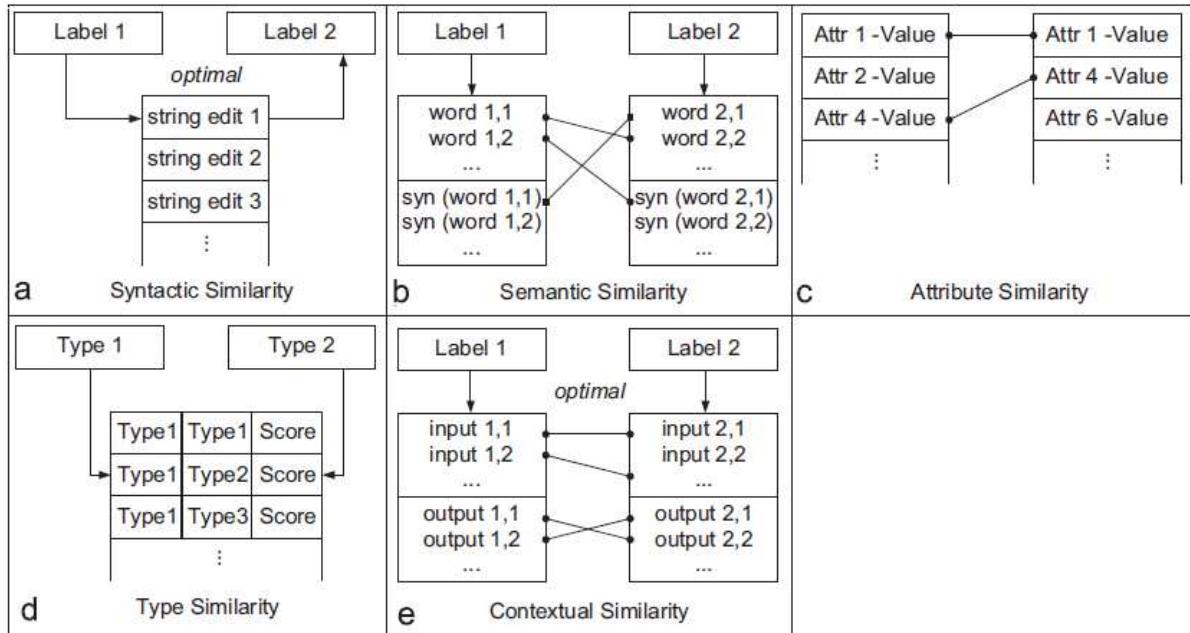
Para realizar as comparações no repositório, os modelos são convertidos para grafos chamados *Business Process Graph* (BPG), que possuem representações genéricas de fluxos de processos sem simbolizar a sintaxe específica de uma notação de modelagem como, por exemplo, BPMN, EPC e Diagrama de Atividades da UML.

Assim como em outros trabalhos citados neste Capítulo, um conjunto de definições foi elaborado para descrever a formação e a estrutura dos modelos, bem como para compreender

as regras que regem as técnicas que são apresentadas para identificação de similaridades. Essas definições e regras envolvem os elementos e suas inter-relações comportamentais nos grafos, assim como as fórmulas que estabelecem os cálculos de similaridade.

A técnica de similaridade de correspondência de rótulos de elementos realiza comparações entre os elementos de ambos os grafos com o objetivo de identificar nós (elementos) similares. Para isso, utiliza métricas de similaridade sintática, semântica, de atributo, de tipo e contextual, conforme ilustrado pela Figura 19.

Figura 19: Similaridade de Correspondência de Rótulos - métricas

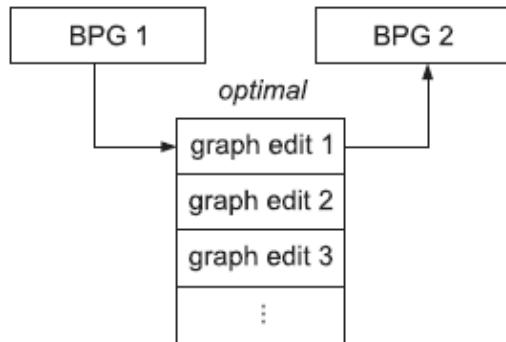


Fonte: Adaptado de Dijkman et al. (2011)

A similaridade sintática calcula a distância entre as *strings* das etiquetas dos rótulos com base no trabalho de Dijkman et al. (2009). Esse trabalho utiliza a distância de Levenshtein (1966) para calcular o número de operações necessárias para transformar uma *string* em outra. A similaridade semântica utiliza um dicionário de sinônimos para encontrar palavras diferentes que possuem aproximadamente o mesmo significado. As similaridades de atributo e de tipo devem ser utilizadas em conjunto com as similaridades sintética ou semântica e estabelecem comparações de valores de atributo e de tipos de elementos, respectivamente. Por último, a similaridade contextual leva em consideração os nodos que representam os elementos que sucedem e antecedem os elementos que estão sendo comparados nos modelos.

A técnica de similaridade estrutural leva em consideração a diferença de estrutura entre os grafos que representam a estrutura dos modelos comparados. Essa métrica calcula a distância entre grafos com base no número de operações necessárias para transformar um grafo em outro como, por exemplo, a eliminação, a substituição e a inclusão de vértices e de arestas, conforme ilustrado pela Figura 20.

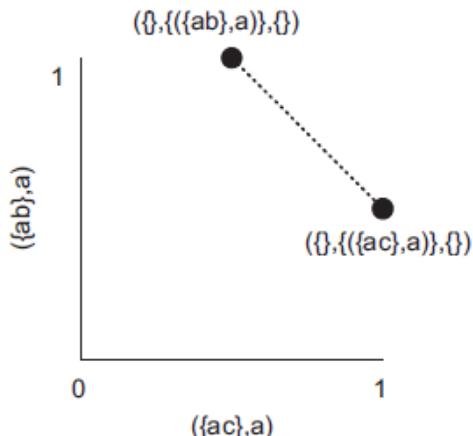
Figura 20: Similaridade Estrutural – distância entre grafos



Fonte: Dijkman et al. (2011)

A técnica de similaridade comportamental leva em consideração o comportamento dos modelos de processos comparados. Nessa técnica, ligações indiretas entre os elementos são consideradas, ao contrário da técnica de similaridade estrutural que não realiza essa consideração. O grau de similaridade é computado pela distância de espaço de um vetor construído que permite aumentar o escore de similaridade de comparação geral dos grafos, conforme ilustra a Figura 21.

Figura 21: Similaridade Comportamental – vetor de espaços entre dois grafos



Fonte: Dijkman et al. (2011)

Por fim, os autores implementaram um *plug-in* para o *framework* ProM - *Process Mining and Analysis*, com o objetivo de validar as técnicas apresentadas nesse trabalho. Essa avaliação se deu pela execução de um cenário compreendido por 100 modelos de processos de negócio, cujo objetivo estava na identificação do grau de similaridade entre os modelos pesquisados.

3.5 Business Process Model Merging

Nesse trabalho, La Rosa et al. (2013) abordam a composição de modelos de processos de negócio que compartilham de fragmentos comuns para produzir um modelo configurável. Essa composição é realizada levando em consideração três questões: (1) a preservação, que garante a conservação do comportamento dos modelos de origem no modelo composto; (2) a rastreabilidade, que permite identificar no modelo composto os elementos de cada modelo de entrada; e (3) a reversibilidade, que possibilita derivar os modelos de entrada a partir do modelo

composto. Além disso, a proposta defendida pelos autores apresenta técnicas de identificação de similaridade para apontar elementos comuns nos modelos de entrada.

O algoritmo de composição proposto não utiliza uma notação conhecida de modelagem de processos de negócio como, por exemplo, BPMN, EPC e Diagrama de Atividades da UML. Em vez disso, define uma abstração de modelagem a partir de um grafo direcionado com nós rotulados e conectores como *AND*, *OR* e *XOR*, chamado *Business Process Graph* (BPG), aparentemente parecida com a notação EPC.

Um conjunto de definições foi elaborado para descrever os critérios e as regras de identificação de similaridade e de composição. Essas definições estabelecem como ocorrem os mapeamentos de nós nos grafos para identificação de correspondências, as técnicas de identificação de similaridade dos rótulos dos elementos, bem como de integração desses elementos no modelo composto. Além disso, estabelecem escores de similaridade, que representam índices calculados que indicam os níveis de similaridade entre os elementos dos modelos de origem.

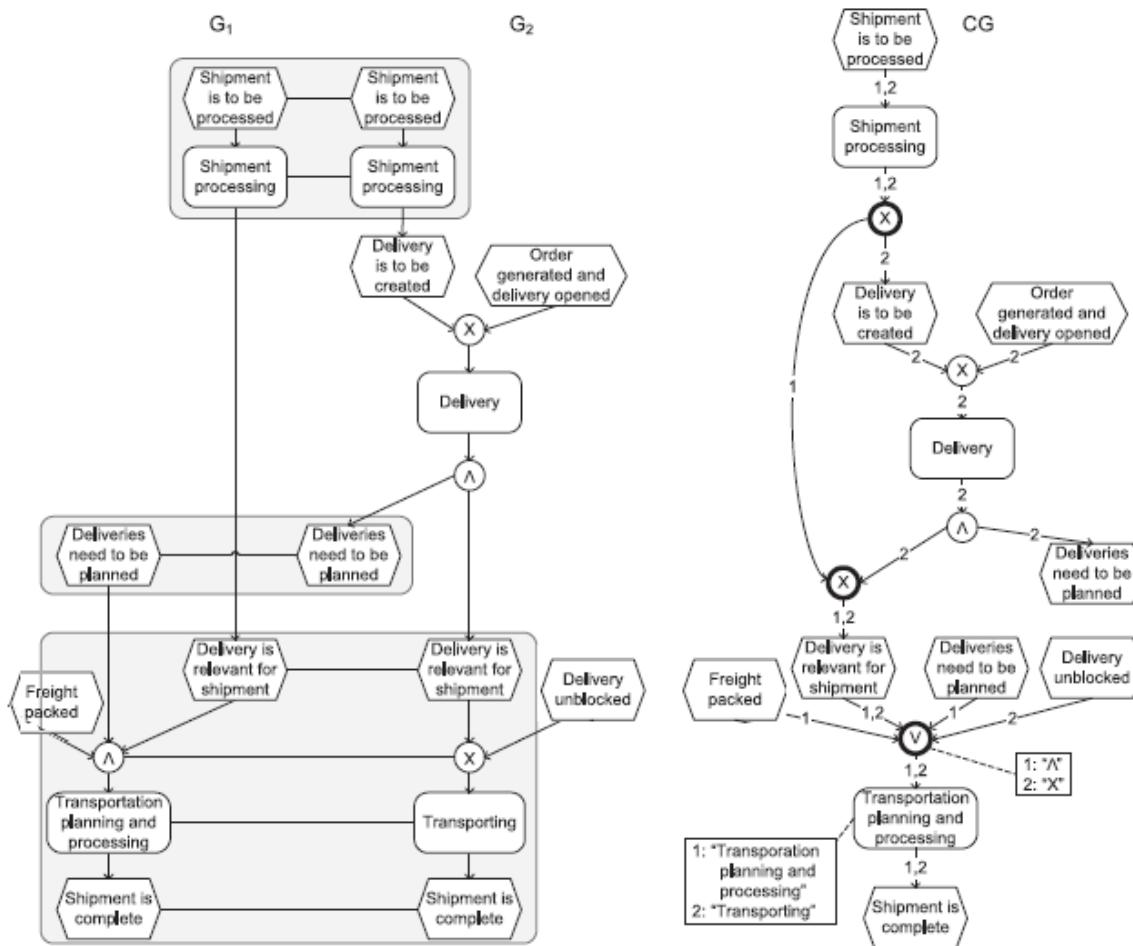
Assim como no trabalho de Gottschalk et al. (2008), essa abordagem também converte os modelos para grafos para realizar as comparações e integrações dos modelos de entrada. Essa conversão é realizada para um grafo de processo configurável, que permite a aplicação das técnicas de identificação de similaridade e integração de elementos. Depois de gerado o grafo composto, resultado da integração dos grafos configuráveis de origem, executa-se uma rotina de otimização para eliminar os nós emaranhados que surgiram na composição para, por fim, transformar esse grafo configurável resultante no grafo da notação BPG.

A similaridade entre os pares de rótulos de elementos dos grafos configuráveis de entrada é calculada pela combinação de duas medidas: a medida similaridade sintática e a medida de similaridade linguística (semântica). A primeira calcula a distância entre esses rótulos, isto é, pelo número de operações necessárias para transformar uma descrição em outra, através do cálculo de Levenshtein (1966). A segunda medida se baseia no pacote *Wordnet* de Pedersen et al. (2004) para encontrar os sinônimos das palavras. Em conjunto com essas medidas, *stop-words* (por exemplo, artigos e conjunções) e *word-endings* (por exemplo, “-ing”) são retirados desses rótulos para melhorar a eficácia da identificação de similaridades.

Um exemplo de composição desse trabalho pode ser visualizado na Figura 22. Nessa ilustração, dois dos grafos configuráveis são apresentados como entrada para a produção do grafo composto, resultado da integração, sendo que os elementos similares estão destacados pela associação entre G_1 e G_2 . Ainda nessa figura, identifica-se o acréscimo de conectores para representar o comportamento de ambos grafos de origem no grafo composto CG.

Uma ferramenta foi construída para avaliar o algoritmo proposto e um estudo de caso foi utilizado para avaliar a utilidade dessa ferramenta. Essa avaliação se deu em uma companhia de seguros que possui mais de 6000 modelos de processos de negócio ao longo de uma série de fusões e aquisições. Foi comparado o desempenho da ferramenta com o trabalho manual de analistas da empresa, na qual foram apontadas situações em que os analistas encontraram mais similaridades do que a ferramenta, uma vez que muitos modelos tinham pequenas omissões, diferenças de granularidade e diferentes convenções de nomenclaturas, embora a ferramenta tenha trazido ganhos de produtividade.

Figura 22: Composição de grafos configuráveis



Fonte: Adaptado de La Rosa et al. (2013)

3.6 Análise de Similaridade e Unificação de Processos de Negócio

Nesse trabalho, Santos (2014) propõe uma técnica para identificação de similaridades e integração de modelos de processos de negócio. Essa técnica compreende duas métricas principais: textual e estrutural. A métrica textual possui como característica a comparação dos rótulos dos elementos dos modelos e compreendem análise sintática, semântica, atributiva, posicional e contextual. Já a métrica estrutural possui como característica a análise da estrutura dos modelos, partir de sua transformação em grafos.

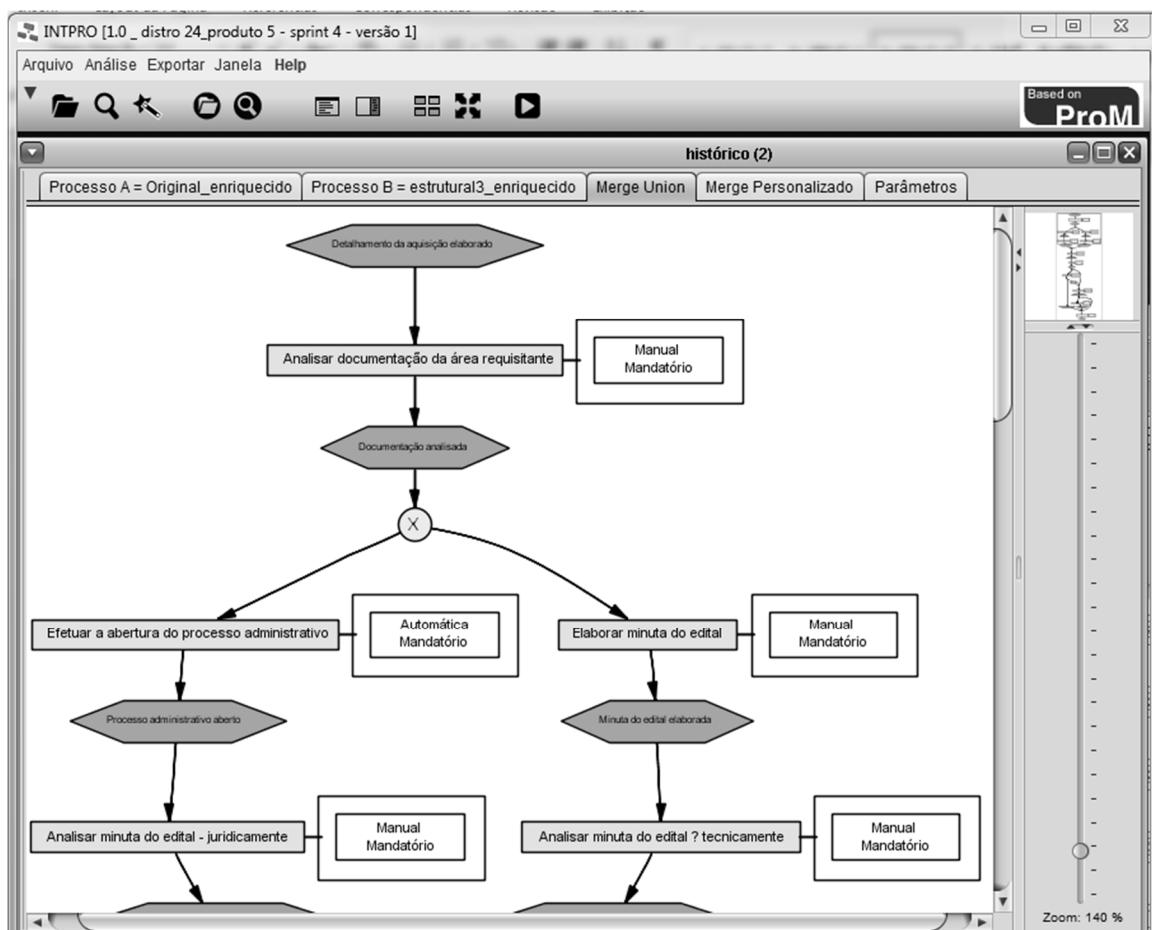
A proposta elaborada pelo autor se baseia na notação EPC devido sua larga utilização em sistemas integrados de gestão ERP. Para isso, utiliza o padrão *EPC Markup Language* (EPML) apresentado por Mendling e Nüttgens (2006), cuja sintaxe estabelece um conjunto de tags XML que representam funções, eventos e conectores dos modelos de processos de negócio desenhados nessa notação.

Um conjunto de definições foi elaborado pelo autor para garantir a correta utilização da técnica proposta. Essas definições estabelecem um modelo bem formado na notação EPC como requisito principal, assim como regras que garantem a correta aplicação das métricas utilizadas pela técnica proposta. Além disso, escores de similaridade são calculados com base em pesos definidos por analistas para cada uma das métricas apresentadas.

Assim como em outros trabalhos relacionados neste Capítulo, a similaridade sintática apresentada pelo autor utiliza o cálculo de distância entre *strings* de Levenshtein (1966). Para o cálculo de similaridade semântica, destaca-se a busca de termos sinônimos na base de dados lexical para a língua portuguesa TeP 2.0, apresentada por Maziero et al. (2008).

Uma ferramenta chamada INTPRO foi implementada para validar a proposta apresentada nesse trabalho. Essa ferramenta, desenvolvida com base no *framework* ProM, realiza a leitura de modelos produzidos na notação EPC, identifica as similaridades dos elementos dos modelos e produz o modelo composto, fruto da integração dos modelos de entrada. Um exemplo da tela de composição de modelos dessa ferramenta está ilustrado na Figura 23.

Figura 23: INTPRO – Tela de composição de modelos EPC



Fonte: Adaptado de Santos (2014)

Um estudo de caso foi realizado com dois cenários que compreendem um modelo de processos de negócio cada um, de uma empresa de energia. A composição foi realizada entre cópias de cada modelo, com pequenas alterações em cada um que envolvem modificações de dois tipos: textual e estrutural. Três tipos de experimentos foram realizados com relação ao índice de similaridade mínima, chamado de limiar nesse trabalho, que compreendem 0,5, 0,6 e 0,7 respectivamente, para caracterizar que os elementos comparados possam ser considerados similares.

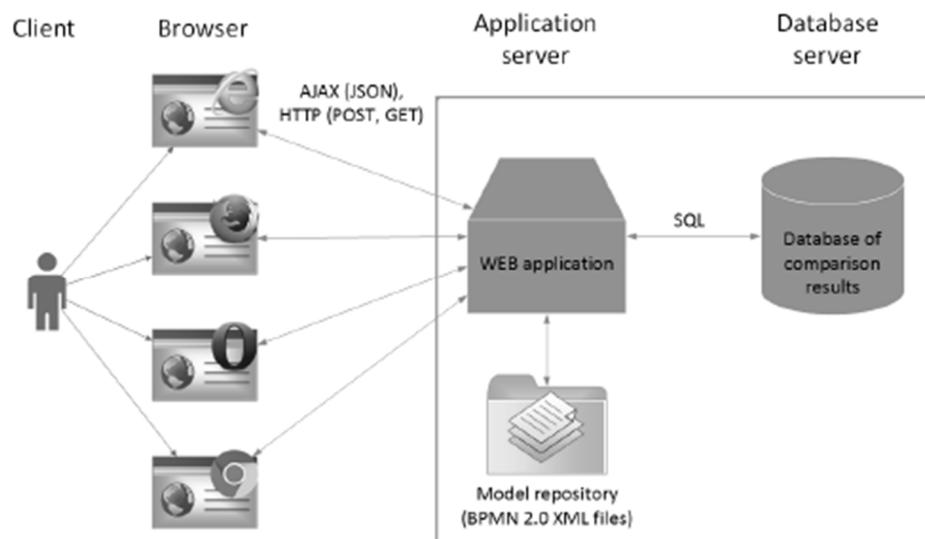
3.7 BPMNDiffViz – A Tool for BPMN Models Comparison

Nesse trabalho, Ivanov et al. (2015) apresentam uma ferramenta para encontrar e visualizar diferenças entre modelos de processos de negócio. Essa ferramenta realiza uma análise comparativa de modelos com base na estrutura de grafos e implementa o algoritmo A* para realizar a correspondência de elementos através de análise de similaridade sintática.

A ferramenta proposta possui uma interface Web e trabalha com a notação BPMN 2.0 que, segundo os autores, é o padrão industrial mais popular para a modelagem de processos de negócio. Com relação às funcionalidades da ferramenta, destacam-se a identificação visual explícita das diferenças através de cores que diferenciam os elementos similares e dados estatísticos que auxiliam na análise de discrepâncias.

A arquitetura da ferramenta é composta de um conjunto de aplicativos e frameworks que gerenciam a interface, a aplicação e a persistência, conforme ilustra a Figura 24. A arquitetura compreende servidor de aplicação Apache Tomcat 7.0⁶ para rodar aplicação, banco de dados PostgreSQL⁷ para armazenar os dados estatísticos, framework Spring MVC⁸ para gerenciar as camadas da aplicação e interface gráfica bpmn.io⁹ para visualizar os modelos em um navegador Web a partir de instruções Javascript (Json).

Figura 24: Arquitetura da ferramenta BPMNDiffViz



Fonte: Ivanov et al. (2015)

O resultado da comparação dos modelos é mostrado na Figura 25, onde é possível visualizar os elementos similares e as diferenças entre os modelos. Nesse resultado, os elementos correspondentes são destacados em azul, os elementos que devem ser excluídos são destacados em vermelho e, destacados em verde, os elementos que devem ser incluídos.

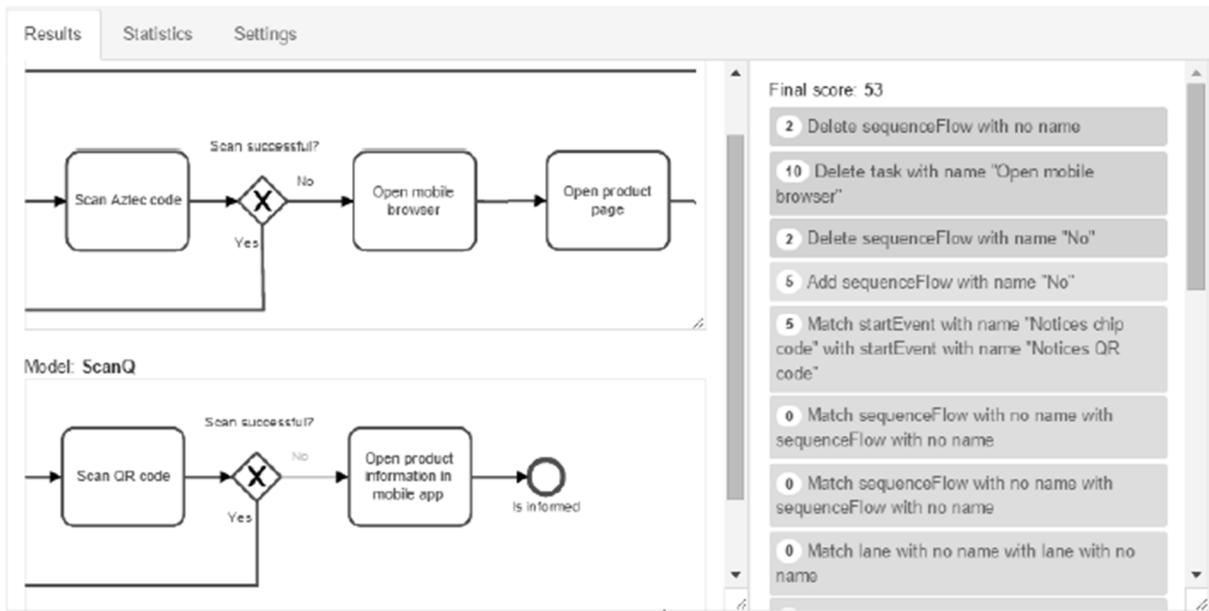
⁶ <http://tomcat.apache.org/>

⁷ <http://www.postgresql.org/>

⁸ <http://docs.spring.io/>

⁹ <https://bpmn.io/>

Figura 25: Lista de correspondências entre dois modelos



Fonte: Adaptado de Ivanov et al. (2015)

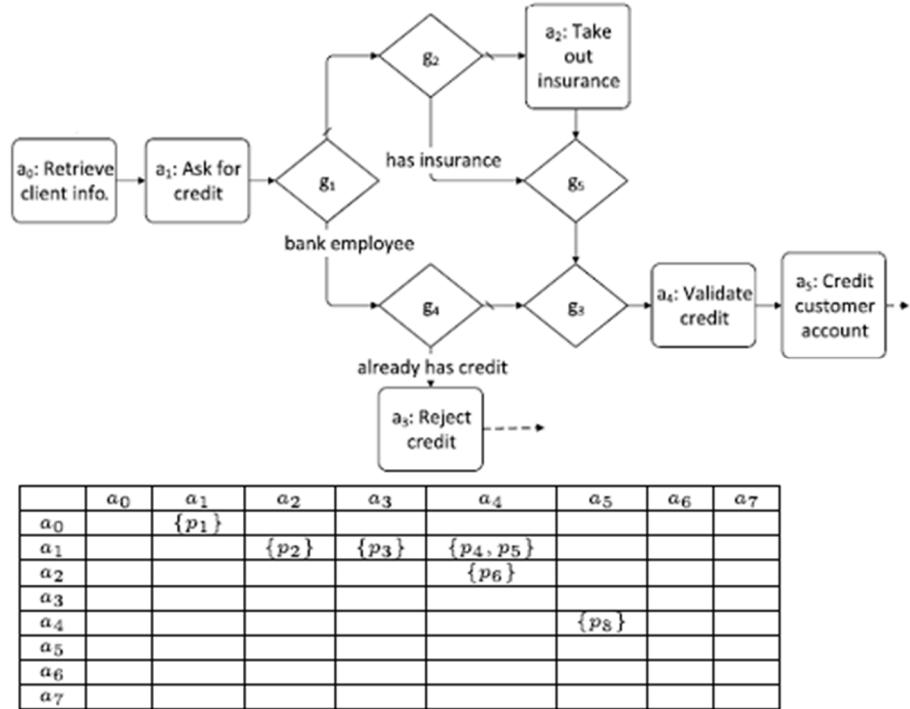
As estatísticas são mostradas em gráfico de pizza e compreendem os totais por tipo de elemento como, por exemplo, *tasks*, *gateways*, *events*, *sequenceFlows* além de informar o total de operações necessárias para transformar um modelo no outro, através das operações incluir, alterar e excluir elementos dos modelos. Por fim, testes foram realizados em vários modelos, cujo desempenho se mostrou satisfatório em modelos com até 30 elementos.

3.8 An Automatized Approach for Merging Business Process Fragments

Nesse trabalho, Zemni et al. (2016) apresentam uma ferramenta para integrar modelos de processos de negócio a partir de fragmentos de modelos. Essa ferramenta implementa uma técnica automática de composição de fragmentos baseada em uma matriz de caminhos como um modo de representar a integração dos modelos mapeados em grafos, com o objetivo de manter o comportamento inicial dos fragmentos originais.

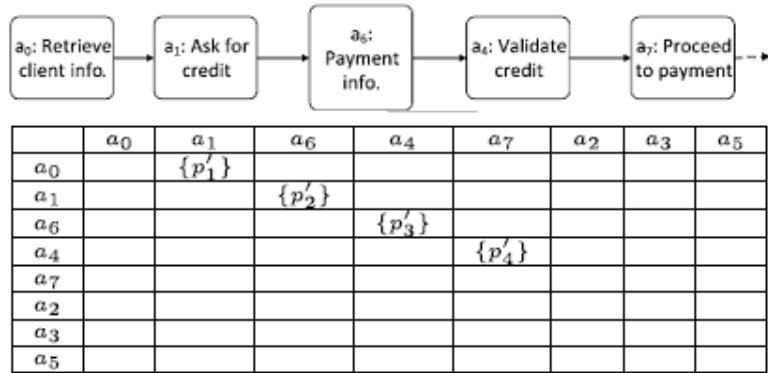
A proposta apresentada pelos autores utiliza uma notação simplificada que se baseia na BPMN, tratando apenas os elementos mais básicos como *tasks* e *gateways*. Um conjunto de definições foi estabelecido para especificar as regras de composição, bem como as características dos fragmentos de modelos, visando garantir a correta integração dos fragmentos de entrada. A conversão dos fragmentos de origem para matriz de caminhos, dos respectivos fragmentos F_1 e F_2 , são apresentadas pela Figura 26 e Figura 27.

Figura 26: Associação do fragmento F₁ de modelo para matriz



Fonte: Adaptado de Zemni et al. (2016)

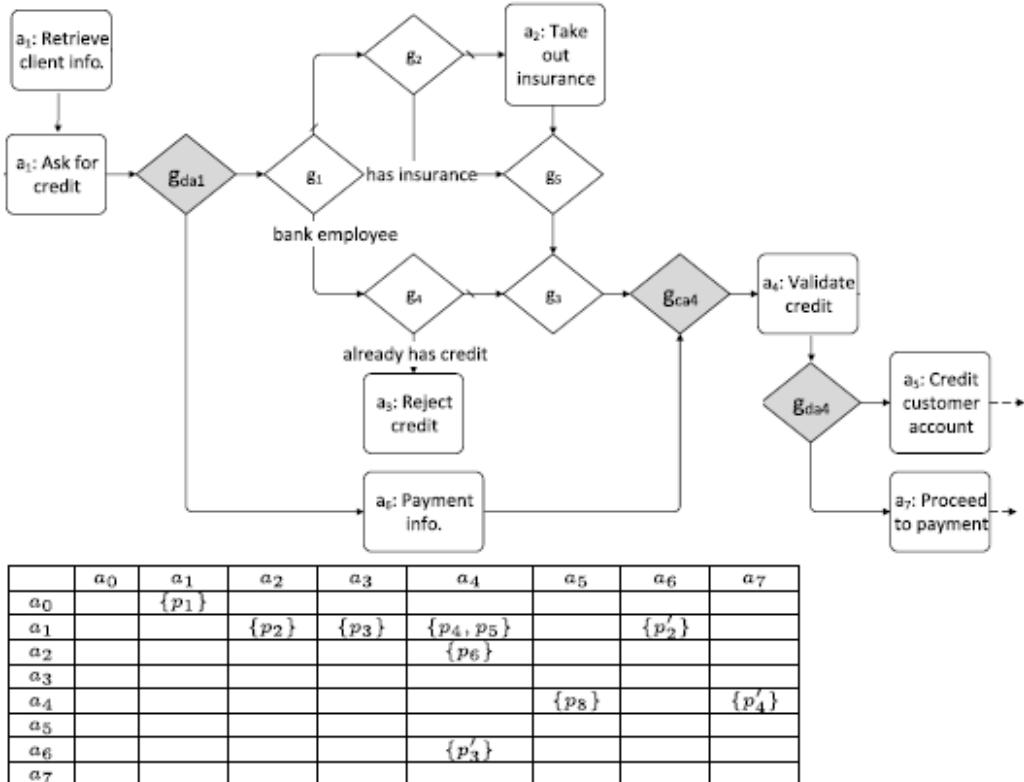
Figura 27: Associação do fragmento F₂ de modelo para matriz



Fonte: Adaptado de Zemni et al. (2016)

O resultado final da integração, apresentado pela Figura 28, ilustra a completa integração dos fragmentos de origem, de forma que o comportamento de ambos é preservado no fragmento composto, através da inclusão de diversos conectores (*gateways*). Ressalta-se que a técnica proposta pelos autores realiza análise comportamental de similaridade, deixando de lado a identificação de similaridade dos rótulos dos elementos.

Figura 28: INTPRO – Tela de composição de modelos EPC



Fonte: Adaptado de Zemni et al. (2016)

Experimentos foram realizados para verificar a corretude da abordagem apresentada pelos autores. Esses experimentos realizam a integração de fragmentos de modelos em uma biblioteca contendo 560 exemplares, com 8874 tasks diferentes, sendo que 1661 delas aparecem em mais de um fragmento. Além disso, foram realizados experimentos de integração de fragmentos com eles mesmos, para comprovar que o resultado final é idêntico aos fragmentos de origem. Por fim, nenhum detalhe a respeito da ferramenta implementada foi divulgado.

3.9 Comparação dos trabalhos relacionados

Visando complementar os estudos apresentados nos trabalhos relacionados neste Capítulo, um conjunto de características foi elaborado para avaliar como tais características se manifestam nesses trabalhos, bem como identificar lacunas e apontar oportunidades de pesquisa na área de identificação de similaridades e integração de modelos de processos de negócio. Para isso, as seguintes características foram analisadas nos trabalhos apresentados, conforme resumo apresentado na Tabela 4.

- **Similaridade sintática.** Analisa os elementos dos modelos pela descrição dos rótulos para identificar as similaridades. Essa análise deve permitir associar não apenas rótulos idênticos, mas também rótulos similares com descrição diferente. Essa é uma característica importante no contexto de análise de similaridade, uma vez que nem sempre os rótulos terão descrição idêntica das atividades, visto que cada empresa possui seu vocabulário corporativo próprio. Nos estudos apresentados, quatro abordagens, de um total de sete, atendem essa característica, sendo esse atendimento através do método de Levenshtein (1966). Esse método calcula o número de operações necessárias para

transformar uma *string* em outra, visando mensurar a diferença entre os rótulos dos elementos dos modelos.

- **Similaridade semântica.** Utiliza um dicionário de sinônimos para correlacionar os rótulos dos elementos nos modelos de entrada. Essa técnica permite que palavras diferentes que possuem o mesmo ou aproximadamente o mesmo significado sejam substituídas por sinônimos para aumentar o grau de identificação de similaridade. Nos trabalhos estudados, três abordagens fazem uso dessas técnicas através do *Wordnet* de Pedersen et al. (2004) e da base de dados lexical para a língua portuguesa TeP 2.0 de Maziero et al. (2008).
- **Similaridade contextual.** Essa característica leva em consideração o contexto em que estão inseridos os elementos, preservando o comportamento e a ordem das atividades dos modelos de entrada no modelo composto. Para isso, em alguns trabalhos, novos conectores (*AND*, *OR* e *XOR*) são incluídos no modelo produzido para garantir a preservação contextual. Nos estudos apresentados, seis abordagens utilizam essa técnica.
- **Similaridade estrutural.** Realiza uma análise estrutural dos modelos como um todo. Essa característica considera a estrutura geral dos modelos para verificar o grau de compatibilidade entre eles, normalmente calculada pela distância entre grafos, que caracteriza o número de ações para transformar um grafo em outro. Nesse quesito cinco abordagens atendem essa característica.
- **Notação BPMN 2.0.** Essa característica procura investigar se a abordagem atende a especificação da notação BPMN 2.0 conforme o padrão estabelecido pela OMG-BPMN (2013). Essa notação, nos dias atuais, tem sido largamente conhecida e utilizada nas organizações e tem se tornado o padrão para modelagem BPM. Nessa condição, apenas duas abordagens mostraram que trabalham com essa notação, sendo que uma delas apresentou uma versão simplificada, referenciada pela versão 1.2.
- **Composição de modelos.** A composição de modelos se caracteriza pela união de dois modelos de entrada, M_A e M_B , para criar o modelo composto M_C , que contém a união dos modelos de origem. Essa é a característica principal buscada nos trabalhos relacionados, uma vez que tem relação direta com o objeto deste estudo. Nesse quesito, cinco abordagens apresentaram algoritmos e/ou ferramentas de composição de modelos de processos de negócio, sendo que o único trabalho que não faz integração de modelos, apresentou uma boa base teórica sobre identificação de similaridades, que também é alvo desta pesquisa.
- **Ferramenta de composição.** Quatro das sete abordagens estudadas implementaram uma ferramenta (*plug-in* ou *framework*) para realizar a integração de modelos de processos de negócio. Entretanto, todas as ferramentas apresentadas realizam a composição de forma automática, sem permitir que o analista possa interagir com a ferramenta em tempo de execução, para decidir sobre os conflitos com base nas regras de negócio que devem orientar o comportamento do modelo produzido.

Ao realizar uma análise das características dos trabalhos relacionados, conforme descrito na Tabela 4, observa-se que as técnicas de identificação de similaridade sintática, semântica, contextual e estrutural estão presentes em vários deles, especialmente as que possuem relação com a comparação de contexto e de estrutura entre modelos. Isso ocorre principalmente porque esses trabalhos atuam na conversão dos modelos para grafos

configuráveis, que visam manter o comportamento dos modelos de origem no modelo produzido.

Apesar de a notação BPMN 2.0 ser considerada atualmente um padrão de modelagem de processos, apenas um trabalho implementa a identificação de similaridade e nenhum implementa a composição de modelos com base nessa notação, uma vez que utilizam de notação genérica baseada elementos simplificados como funções, atividades e conectores, que representam notações mais simplificadas como a EPC, presente em vários trabalhos.

Por fim, as ferramentas apresentadas não são amigáveis para a prática de composição de modelos, uma vez que realizam a composição de modelos de forma automática visando manter o comportamento de ambos os modelos de origem. Essa característica impede a interação do analista no processo de composição para decidir sobre os conflitos com base nas regras de negócio.

Tabela 4: Comparação dos trabalhos relacionados

Abordagem	Similaridade Sintática	Similaridade Semântica	Similaridade Contextual	Similaridade Estrutural	Notação BPMN 2.0	Composição de modelos	Ferramenta de composição
Sun et al. (2006)	-	-	+	+	-	+	-
Gottschalk et al. (2008)	-	-	+	-	-	+	+
Dijkman et al. (2011)	+	+	+	+	-	-	-
La Rosa et al. (2013)	+	+	+	+	-	+	+
Santos (2014)	+	+	+	-	-	+	+
Ivanov et al. (2015)	+	-	-	+	+	-	-
Zemni et al. (2016)	-	-	+	+	~	+	+

Notas: (+) Atende, (-) Não atende, (~) Atende parcialmente

Fonte: Elaborado pelo autor

3.10 Oportunidades de pesquisa

Após análise dos estudos apresentados neste Capítulo, identificam-se oportunidades de pesquisa que podem agregar valor ao processo de composição de modelos de processos de negócio. Essas oportunidades de pesquisa são exploradas como possíveis contribuições científicas desta pesquisa e levam em consideração as características e fragilidades identificadas nos trabalhos relacionados, bem como novas formas de aplicação das estratégias apresentadas, conforme segue:

- **Comparação de modelos BPMN 2.0:** criar um método de reconhecimento de similaridades, entre elementos de modelos de processos de negócios sensíveis à notação BPMN 2.0, que permita aplicar múltiplas estratégias de comparação e identificação de elementos equivalentes, baseando-se em características sintáticas, semânticas e contextuais. Esse método de reconhecimento de similaridades visa suprir a lacuna existente em relação ao uso da notação BPMN 2.0 em conjunto com métricas de similaridade sintática e semântica.
- **Integração de modelos BPMN 2.0:** criar um método de integração de modelos de processos de negócio que atenda à especificação BPMN 2.0. Esse método leva em

consideração os métodos de identificação de similaridade para associar elementos comuns nos modelos para realizar a integração deles. O objetivo de criar esse método é preencher uma lacuna identificada nos trabalhos relacionados, uma vez que nenhuma abordagem estudada realiza composição de modelos baseados na especificação BPMN 2.0.

- **Ferramenta de composição semiautomática:** implementar uma ferramenta que aplique os métodos de identificação de similaridade e de integração de modelos para produzir modelos de processos de negócio de forma semiautomática. Ao criar essa ferramenta, espera-se pôr em prática os métodos apresentados, bem como disponibilizar um instrumento que agilize o trabalho de analistas que precisam compor modelos, visando reduzir o esforço e a propensão a erros, proporcionando economia de tempo e de custos.
- **Conhecimento empírico sobre composição de modelos:** realizar experimentos sobre composição de modelos de processo de negócio permite mensurar o esforço que analistas investem para integrar modelos, bem como quantificar os erros que são apresentados nos modelos produzidos. Ao realizar esses experimentos, espera-se produzir dados que podem servir de objeto de estudos sobre o processo composição de modelos, realizado tanto de forma manual como com o apoio de uma ferramenta que auxilie nessa tarefa.

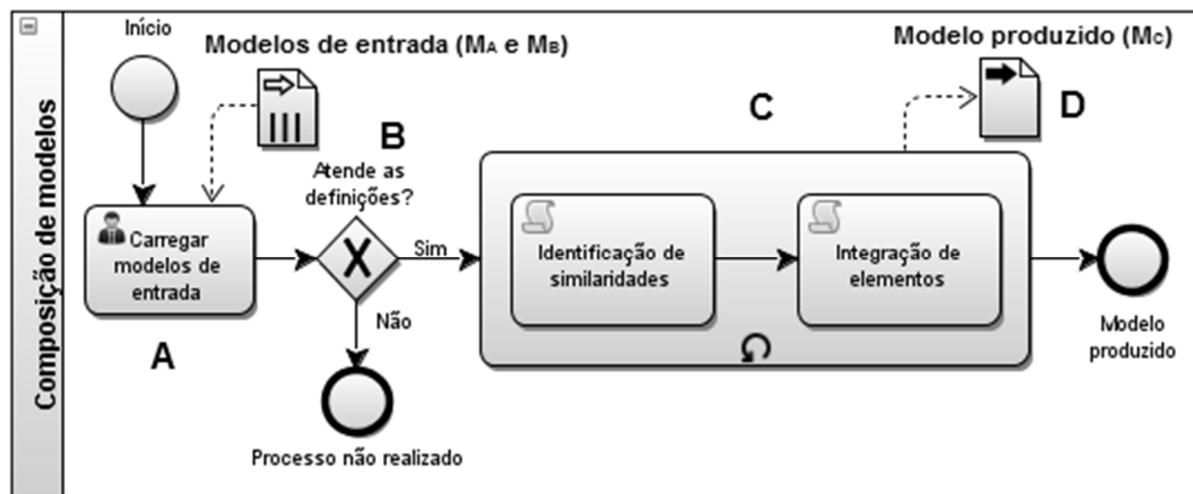
4 TÉCNICA DE COMPOSIÇÃO DE MODELOS

Para explorar as oportunidades de pesquisa elencadas na Seção 3.9, assim como atender ao primeiro objetivo específico enumerado na Seção 1.3, propõe-se uma técnica de composição sensível à semântica dos modelos de processos de negócio, que é capaz de identificar as similaridades entre os elementos dos modelos de entrada, bem como integrá-los. Para apresentar essa técnica, este Capítulo foi organizado da seguinte forma. A Seção 4.1 apresenta uma visão geral da técnica de composição. A Seção 4.2 apresenta os conceitos básicos necessários para o entendimento dessa técnica. A Seção 4.3 descreve como a técnica identifica a similaridade entre os elementos dos modelos de entrada. Por fim, a Seção 4.4 detalha como os elementos dos modelos de entrada, identificados como similares, são integrados.

4.1 Visão geral da técnica de composição

A técnica de composição proposta neste trabalho é suportada por diversos procedimentos que atuam em conjunto. Esses procedimentos realizam o carregamento dos modelos de entrada, a verificação do atendimento das definições de validação, a identificação e integração de elementos e a exportação do modelo composto, conforme ilustra a Figura 29.

Figura 29: Visão geral da técnica de composição



Fonte: Elaborado pelo autor

A seguir são explicados os procedimentos que compõem a técnica de composição proposta.

- (A) **Carregamento dos modelos de entrada.** O carregamento dos modelos de entrada ocorre pela leitura e interpretação desses modelos, sendo que os elementos de cada modelo são lidos e traduzidos para realizar a validação de seu conteúdo.
- (B) **Validação dos modelos de entrada.** A validação dos modelos de entrada é realizada a partir do atendimento de um conjunto de definições criado especificamente para garantir a correta aplicação da técnica. Essas definições atuam como requisitos que devem ser cumpridos para que os modelos de entrada possam ser integrados corretamente.
- (C) **Composição dos modelos de entrada.** A composição é formada por duas atividades centrais: identificação de similaridades e integração de elementos. Enquanto

a primeira atividade visa identificar a similaridade entre os elementos dos modelos, a segunda foca em integrar os elementos definidos como similares para gerar o modelo composto.

- **(D) Exportação do modelo composto.** A produção do modelo composto é a última etapa da técnica de composição, que apresenta como resultado o modelo produzido que comprehende a integração dos modelos de entrada.

Em resumo, a técnica lê e avalia dois modelos de entrada, executa procedimentos de identificação de similaridades e integração de elementos, visando e produzir o modelo composto. As definições que estabelecem os critérios de validação e os procedimentos de identificação e integração de elementos são apresentados a seguir, ao contrário dos procedimentos de leitura e exportação que serão detalhados no Capítulo 5, que detalha a ferramenta Common que foi implementada para pôr em prática a técnica apresentada neste Capítulo.

4.2 Definições de composição de modelos

As definições de composição, criadas neste trabalho, estabelecem os critérios gerais baseados na especificação *Business Process Definition Metamodel* (BPDM), definida por (OMG-BPDM, 2008), para regular os procedimentos de composição de modelos de processos de negócio da técnica apresentada neste Capítulo. Essas definições são baseadas em critérios ou comportamentos que devem ser atendidos, de forma que seja possível aplicar a técnica corretamente para alcançar os objetivos propostos, conforme definido a seguir.

4.2.1 Definição 1 – Conjunto de modelos

Seja M um conjunto de modelos de processos de negócio BPMN 2.0, conforme especificação da OMG-BPMN (2013), define-se:

- Os modelos de entrada M_A e M_B devem pertencer ao conjunto M , tal que

$$M_A \in M \wedge M_B \in M$$

- Realizar a composição de um modelo com ele mesmo ou de dois modelos similarmente idênticos, retornará o próprio modelo.

$$M_A \cup M_B, \text{ tal que } M_A = M_B \rightarrow M_A \text{ ou } M_B$$

- A ordem de integração dos modelos não deve importar. Isso significa que realizar a composição entre M_A e M_B vai gerar aproximadamente o mesmo resultado que compor M_B e M_A , ressalvadas pequenas diferenças de nomenclatura, que não descaracterizam a similaridade.

$$M_A \cup M_B \cong M_B \cup M_A$$

- Para realizar a composição de mais de dois modelos, primeiro integra-se os dois primeiros, depois utiliza-se o resultado dessa integração para fundir com o terceiro modelo, e assim por diante.

$$((M_A \cup M_B) \cup M_C) \cup M_D \dots$$

4.2.2 Definição 2 – Um modelo BPMN 2.0

Considerando um modelo de processo de negócio M , obtém-se as seguintes considerações (BRAGHETTO, 2011):

- $M = (F, C, t, r, a) \rightarrow$ **modelo de processo de negócio.** M é o conjunto finito de elementos que pertencem ao modelo.
- $F \rightarrow$ **conjunto finito de elementos.** F é o conjunto finito de itens do modelo. Esses itens compreendem *tasks*, *events*, *gateways* e *data* (dados de entrada, saída e de armazenamento), que representam os elementos do modelo.
- $C \subseteq (F \times F) \rightarrow$ **conjunto finito de conexões.** C é o conjunto finito de objetos de conexão do modelo. Esses elementos compreendem fluxo de sequência, fluxo de mensagem e fluxo associativo, que representam as arestas do modelo.
- $t \subseteq (F \cup C) \rightarrow$ **associação entre elementos e tipos.** t é o tipo do elemento que caracteriza a especialização de F como, por exemplo, um evento de início, um evento de fim, um *gateway* paralelo ou uma atividade.
- $r \subseteq (F \cup C) \rightarrow$ **associação entre elementos e rótulos.** r é o rótulo do elemento que caracteriza a descrição de F ou C como, por exemplo, descrição de uma atividade, de um *gateway* de decisão ou de um fluxo de sequência nomeado.
- $a \subseteq (F \cup C) \rightarrow$ **associação entre elementos e atores.** a é o ator responsável pela realização de F .
- Uma conexão C em M é um par (c_1, c_2) , onde c_1 e $c_2 \in C$ e indicam o fluxo de sequência de c_1 para c_2 .
- Um elemento F em M é um elemento que recebe uma ou mais entradas ou uma ou mais saídas, onde $f_1 \in F$, tal que $\bullet f_1$ indica um fluxo de entrada e $f_2 \bullet$ um fluxo de saída.
- O caminho em um modelo M é representado pela existência de uma sequência entre os elementos F desse modelo, sendo que cada elemento é ligado por uma ou mais conexões.

4.2.3 Definição 3 – Modelo BPMN bem formado

Braghetto (2011) cita que a notação BPMN 2.0 possui uma semântica operacional rica e complexa que, se mal aplicada, pode levar a construção de modelos com mais de uma interpretação, podendo gerar impasse no seu entendimento. Para minimizar esses efeitos e evitar prejuízos ao processo de composição, o conjunto de premissas, listado a seguir, foi estabelecido com base na OMG-BPDM (2008) para garantir que um modelo de processo de negócio seja bem formado, garantindo assim a correta aplicação da técnica de composição.

- Possuir um único evento de início.
- Um evento de início deve ter apenas uma conexão de saída.
- Um evento de início não pode ter conexões de entrada.
- Possuir um ou mais eventos de fim.
- Um evento de fim não deve ter conexão de saída.
- Um evento de fim deve ter uma ou mais conexões de entradas.
- Nenhum elemento pode estar solto no modelo (sem conexões de entrada ou saída).

- Um elemento fluxo de decisão deve ter uma conexão de entrada e duas ou mais conexões de saída.
- Um elemento fluxo paralelo deve ter uma conexão de entrada e duas ou mais conexões de saída, no caso de conexão de divergência; ou duas ou mais conexões de entrada e uma conexão de saída, no caso de convergência.
- Todo elemento de conexão deve ter um elemento de origem e um elemento de destino.
- Todo artefato (produto de entrada, saída ou armazenamento) deve estar vinculado a um elemento pai no modelo.
- O modelo não deve ter elementos em situações inatingíveis.
- O processo deve seguir de modo que tenha um início e pelo menos um fim.

4.2.4 Definição 4 – Comparação de elementos

A comparação de elementos entre os modelos de entrada é fator crucial para a realizar a composição de modelos, uma vez que é a partir dessa comparação que os elementos podem ser identificados como similares. Para isso, algumas regras foram construídas visando estabelecer critérios de comparação, conforme segue:

- **Regra 1: tipo do elemento.** Os elementos devem ser comparados com elementos do mesmo tipo, conforme a especificação dos tipos listada na Seção 2.5. Essa regra permite restringir o espaço de busca por elementos similares nos modelos, uma vez que considera apenas elementos passíveis de comparação.
 $\forall f_1 \in F \text{ em } M_A \wedge f_2 \in F \text{ em } M_B, \text{ tal que } tipo(f_1) = tipo(f_2)$
- **Regra 2: elemento sem rótulo.** Os elementos que não possuem rótulo como, por exemplo, *parallel gateway* e *sequence flow*, devem ser comparados através da similaridade contextual (elementos que antecedem ou sucedem eles nos modelos).
- **Regra 3: elemento similar.** Elementos similares, identificados ou não por exatidão, devem ser sobrepostos no modelo composto.
- **Regra 4: elemento não similar.** Elementos não similares, isto é, que não foram identificados como similares, podem ser incluídos ou não no modelo composto, conforme estiver estabelecido pelas regras de negócio que guiarão o analista durante o processo de composição.

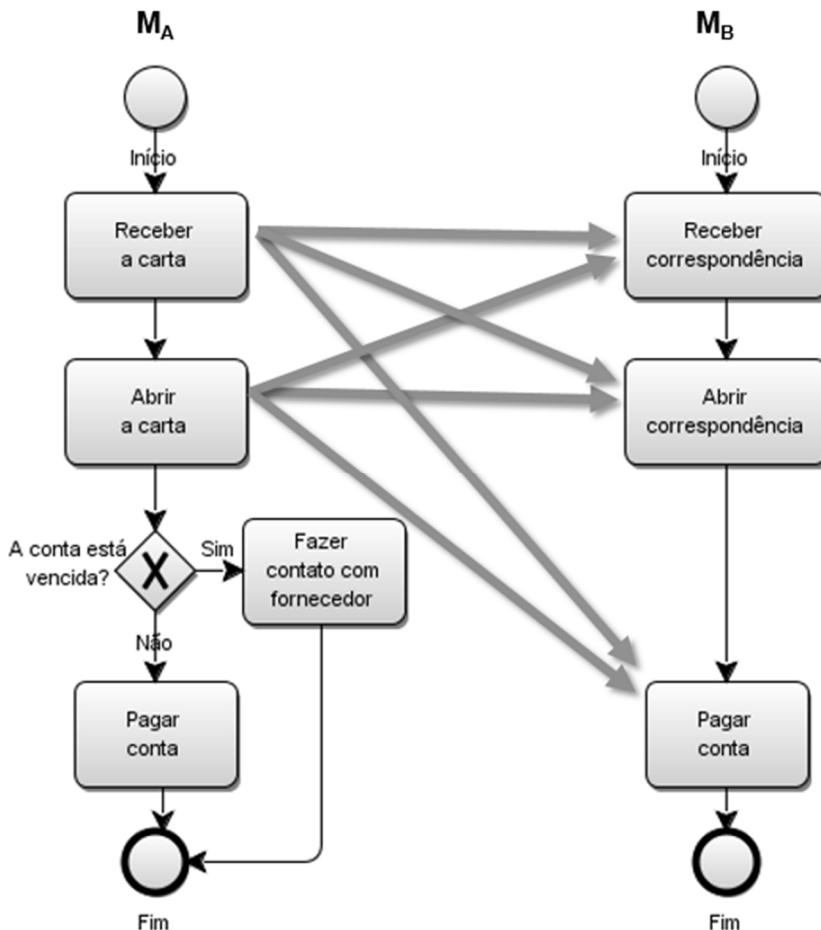
4.3 Identificação de similaridades

Comparar os elementos dos modelos de entrada M_A e M_B é o primeiro passo do processo de identificação das similaridades. O procedimento inicia com uma varredura nos modelos visando identificar, para cada elemento, os seguintes atributos: tipo, rótulo, entradas, saídas e artefatos vinculados. Dessa forma, para cada elemento de um modelo, o procedimento compara com todos os elementos do outro, obedecendo necessariamente aos critérios estabelecidos nas definições da Seção 4.2.

Essa atividade ocorre através da comparação dos elementos pela associação $A \times B$, onde A representa os elementos do modelo M_A e B representa os elementos do modelo M_B . Os elementos de M_A são comparados com os elementos de M_B através de um produto cartesiano,

de modo que todas as combinações possíveis sejam testadas. A Figura 30 apresenta um exemplo da comparação de elementos, onde os elementos do tipo *task* “Receber a carta” e “Abrir a carta” de M_A são comparados com todos os elementos tipo *task* de M_B .

Figura 30: Comparação de elementos entre M_A e M_B



Fonte: Elaborado pelo autor

Muito embora a busca de elementos nos modelos seja uma atividade simples, que pode ser resolvida com algoritmos conhecidos na literatura como *Depth-First Search* (DFS) explicado por Cormen et al. (2009), identificar corretamente as similaridades entre elementos de modelos distintos pode ser considerado uma atividade desafiante, citam Dijkman et al. (2011), uma vez que as estruturas e os rótulos desses elementos usualmente não são iguais, exigindo a aplicação de estratégias que facilitem a identificação das similaridades existentes.

Para mitigar esse desafio, este trabalho propõe as seguintes estratégias de identificação de similaridades que foram baseadas nos estudos apresentados no Capítulo 3. Entre os tipos de identificação de similaridade estudados, optou-se por descartar as similaridades estrutural e contextual por serem mais adequadas para comparação de modelos como um todo, pela transformação dos modelos em grafos.

4.3.1 Estratégia de Similaridade Semântica (ESE)

A similaridade semântica considera as diferentes formas de descrever os rótulos dos elementos com base nos mesmos sentidos. Ou seja, devido às diferenças de vocabulário

existentes entre equipes de modelagem, muitos elementos similares são escritos com palavras diferentes, mas que possuem o mesmo significado (LA ROSA et al., 2013). Para resolver esse problema, esta estratégia utiliza um dicionário de sinônimos para procurar e identificar palavras que tenham o mesmo ou aproximadamente o mesmo significado, sem que haja prejuízo ou alteração de sentido.

O dicionário de sinônimos utilizado neste trabalho é o WordNet.Br, criado por Dias-da-Silva et al. (2005), que possui um conjunto rico de verbos, substantivos e adjetivos escritos em português do Brasil. O WordNet.Br possui cerca de 44 mil palavras agrupadas em 18 mil synsets, que representam a palavra principal e seus sinônimos, sendo que esses dados podem ser acessados através de diversas APIs criadas especialmente para isso.

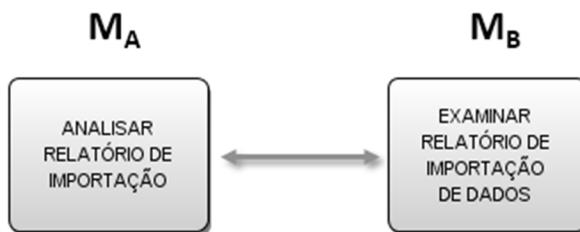
Para acessar o WordNet.Br, esta estratégia utiliza a API JWN-Br desenvolvida por Oliveira e Roman (2013). Essa API, distribuída gratuitamente sob licença GPL, permite encontrar sinônimos de forma simples e rápida, através métodos específicos como *getWords*, *getSynsets* e *getSinonymous*.

Além disso, essa API permite ampliar a gama de palavras do dicionário a partir do uso de métodos de inclusão como, por exemplo, *addWord* e *setSinonymousPointer*. Essa facilidade possibilita que novas palavras sejam adicionadas à base de dados local do WordNet.Br, de modo que palavras específicas, de determinados contextos, possam ser associadas. Exemplo disso são as palavras *mala*, *recipiente* e *caixeta* que não estão originalmente associadas no WordNet.Br, mas que são usadas como sinônimos para caracterizar invólucro para transporte de objetos postais, nos Correios.

Após a substituição dos sinônimos, retiram-se as preposições e as conjunções dos rótulos dos elementos de cada modelo que estão sendo comparados. Essa retirada faz com que apenas as palavras importantes para o entendimento do processo sejam utilizadas na comparação. Além disso, palavras acentuadas e não acentuadas e caracteres maiúsculos e minúsculos são identificados naturalmente, visto que esta estratégia realiza a conversão em tempo de execução, aumentando ainda mais a identificação das similaridades.

No exemplo ilustrado pela Figura 31, as conjunções “de” são retiradas do conjunto de palavras que serão comparadas nos elementos dos modelos, ficando apenas “Analizar relatório importação” de M_A e “Examinar relatório importação dados” de M_B . Além disso, a estratégia deve ser capaz de identificar que as palavras “Analizar” e “Examinar” são sinônimas e, portanto, possuem o mesmo significado.

Figura 31: Comparação de elementos nos modelos M_A e M_B



Fonte: Elaborado pelo autor

Assim que os rótulos foram preparados, o cálculo da Estratégia de Similaridade Semântica é realizado pela Equação 1, descrita a seguir:

$$ESE = \frac{s}{n} \quad (1)$$

onde:

s é o número de palavras iguais ou similares entre os elementos comparados;

n é o número total de palavras do rótulo com maior número de palavras, tal que $n > 0$.

O escore a ser calculado por esta estratégia compreende um valor entre 0 e 1, onde 0 indica que nenhuma palavra é igual, mesmo após a identificação dos sinônimos e 1 quando todas as palavras dos rótulos são iguais ou similares. No exemplo ilustrado pela Figura 13, o valor de ESE é $3 / 4 = 0,75$, referente à comparação dos seguintes rótulos $r_A = \{\text{analisar, relatório, importação}\}$ e $r_B = \{\text{examinar, relatório, importação, dados}\}$, sendo que 3 palavras do rótulo r_A foram associadas com outras 3 palavras do rótulo r_B , que possui 4. O índice 0,75 calculado indica que os rótulos possuem 75% de similaridade.

4.3.2 Estratégia de Similaridade Sintática (ESI)

Esta estratégia considera a sintaxe dos rótulos dos elementos que estão sendo comparados. Apesar de também usar os rótulos, como a Estratégia de Similaridade Semântica, utiliza-os de outra maneira, com o objetivo de identificar e calcular a diferença sintática entre eles, através da distância entre *strings*.

A distância entre *strings*, obtida pelo cálculo de Levenshtein (2016), calcula o número de operações necessárias para transformar uma descrição em outra, identificando assim o quanto distante estão, em termos sintáticos, os elementos que estão sendo comparados. Além disso, entende-se que aplicar essa técnica em conjunto com o uso do dicionário de sinônimos e a retirada de conjunções e preposições, utilizados na ESE, pode ser mais vantajoso, visto que a distância entre as palavras sinônimas, mas diferentes, será sempre maior que entre palavras iguais.

Para a Estratégia de Similaridade Sintática, o cálculo é realizado pela Equação 2 descrita a seguir:

$$ESI = 1 - \frac{d \cdot 2}{l + r} \quad (2)$$

onde,

d é o número de operações (inclusões, exclusões e substituições) que devem ser realizadas para transformar um rótulo em outro, entre os elementos comparados, a partir do cálculo de Levenshtein (1966);

l é o número total de caracteres no rótulo do elemento de M_A ;

r é o número total de caracteres do rótulo do elemento de M_B , tal que $l + r > 0$.

O escore a ser calculado por esta estratégia compreende um valor de 0 a 1, onde 0 indica que o número de operações para transformar um rótulo no outro é maior que a média do tamanho dos rótulos, e 1 quando os rótulos são iguais e nenhuma operação é necessária para transformar um rótulo no outro.

Tomando como exemplo a Figura 31, cujos conjuntos de palavras dos rótulos de M_A e M_B são, respectivamente, $r_A = \{\text{analisar, relatório, importação}\}$ e $r_B = \{\text{examinar, relatório, importação, dados}\}$, o valor de ESI é $1 - ((6 \cdot 2) / (29 + 35)) = 0,81$, referente as 6 operações calculadas pela distância entre *strings* e ao comprimento dos dois rótulos, com 29 e 35 caracteres respectivamente. O índice 0,81 calculado indica que os rótulos possuem 81% de similaridade.

4.4 Integração de elementos

Este procedimento visa realizar a integração dos elementos dos modelos de entrada. Para isso, capture os índices calculados pelas estratégias ESE e ESI, calcula o Escore Final de Similaridade (EFS) e realiza a integração dos elementos no modelo composto.

Dessa forma, o procedimento de integração realiza a unificação desses escores em um único valor, que contém a média aritmética ponderada dos escores individuais ESE e ESI, conforme escrito pela Equação 3:

$$EFS = \frac{(p1 \cdot ESE) + (p2 \cdot ESI)}{p1 + p2} \quad (3)$$

onde,

$p1$ é o peso de ESE;

$p2$ é o peso de ESI, tal que $p1 + p2 > 0$.

Os pesos $p1$ e $p2$ podem ser configurados de modo que o usuário poderá estabelecer pesos diferentes para cada uma das estratégias, baseando-se em experimentos empíricos de composição de modelos. Essa configuração permite desativar uma estratégia, caso seu peso seja zero para usar apenas a outra estratégia ou atribuir valores diferentes, caso pretenda impor maior importância para determinada estratégia. Tomando como exemplo a Figura 31, o escore final EFS é calculado conforme segue: $EFS = \frac{(1 \cdot 0,75) + (2 \cdot 0,81)}{1+2} = 0,79$, onde $p1 = 1$, $p2 = 2$, $ESE = 0,75$ e $ESI = 0,81$.

Como mencionado anteriormente, EFS indica o escore final de similaridade entre dois elementos após a aplicação das estratégias ESE e ESI. Contudo, é necessário que esse escore atinja um índice mínimo para que se possa considerar que os elementos comparados são similares. Para isso, esse escore é comparado com o Valor Limite de Similaridade (VSL).

O VLS é um índice fornecido pelo usuário que estabelece um valor mínimo a ser superado pelo EFS para que tais elementos sejam considerados similares. O valor do VLS não deve ser inferior a 0,5, para evitar que elementos diferentes ou com pouca similaridade sejam considerados similares. Assim, neste trabalho utilizou-se o VLS igual a 0,75 para estabelecer uma margem de identificação de similaridade que compreende valores pertencentes a índices utilizados na literatura como, por exemplo, Farias et al. (2014), Santos (2014) e Dijkman et al. (2011), que utilizam índices de 0,8, 0,7 e 0,7 respectivamente.

Com base nisso e, por ser uma técnica de composição semiautomática, a comparação dos índices de similaridade atua da seguinte maneira:

Considerando VLS = 0,75;

- **EFS = 1.** Elementos 100% similares que entram automaticamente no modelo composto.
- **EFS $\geq 0,75$ e $\leq 0,99$.** Elementos com alguma similaridade que dependem de aprovação do usuário para entrar no modelo composto. Nessa situação, podem se apresentar vários elementos para o usuário, que se baseará nas regras de negócio para decidir qual é o elemento indicado.
- **EFS < 0,75.** Elementos com pouca ou nenhuma similaridade. Nesses casos, o procedimento considera que os elementos não são similares e então solicita aprovação do usuário para que seja incluído no modelo composto, sem apontar possíveis elementos similares.

No caso do exemplo da Figura 31, EFS ficou com o índice de 0,79, que supera o VLS estabelecido em 0,75. Dessa forma, é apontado para o usuário como elementos possivelmente similares, aguardando então aprovação do mesmo para inclusão no modelo composto. A Tabela 5 apresenta o resumo do procedimento de identificação de similaridade entre os elementos dos modelos.

Tabela 5: Resumo da similaridade entre elementos (VLS = 0,75)

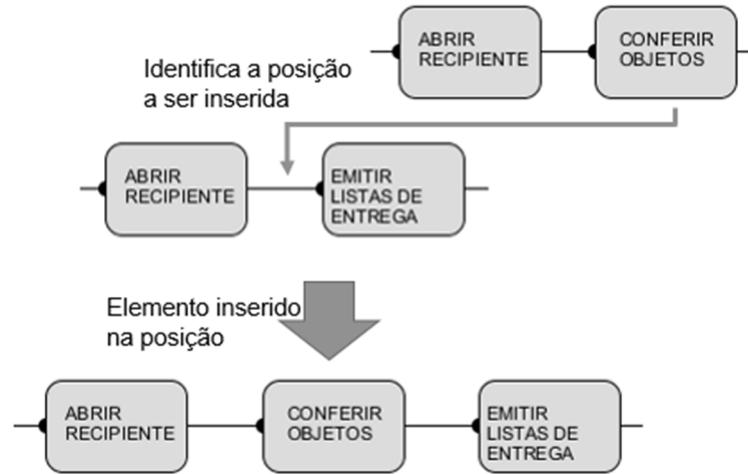
Índices	Forma de inclusão no modelo composto
EFS = 1	Automática. Inclui o elemento automaticamente no modelo composto
EFS $\geq 0,75$ e $\leq 0,99$	Semiautomática. Apresenta lista para seleção de elementos similares
EFS < 0,75	Semiautomática. Solicita interferência do usuário para incluir no modelo composto

Fonte: Elaborado pelo autor

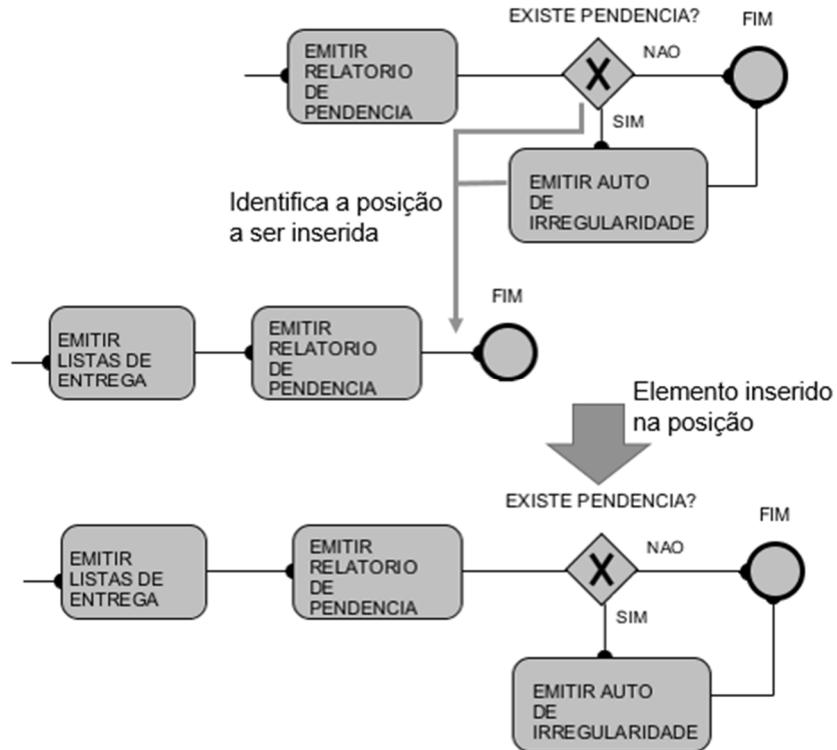
Assim que a forma de inclusão é identificada, conforme mencionado no resumo ilustrado pela Tabela 5, há a necessidade de inserir esse elemento identificado no modelo composto. Essa inclusão pode ocorrer de diversas formas, como segue:

Inclusão de elemento entre elementos. Nesse caso, o procedimento de integração localiza a posição de inserção pela identificação do elemento anterior. Após identificar a posição, abre espaço para inseri-lo entre os elementos, de modo que o elemento anterior irá apontar para ele e ele apontará para o sucessor do elemento anterior, conforme é apresentado pela Figura 32.

Inclusão de elemento gateway. Elementos gateways indicam fluxos paralelos ou de decisão, sendo que esses elementos apresentam sucessores que dependem deles. Nesse caso, o procedimento incorpora automaticamente seus sucessores no modelo composto se o gateway for incluído, conforme ilustrado pela Figura 33.

Figura 32: Inclusão de elemento entre elementos

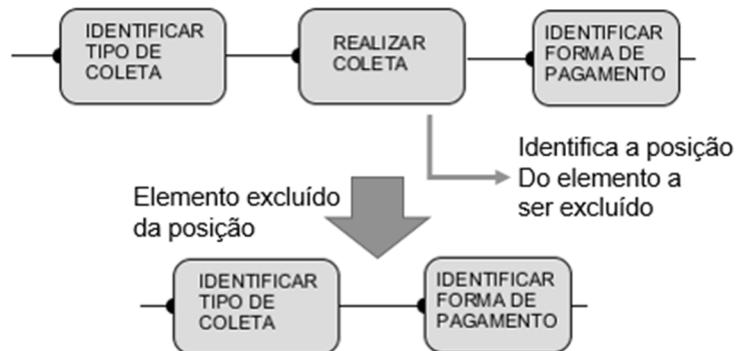
Fonte: Elaborado pelo autor

Figura 33: Inclusão de elemento entre elementos

Fonte: Elaborado pelo autor

Exclusão de elemento. Nesse procedimento considera-se exclusão quando um elemento não é inserido no modelo composto por decisão do usuário. Nesse caso, o método mantém a referência do elemento anterior e faz o apontamento para o próximo elemento que será inserido, conforme ilustrado pela Figura 34.

Figura 34: Exclusão de elemento na integração



Fonte: Elaborado pelo autor

O procedimento de integração mantém ponteiros para os elementos dos modelos M_A e M_B , de modo que todas as comparações possíveis são realizadas e inseridas nas posições corretas no modelo M_C . Além disso, ele atribui um status para cada elemento já que foi incluído ou excluído, de maneira que esses elementos não sejam mais comparados, diminuindo assim o espaço de busca e reduzindo o tempo de processamento.

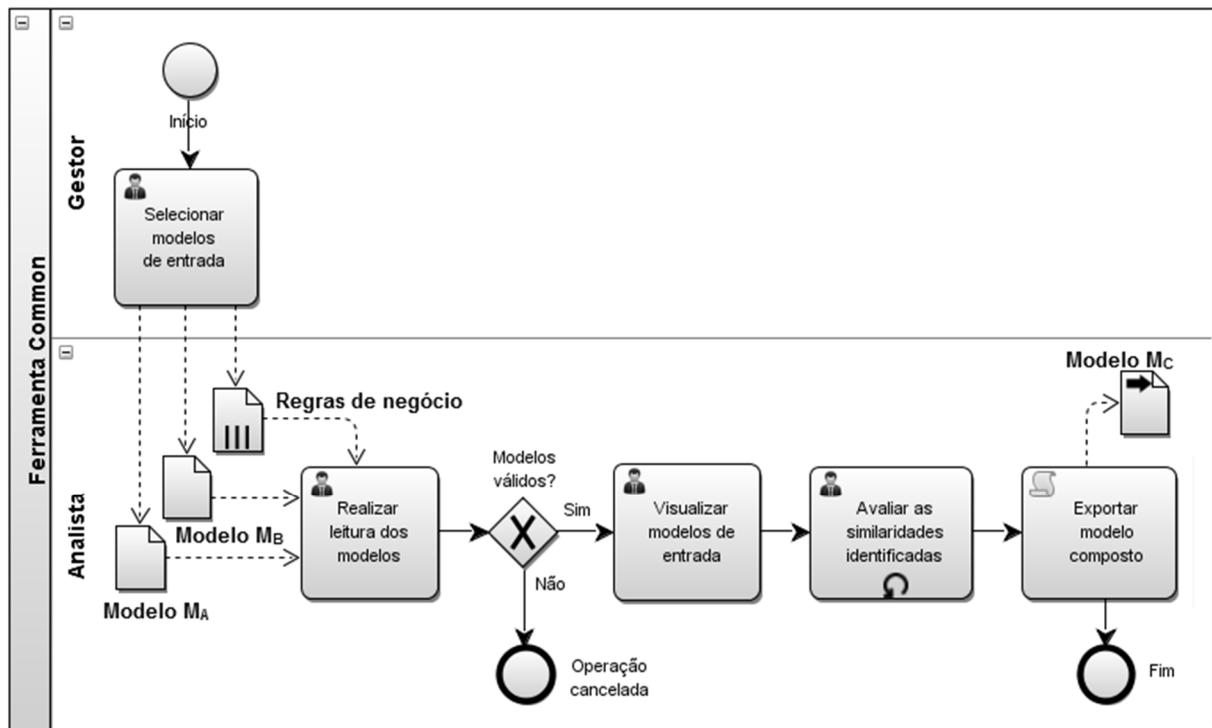
5 FERRAMENTA DE COMPOSIÇÃO DE MODELOS

Com a intenção de pôr em prática a técnica apresentada no Capítulo 4, bem como atender ao segundo objetivo específico enumerado na Seção 1.3, propõe-se uma ferramenta de composição de modelos que visa auxiliar analistas que realizam a integração de modelos de processos de negócio, sensíveis à notação BPMN 2.0. Com a implementação dessa ferramenta, chamada Common – Composição de Modelos de Negócio, pretende-se reduzir o esforço de composição e minimizar a propensão de erros que produzem inconsistências nos modelos integrados. Assim, este Capítulo está organizado da seguinte forma. A Seção 5.1 apresenta o processo de integração de modelos. Em seguida, a Seção 5.2 apresenta a arquitetura da ferramenta Common. Os algoritmos implementados (interpretação de modelo, identificação de similaridade e processo de composição e gravação do modelo composto) são apresentados na Seção 5.3. Por fim, a Seção 5.4 apresenta as funcionalidades e interfaces da ferramenta proposta.

5.1 Processo de integração de modelos

O processo de integração de modelos, proposto na ferramenta Common, comprehende a execução de uma sequência de passos que deve ser seguida, conforme ilustra a Figura 35. Esses passos incluem a seleção dos modelos de entrada e a compreensão das regras de negócio, a leitura, importação, validação e visualização dos modelos M_A e M_B , a interação com a ferramenta para avaliar, aceitar ou rejeitar as similaridades indicadas e a exportação do modelo integrado M_C .

Figura 35: Processo de integração de modelos



Fonte: elaborado pelo autor

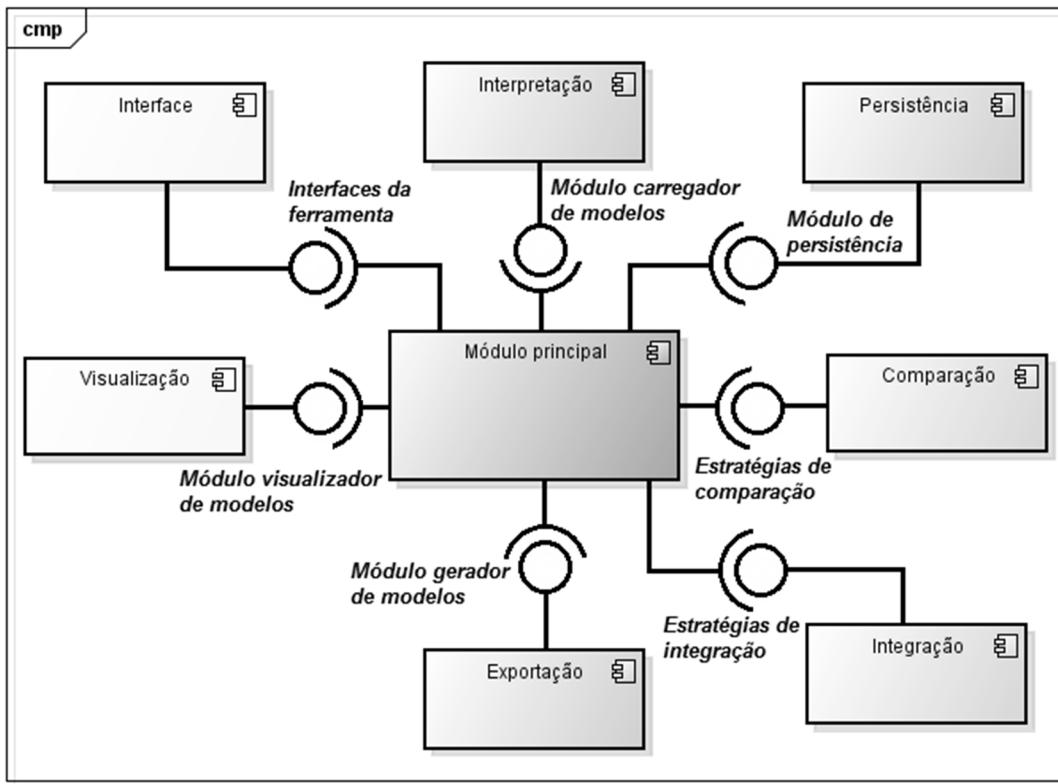
A sequência detalhada dos passos, bem como seus respectivos responsáveis, é descrita a seguir:

- **Selecionar modelos de entrada.** O gestor seleciona os modelos de entrada M_A e M_B que devem ser integrados, juntamente com as regras de negócio que descrevem o comportamento do modelo que será produzido, e envia-os ao analista dar início ao processo de composição, que se dá pela compreensão das regras de negócio e pela leitura e importação desses modelos.
- **Realizar leitura dos modelos.** O analista executa uma rotina automatizada que realiza a leitura e importação dos modelos de entrada para que esses modelos possam ser validados e utilizados pela ferramenta.
- **Validação dos modelos.** Assim que os modelos de entrada são importados, uma rotina automática realiza a validação desses modelos conforme as definições estabelecidas na Seção 4.2. Essa validação visa garantir que os modelos são bem formados e atendem aos requisitos especificados. Se os modelos forem inválidos a operação é cancelada, caso contrário, a rotina segue o fluxo do processo.
- **Visualizar modelos de entrada.** Após a validação dos modelos, a ferramenta apresenta a visualização deles na tela, para que o analista possa analisar e compreender os fluxos dos dois processos de negócio que eles representam.
- **Avaliar similaridades identificadas.** Essa é a atividade principal do processo de composição, uma vez que o analista deve avaliar, aceitar ou rejeitar as similaridades sugeridas pela ferramenta, para realizar a integração dos modelos de entrada.
- **Exportar modelo composto.** Após concluir a integração dos modelos de entrada em um modelo final M_C , que compreende a composição dos modelos M_A e M_B , uma rotina automática realiza a exportação do modelo produzido com base na especificação da notação BPMN 2.0.

5.2 Arquitetura da ferramenta Common

A arquitetura da ferramenta Common possibilita uma visão mais técnica do processo que a modelagem de negócio apresentada na Seção anterior, pois ilustra os módulos da ferramenta e suas respectivas interações entre eles, conforme pode ser visualizado na Figura 36. As principais características apresentadas se referem ao repositório de modelos, camada de apresentação, interpretador de modelos, técnica de composição, parâmetros de configuração, regras de negócio e integração de modelos, que são descritos a seguir:

Figura 36: Diagrama de componentes da ferramenta Common



Fonte: Elaborador pelo autor

- **Interface.** É o componente que faz a interface entre o analista - usuário da ferramenta e o módulo principal. Compreende as telas da ferramenta que permitem a interação do usuário através da seleção de opções de configuração e execução das rotinas de composição. A *Application Programming Interface* (API) Java Swing¹⁰ foi utilizada para a criação das telas exibidas pela ferramenta.
- **Visualização.** Esse componente possibilita a exibição dos modelos que são utilizados durante o processo de composição (M_A , M_B e M_C). Esses modelos são exibidos como imagem em formato *Scalable Vectorial Graphics* (SVG¹¹) mantido pela Word Wide Web Consortium (W3C). O framework Apache Batik¹² foi utilizado para exibir as imagens geradas através de comandos XML.
- **Interpretação.** Esse componente tem como finalidade ler os modelos M_A e M_B , verificar sua aderência à especificação da notação de modelagem utilizada e realizar a identificação dos elementos dos modelos, bem como suas propriedades e comportamentos.

¹⁰ Java Swing - <http://docs.oracle.com/javase/7/docs/technotes/guides/swing/>

¹¹ SVG - <https://www.w3.org/Graphics/SVG/>

¹² Apache Batik - <https://xmlgraphics.apache.org/batik/>

- **Persistência.** É a camada que armazena os modelos carregados pelo componente de interpretação e o modelo produzido pelo processo de composição. O software SQLite¹³ foi utilizado para armazenar esses modelos no disco rígido.
- **Comparação.** É o componente responsável por executar as estratégias de comparação definidas na Seção 4.3, que visam a identificação da similaridade dos elementos dos modelos de entrada, tomando por base os elementos avaliados nos modelos e os parâmetros de configuração.
- **Integração:** O componente de integração é invocado após a avaliação das similaridades encontradas e atua na de integração do modelo M_C , através da interação do usuário com a ferramenta, com base nos procedimentos de integração definidos na Seção 4.4.
- **Exportação.** Esse componente é responsável pela geração do modelo composto, produzido pelo processo de composição de modelos. O modelo produzido segue a especificação da notação BPMN 2.0.
- **Módulo principal:** É o motor da ferramenta, responsável pelo gerenciamento de todos os demais módulos que possuem funções independentes. Foi implementado em Java¹⁴ e seu funcionamento se assemelha ao padrão arquitetural Microkernel, em que um módulo central é responsável por conectar as extensões e coordenar a colaboração dos demais componentes (BUSCHMANN et al., 1996).

5.3 Algoritmos da ferramenta Common

Três algoritmos foram implementados na ferramenta Common para realizar as etapas de leitura e interpretação dos modelos, identificação de similaridades e integração de modelos e gravação do modelo composto, que compreendem as principais etapas do uso da ferramenta. A descrição de cada um, bem como seu detalhamento, segue na sequência desta Seção.

5.3.1 Leitura e interpretação de modelo

Conforme detalhado na Seção 2.5 deste trabalho, os modelos de processos de negócio que seguem a notação BPMN 2.0, atendendo a especificação OMG-BPMN (2013), utilizam o padrão XPDL para descrever os elementos e as ligações deles, formando, assim, o conteúdo de um modelo em sua forma estrutural, em baixo nível. Dessa forma, o conteúdo de um modelo, em sua estrutura de arquivos, é constituído por *tags* XML.

Para realizar a leitura dos modelos de entrada, o módulo interpretador executa uma rotina de conversão dessas *tag* XML para capturar e armazenar, no banco de dados, todas as propriedades e atributos dos elementos dos modelos. Cada tipo de elemento (*event*, *task*, *gateway*, *sequenceFlow*, etc.) possui suas próprias características, que são representadas por atributos e propriedades nas *tags* XML, de modo que o interpretador precisa reconhecer cada uma delas para identificar corretamente as informações contidas na *tag*.

A Figura 37 ilustra um exemplo em que as *tags* task apresentam as propriedades *completionQuantity*, *id*, *isForCompensation*, *name* e *startQuantity* e os atributos *incoming* e

¹³ SQLite - <https://www.sqlite.org/>

¹⁴ Oracle Java - <http://www.oracle.com/br/java/overview/index.html>

outgoing. A representação completa dos atributos e propriedades da notação BPMN 2.0 pode ser obtida em OMG-BPMN(2013).

Figura 37: Tags XML de modelos BPMN

```

<task completionQuantity="1" id="_3" isForCompensation="false"
      name="RECEBER RECIPIENTE" startQuantity="1">
  <incoming>_28</incoming>
  <outgoing>_26</outgoing>
</task>
<task completionQuantity="1" id="_4" isForCompensation="false"
      name="ABRIR RECIPIENTE" startQuantity="1">
  <incoming>_26</incoming>
  <outgoing>_7</outgoing>
</task>
<sequenceFlow id="_26" sourceRef="_3" targetRef="_4"/>
```

Fonte: Elaborador pelo autor

O Algoritmo 1 apresenta a rotina de interpretação do conteúdo do modelo, que realiza um laço de repetição no arquivo (modelo) para capturar todas as *tags* que representam os elementos dele. Assim que uma *tag* (elemento) é capturada, seu tipo é identificado e seus atributos são armazenados no banco de dados.

Algoritmo 1: Leitura e importação de modelo

Entrada: Modelo de processo de negócio

Saída: Elementos do modelo carregados em memória

- 1: **Enquanto** não fim do arquivo **Faça**
 - 2: *simbolo* = *tagXML*;
 - 3: *tipo* = *simbolo*.identificaTipoElemento();
 - 4: **Para** *i* = 1 **até** *tipo*.qtdeAtributos()
 - 5: *objeto* = *simbolo*.analisaXML(*tipo*);
 - 6: *persiste* (*objeto*);
 - 7: **Fim para**
 - 8: **Fim enquanto**
-

5.3.2 Identificação de similaridade e processo de composição

A identificação das similaridades entre os elementos dos modelos de entrada (M_A e M_B) e a respectiva integração desses elementos, para a produção do modelo composto (M_C), é realizada pelo Algoritmo 2, apresentado nesta Seção. Esse algoritmo, é responsável pela implementação das técnicas de comparação e integração apresentadas no Capítulo 4.

Inicialmente, esse algoritmo realiza dois laços de repetição, sendo que o primeiro laço percorre os elementos do modelo M_A e o segundo laço percorre elementos do modelo M_B . Essa varredura permite que cada elemento do modelo M_A seja comparado com todos os elementos do modelo M_B , com o objetivo de encontrar os elementos similares.

Algoritmo 2: Identificação de similaridade e processo de composição

Entrada: Modelos M_A e M_B de processo de negócio

Saída: Modelo de processo de negócio, carregado em memória

```

1: pInf = 90%;
2: pSup = 99%;
3: Enquanto existirem elementos no modelo  $M_A$  Faça
4:   Enquanto existem elementos no modelo  $M_B$  Faça
5:     Se elementoA.indiceSimilaridade(elementoB) == 100% Então
6:       modeloMC.adiciona(elementoA);
7:     Senão Se elementoA.indiceSimilaridade(elementoB) Entre pInf e pSup Então
8:       listaElementosSimilares.adiciona(elementoA);
9:     Fim Se
10:   Fim enquanto
11:   Se listaElementosSimilares.tamanho > 0 Então
12:     elemento = selecionaElementoDaLista();
13:     modeloMC.adiciona(elemento);
14:   Senão
15:     Se confirmaInclusao("elemento não similar!") Então
16:       modeloMC.adiciona(simboloA);
17:     Fim Se
18:   Fim Se
19: Fim enquanto
20: Enquanto existirem elementos no modelo  $M_B$ 
21:   Se confirmaInclusao("elemento não similar!") Então
22:     modeloMC.adiciona(simboloB);
23:   Fim se
24: Fim enquanto
```

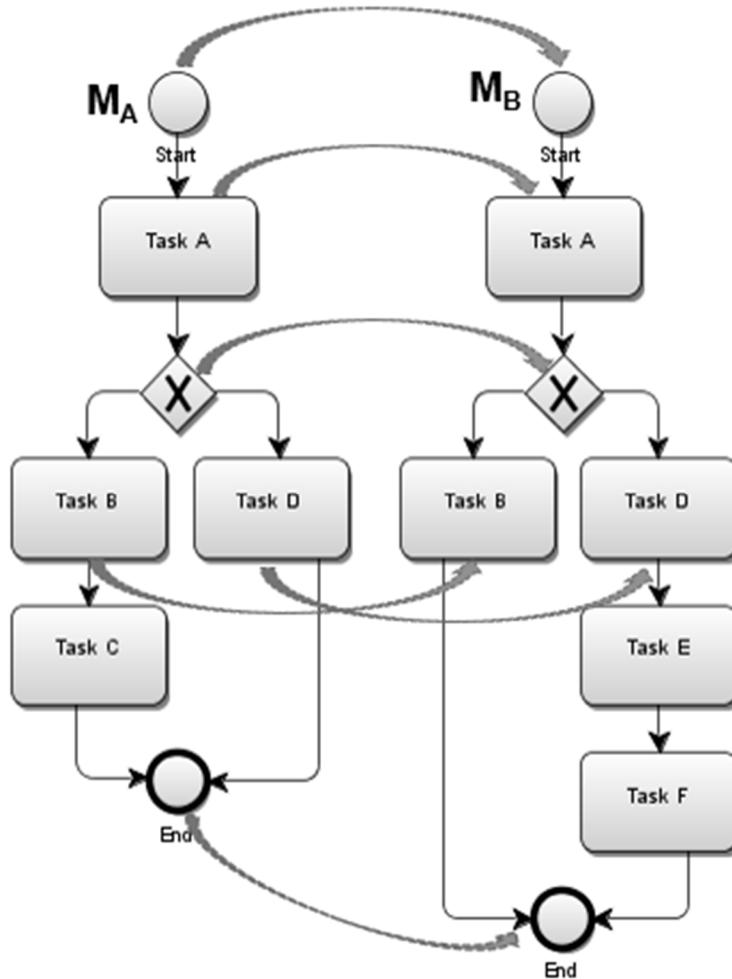
A cada execução completa do primeiro laço de repetição, os elementos comparados são avaliados conforme o grau de similaridade calculado, sendo que os elementos serão incluídos no modelo M_C automaticamente quando sua similaridade alcançar o índice de 100%. Quando nenhum elemento for identificado com 100% de similaridade, o algoritmo apresentará todos os elementos que ficarem na faixa entre 90% e 99% (representada pelas variáveis $pInf$ e $pSup$ respectivamente), para que o usuário decida qual é o elemento ideal. Por último, quando todos os elementos comparados ficarem abaixo do índice de 90% ($pInf$) de similaridade, o usuário será questionado se deve ou não incluir o elemento no modelo M_C , pois ficará, então, caracterizado que esse elemento não é comum entre os dois modelos.

Destaca-se, neste ponto, que a opção do usuário de decidir sobre incluir ou não determinado elemento no modelo M_C , quando seu índice de similaridade não atingir 100%, deve ser baseada nas regras definidas pelo gestor do negócio, pois essas regras estabelecem o comportamento que o modelo composto de apresentar.

Essa execução é repetida para cada elemento do modelo M_A até que todos os elementos desse modelo sejam avaliados, restando, eventualmente, alguns elementos do modelo M_B . Neste ponto, o algoritmo executou até a linha identificada pelo número 19.

Uma vez que todos os elementos do modelo M_A foram avaliados, restando apenas no modelo M_B os elementos que não foram selecionados para fazer parte do modelo M_C , seja de forma automática ou por seleção do usuário, o algoritmo deve então percorrer o modelo M_B para que os elementos restantes sejam avaliados pelo usuário com relação a sua inclusão ou não no modelo M_C . Na Figura 38, esses elementos do modelo M_B são representados pelas atividades *Task E* e *Task F*.

Figura 38: Identificação de similaridades entre dois modelos BPMN



Fonte: Elaborador pelo autor

Para percorrer os elementos restantes do modelo M_B , que não alcançaram os índices de similaridade necessários, o algoritmo executará outro laço de repetição para solicitar a confirmação de inclusão dos mesmos, pois poderá haver a possibilidade de eles pertencerem ao modelo composto. Esse trecho do algoritmo é representado pelas linhas 20 até a 24.

Ao final dessa atividade, o modelo M_C , resultado da combinação dos modelos de entrada M_A e M_B , está concluído e pronto para ser exportado em formato adequado, conforme está explicado na próxima Seção.

5.3.3 Exportação do modelo composto

Depois que o Algoritmo 2, responsável pela integração dos modelos, foi concluído, o próximo passo se concentra na exportação do modelo M_C . Ou seja, os objetos armazenados em memória, que representam os elementos do modelo produzido, são gravados em disco, concluindo, assim, o processo de composição dos modelos de entrada M_A e M_B , conforme descrito no Algoritmo 3.

Para isso, a ferramenta faz um laço de repetição entre os objetos do modelo M_C e escreve as *tags* XML conforme o tipo do elemento. Considerando que cada elemento possui seus atributos particulares, o algoritmo faz outro laço de repetição (interno) para ler esses atributos para formar a respectiva *tag* XML completa. Por fim, o algoritmo se encerra com a gravação em disco do modelo produzido.

Algoritmo 3: Gravação do modelo composto

Entrada: Objetos do modelo composto carregados em memória

Saída: Modelo de processo de negócio

```

1: Enquanto existirem elementos no modelo  $M_C$ 
2:    $tipo = elemento.identificaTipoElemento();$ 
3:    $tagXML = elemento.incluiPropriedades();$ 
4:   Para  $i = 1$  até  $tipo.qtdeAtributos()$ 
5:      $tagXML = tagXML + elemento.formata(atributo);$ 
6:   Fim para
7:    $modelo.grava(tagXML);$ 
8: Fim enquanto
```

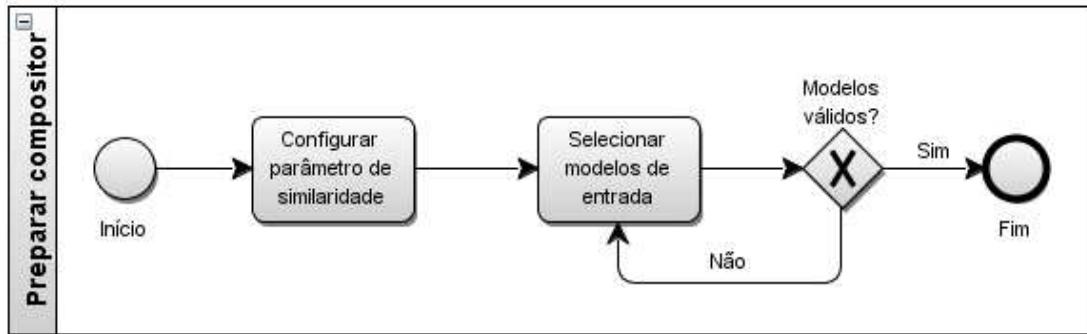
5.4 Apresentação da ferramenta Common

A apresentação da ferramenta Common está dividida em duas etapas, funcionalidades e interface. As funcionalidades são representadas por dois processos principais: preparação do ambiente e execução da composição de modelos. A interface representa o meio pelo qual o usuário interage com a ferramenta, através de telas que apresentam as diversas funções que podem ser realizadas.

5.4.1 Funcionalidades

As funcionalidades da ferramenta Common são duas: (1) preparar o ambiente para realizar a composição e (2) realizar a composição dos modelos de entrada. Para isso, processos foram construídos para representar graficamente essas funcionalidades, conforme ilustram as Figura 39 e Figura 40, respectivamente.

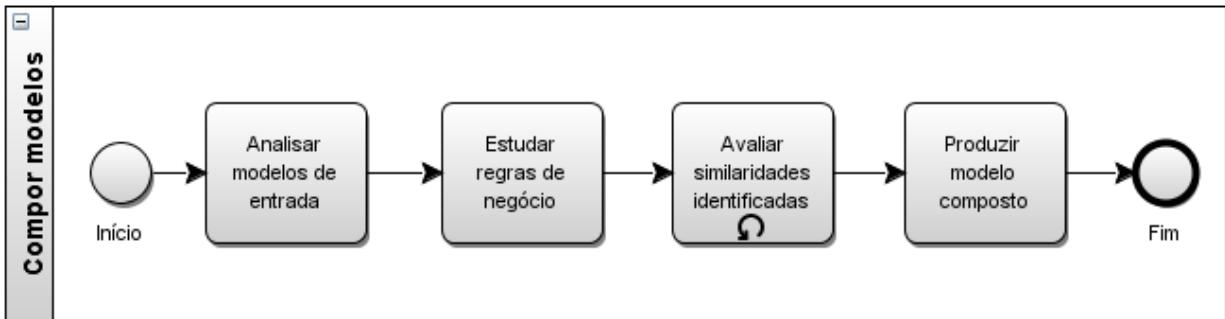
Figura 39: Funcionalidade – preparar a ferramenta



Fonte: Elaborado pelo autor

Para preparar o ambiente, o analista deve configurar o parâmetro de similaridade (VLS) que será utilizado pela ferramenta para identificar o nível de equivalência dos elementos nos modelos que serão compostos. Após, os modelos de entrada são selecionados e validados, permitindo, com isso, que o ambiente esteja pronto para realizar a composição.

Figura 40: Funcionalidade – compor modelos



Fonte: Elaborado pelo autor

Para realizar a composição dos modelos, inicialmente o analista faz uma análise dos modelos carregados em conjunto com as regras de negócio que serão utilizadas para definir o comportamento do modelo composto. Durante o processo de composição, o analista avalia as similaridades identificadas pela ferramenta até que todos os elementos dos modelos de entrada sejam verificados. Por fim, o modelo integrado é produzido.

5.4.2 Interface da ferramenta Common

A interface é o módulo mais importante da ferramenta do ponto de vista do usuário, pois é através dela que ele pode escolher os modelos que serão comparados para produzir o modelo composto, conforme ilustra a Figura 19. A partir da escolha dos modelos de entrada M_A e M_B e com base nas regras de negócio que estabelecem o comportamento do modelo composto M_C , o analista pode se beneficiar dos recursos de identificação de similaridades da ferramenta Common para diminuir seu esforço de composição, bem como evitar erros sintáticos e semânticos durante a produção do modelo final.

Neste trabalho optou-se por apresentar apenas a interface de composição, devido ao fato de que as demais interfaces como, por exemplo, parâmetros de configuração e opções de manipulação de arquivos (Salvar, Salvar como e Sobre) são demasiado simples, uma vez que seguem o padrão dos editores de texto mais conhecidos na indústria. Assim, prioriza-se o espaço

nesta dissertação para a interface principal, que realmente possui importância para o processo de composição.

Os modelos carregados, M_A e M_B , e o modelo em construção, M_C , são exibidos pela ferramenta no formato imagem SVG, de modo que o usuário possa visualizá-los durante o processo de composição. Esse tipo de arquivo é constituído por *tags XML*, que podem ser manipuladas e redefinidas em tempo de execução, podendo, assim, exibir elementos em coloração diferente.

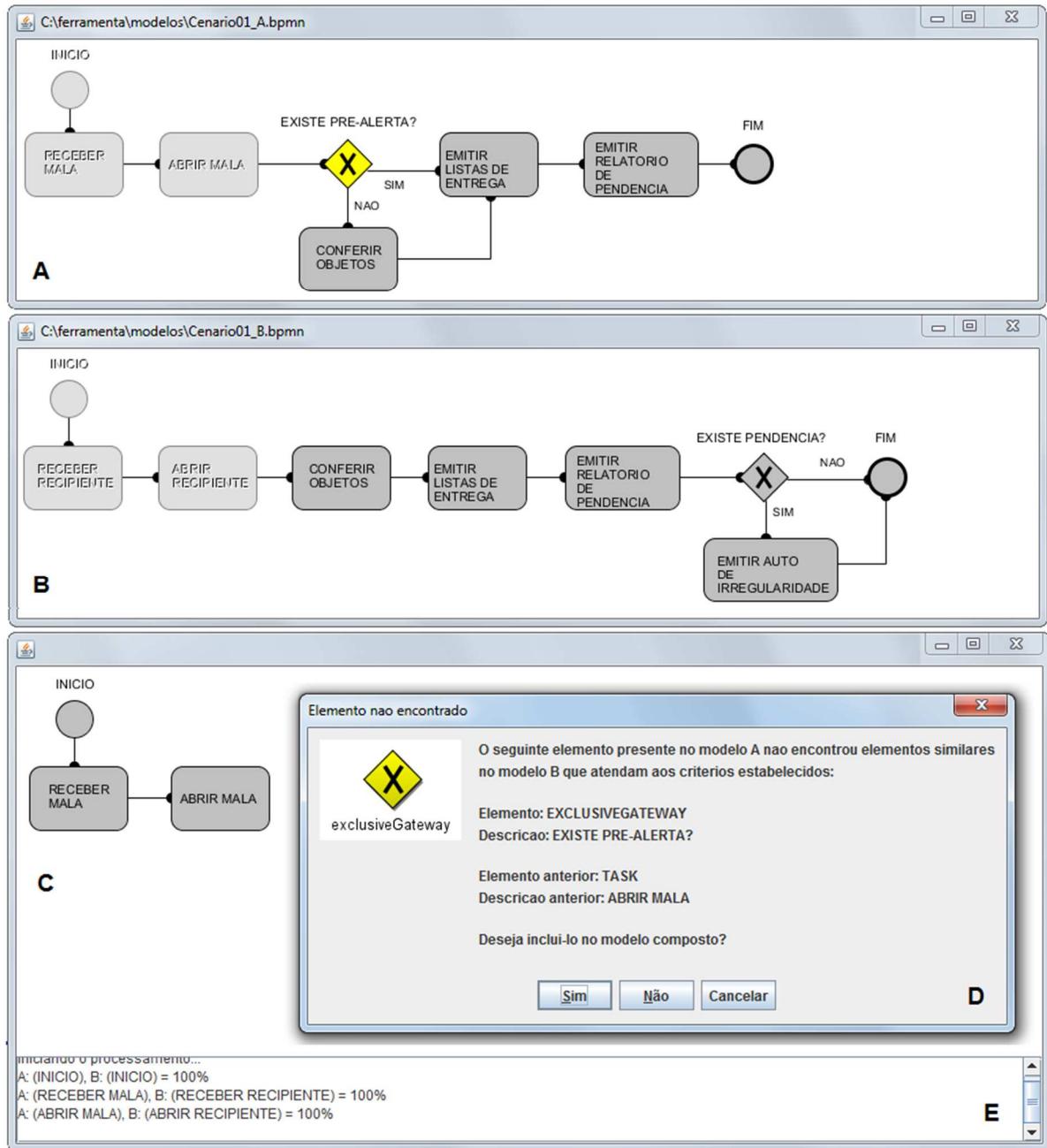
A coloração diferente, conforme ilustrado pela Figura 40, que representa o Cenário 01 do segundo experimento, permite ao usuário identificar facilmente qual é o elemento atual que está em conflito (destacado em amarelo), cuja solução depende da intervenção do mesmo. Além disso, os elementos que já foram avaliados ficam com uma coloração mais clara e com a descrição em baixo relevo, para identificar que já foram avaliados.

Os elementos identificados com 100% de similaridade são incluídos automaticamente no modelo M_C , ao contrário dos elementos que não são encontrados ou com índice de similaridade entre 90 e 99%, conforme parâmetro definido nas configurações, cuja inclusão no modelo composto depende da intervenção do usuário.

No exemplo exibido pela Figura 41, a decisão de inserir o elemento em destaque, nomeado de EXISTE PRÉ-ALERTA, depende do usuário, que se valerá das regras de negócio para decidir sobre a resolução do conflito. Após seleção da resposta, a ferramenta dará seguimento à varredura, percorrendo os demais elementos dos modelos até o final. O detalhamento de cada item da interface é explicado a seguir:

- **(A) Área de visualização do modelo A.** Essa área visualiza os elementos do primeiro modelo de entrada, de modo que o analista possa ter uma compreensão visual desse modelo, facilitando, assim, a interação com a ferramenta.
- **(B) Área de visualização do modelo B.** Essa área visualiza os elementos do segundo modelo de entrada. Assim como em **(A)**, permite ao analista obter uma compreensão visual desse modelo.
- **(C) Área de visualização do modelo composto.** Essa área possibilita a visualização, em tempo de execução, do modelo que está sendo produzido. Durante o processo de composição, o analista consegue visualizar a construção do modelo M_C a partir da inclusão dos elementos durante a execução, permitindo uma visão parcial do modelo final.
- **(D) Caixa de diálogo.** Essa pequena janela apresenta os elementos equivalentes que foram identificados com nível de similaridade inferior a 100%, bem como elementos que não são similares e que podem ser incluídos ou não no modelo composto. Como auxílio ao analista, essa tela apresenta a imagem do elemento em conflito, além da descrição e do tipo dele, bem como a descrição e o tipo do elemento anterior.
- **(E) Área de eventos.** Essa área apresenta as mensagens de andamento da execução da ferramenta, que representam as atividades que já foram realizadas e a indicação do índice de similaridade entre cada elemento incluído no modelo M_C .

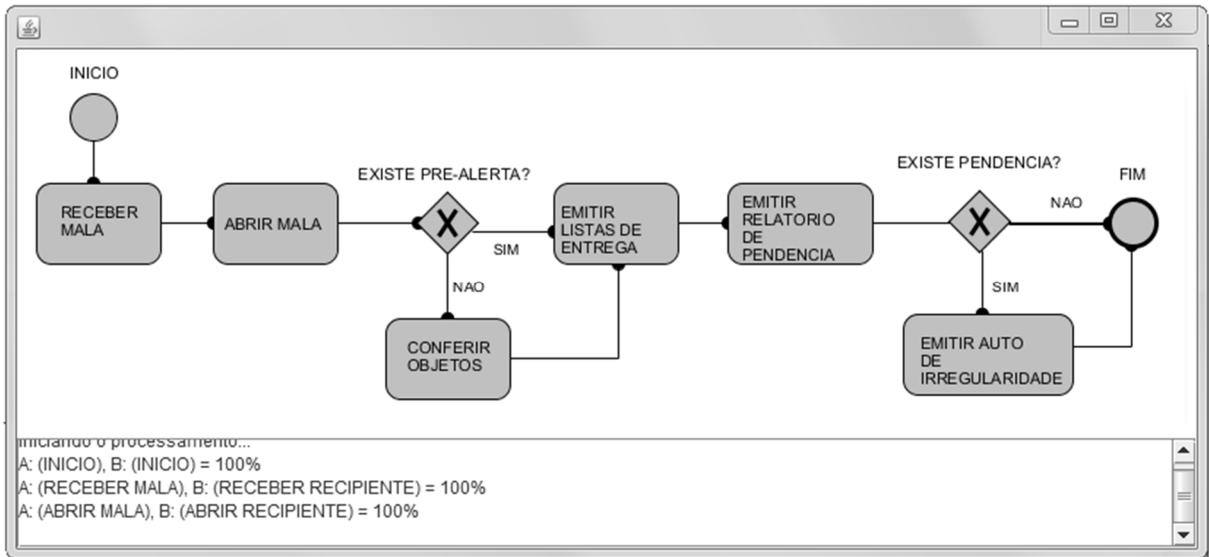
Figura 41: Exibição dos modelos M_A e M_B durante o processo de composição



Fonte: Elaborado pelo autor

Por fim, o modelo composto M_C , que comprehende a integração dos modelos M_A e M_B , produzido pela ferramenta Common no exemplo anterior, é apresentado na Figura 42, cuja representação ilustra a completa integração dos modelos de entrada.

Figura 42: Modelo M_C – resultado da integração de M_A e M_B



Fonte: Elaborado pelo autor

6 AVALIAÇÃO DA TÉCNICA PROPOSTA

Com a finalidade de validar a técnica proposta no Capítulo 4, bem como atender ao terceiro e quarto objetivos específicos enumerados na Seção 1.3, foram realizados dois experimentos de composição de modelos BPMN para a solução de seis cenários de avaliação em cada um. O primeiro experimento fez uso de um editor BPMN, porém, sem apoio de uma ferramenta que pudesse auxiliar na identificação de similaridades e/ou integração dos modelos. Já o segundo experimento foi realizado com o apoio da ferramenta Common, descrita no Capítulo 5 deste trabalho.

Dessa forma, para apresentar a avaliação da técnica proposta, este Capítulo foi organizado da seguinte maneira. A Seção 6.1 apresenta o método de pesquisa. A Seção 6.2 apresenta a formulação das hipóteses. O perfil dos participantes dos experimentos é apresentado na Seção 6.3. A Seção 6.4 apresenta o experimento 1 e a tabulação dos resultados. A tabulação e os resultados do experimento 2 são apresentados na Seção 6.5. A análise comparativa dos dois experimentos é realizada na Seção 6.6. Por fim, a Seção 6.7 apresenta as ameaças à validade do estudo.

6.1 Método de pesquisa

Com o intuito de mensurar esforço, a taxa de erros e a corretude de composição de modelos de processos de negócio, montou-se dois experimentos para avaliar como analistas realizam essa tarefa e quais os principais problemas encontrados nos modelos produzidos. Para isso, foram criados seis cenários de avaliação em cada experimento, sendo que cada cenário compreende dois modelos, identificados pelo conjunto $M = \{(M_{1A}, M_{1B}); (M_{2A}, M_{2B}); \dots (M_{6A}, M_{6B})\}$, onde M_A indica o primeiro modelo de entrada e M_B o segundo, de cada cenário que é identificado pelos números de 1 a 6.

Além disso, foram criados três casos de uso para cada cenário – que representam as regras de negócio, indicados pelo conjunto $C = \{(C_{1A}, C_{1B}, C_{1C}); (C_{2A}, C_{2B}, C_{2C}); \dots (C_{6A}, C_{6B}, C_{6C})\}$, onde C_A indica o caso de uso referente ao modelo M_A , C_B indica o caso de uso referente ao modelo M_B e C_C o caso de uso referente ao modelo M_C que foi produzido pelos participantes, de cada cenário que é identificado pelos números de 1 a 6.

Ao todo, foram produzidos cento e vinte artefatos (entre modelos e casos de uso) para o desenvolvimento de cada experimento, conforme apresentado nas Tabela 6 e Tabela 7, que listam, também, a quantidade de elementos dos modelos de entrada dos experimentos.

Os modelos M_A e M_B representam fluxos diferentes de um mesmo processo, seja pela diferença de granularidade – em que uma ou mais atividades são agrupadas ou expandidas em mais atividades, pela diferença na descrição dos elementos (*events, tasks, gateways, etc.*) ou pelo mesmo modelo em diferentes versões.

Além dos modelos constantes no conjunto M e dos casos de uso constantes no conjunto C de cada cenário, os participantes receberam um questionário de levantamento de informações profissionais, em que puderam responder questões como tempo de experiência com modelagem de software, desenvolvimento de software, idade e formação acadêmica/profissional. Esses dados do questionário serviram para análise de perfil e de experiência dos participantes, conforme será descrito na Seção 6.2.

Tabela 6: Cenários do experimento 1

Cenário	Modelo A		Modelo B		Modelo C	
	Nome	Nº de elementos	Nome	Nº de elementos	Nome	Nº de elementos
1 – Recebimento de objetos postais	M1 _A	10	M1 _B	8	M1 _C	11
2 – Triagem de objetos postais	M2 _A	10	M2 _B	13	M2 _C	16
3 – Postagem de objetos	M3 _A	13	M3 _B	13	M3 _C	16
4 – Coleta de objetos postais	M4 _A	13	M4 _B	9	M4 _C	15
5 – Entrega de correspondência	M5 _A	5	M5 _B	12	M5 _C	13
6 – Processamento de arquivos de postagem	M6 _A	10	M6 _B	10	M6 _C	14

Fonte: Elaborado pelo autor

Tabela 7: Cenários do experimento 2

Cenário	Modelo A		Modelo B		Modelo C	
	Nome	Nº de elementos	Nome	Nº de elementos	Nome	Nº de elementos
1 – Recebimento de objetos postais	M1 _A	8	M1 _B	9	M1 _C	10
2 – Triagem de objetos postais	M2 _A	12	M2 _B	11	M2 _C	16
3 – Postagem de objetos	M3 _A	13	M3 _B	12	M3 _C	17
4 – Coleta de objetos postais	M4 _A	14	M4 _B	11	M4 _C	17
5 – Entrega de correspondência	M5 _A	6	M5 _B	12	M5 _C	13
6 – Processamento de arquivos de postagem	M6 _A	9	M6 _B	9	M6 _C	13

Fonte: Elaborado pelo autor

Apesar dos dois experimentos representarem os mesmos processos de negócio em cada cenário como, por exemplo, recebimento de objetos postais no Cenário 1, eles os apresentam de maneira similar, porém diferente, pois são distintos em relação ao número de elementos em cada modelo, conforme pode ser observado nas Tabelas 7 e 8 e/ou pela descrição ou posicionamento dos elementos nos mesmos.

Essa diferença foi idealizada com o intuito de evitar que os participantes pudessem manipular os resultados do segundo experimento após terem realizado o primeiro, devido ao possível aprendizado sobre os fluxos que eles representam, uma vez que os mesmos participantes foram selecionados para realizar os dois experimentos. Além disso, para reforçar ainda mais a validade dos resultados obtidos, os experimentos foram realizados em datas distintas, com intervalo de dez meses entre eles, conforme ilustrado na Figura 40.

Para realizar cada experimento, os participantes receberam dois modelos de entrada, M_A e M_B, que foram criados usando a notação de BPMN 2.0, e três especificações de caso de uso, C_A, C_B, e C_C, que representam os modelos de entrada M_A e M_B, e o modelo a ser produzido M_C, respectivamente. Destaca-se que o modelo produzido pelos participantes (M_C) foi corrigido com base no modelo desejado, especificado pelo caso de uso C_C.

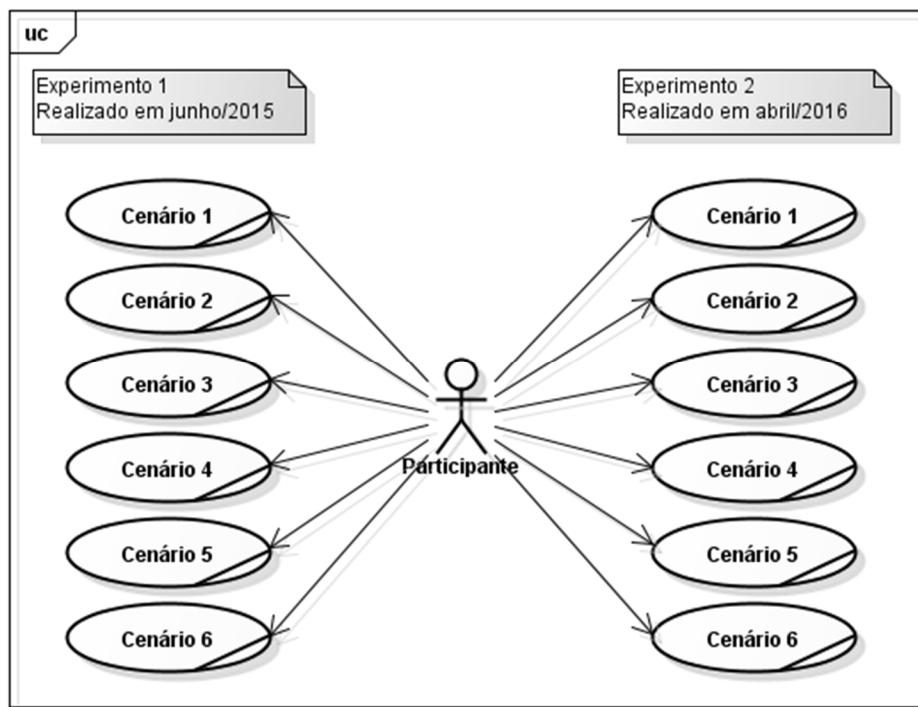
Além dos arquivos lógicos, os modelos e as especificações de casos de uso foram entregues impressos em papel, possibilitando que o participante pudesse escolher o modo pelo qual se sentiria mais confortável para realizar o experimento. Ademais, antes do início do primeiro experimento foi realizado um rápido treinamento para explicar a sintaxe da notação, proporcionando que todos os participantes compartilhassem de um conhecimento comum sobre modelagem BPMN 2.0. Dessa forma, a avaliação deste trabalho foi realizada em duas etapas, através da realização dos experimentos 1 e 2, conforme detalhamento a seguir:

Experimento 1. Essa primeira etapa consistiu em realizar a composição dos modelos de cada cenário sem o uso de uma ferramenta de apoio, isto é, apesar de os participantes do experimento usarem um editor de modelagem BPMN, ele não forneceu auxílio ao processo de composição, nem identificou as similaridades entre os modelos dos cenários.

Experimento 2. Essa segunda etapa consistiu em realizar o experimento 2 com o apoio da ferramenta desenvolvida neste trabalho, cujo objetivo principal foi identificar se essa ferramenta, que implementa a técnica proposta, auxilia os analistas no processo de composição de modelos de processos de negócios, reduzindo o esforço e o número de erros, e aumentando a produtividade.

A visualização da execução dos dois experimentos está representada na Figura 43, cujo diagrama de caso de uso ilustra que os mesmos participantes executaram cada um dos cenários, de ambos experimentos, em períodos distintos.

Figura 43: Diagrama de caso de uso dos experimentos 1 e 2



Fonte: Elaborado pelo autor

Por fim, sabendo-se que o objetivo principal desta avaliação é verificar se a técnica e a ferramenta propostas, de fato, reduzem o esforço e a propensão a erros no processo de composição de modelos BPMN 2.0, justifica-se utilizar os mesmos participantes para comparar, com maior precisão, os resultados dos dois experimentos. Dessa forma, confrontar os desempenhos de integrar modelos de forma manual e integrar modelos com o uso de uma

ferramenta de apoio, utilizando-se os mesmos participantes, pode trazer resultados mais fidedignos, pois essa atividade envolve características particulares de cada um como, por exemplo, capacidade de leitura e compreensão de modelos e casos de uso, e habilidades de construção de fluxos de processos (modelos).

6.2 Formulação das hipóteses

Com base nos experimentos realizados neste trabalho, que visam mensurar o esforço, a taxa de erros e a corretude dos modelos produzidos, tanto de forma manual como semiautomática, três hipóteses foram formuladas para avaliar se, de fato, a técnica proposta proporciona melhoria ao processo de composição de modelos de processos de negócio sensíveis à notação BPMN 2.0.

Considerando que a melhoria do processo de composição se dá pela redução do esforço, pela minimização da taxa de erros e pelo aumento da corretude nos modelos produzidos, as hipóteses que foram formuladas visam avaliar cada um desses aspectos, de forma que seja possível elencar indícios suficientes para responder às questões de pesquisa elaboradas na Seção 1.2 deste trabalho.

Hipótese 1. Realizar a composição de modelos de processos de negócio é uma atividade que pode ser considerada difícil e onerosa, uma vez que envolve procedimentos de análise dos modelos de entrada M_A e M_B para identificar as possíveis similaridades entre os elementos de cada modelo. Dessa forma, esta hipótese avalia se a composição de modelos realizada de forma semiautomática, com base na técnica e na ferramenta propostas, reduz o esforço de composição, ou seja, ajuda os analistas a integrar os modelos consumindo menos tempo.

Hipótese Nula 1, H_{1-0} : A técnica e a ferramenta propostas não reduzem o esforço de composição ao produzir o modelo M_C a partir dos modelos M_A e M_B .

H_{1-0} : Esforço (M_A, M_B) Semiautomática \geq Esforço (M_A, M_B) Manual

Hipótese Alternativa 1, H_{1-1} : A técnica e a ferramenta propostas reduzem o esforço de composição ao produzir o modelo M_C a partir dos modelos M_A e M_B .

H_{1-1} : Esforço (M_A, M_B) Semiautomática $<$ Esforço (M_A, M_B) Manual

O cálculo do esforço é realizado pela soma dos tempos de execução investidos pelos analistas para integrar os modelos, sendo Esforço (M_A, M_B)_{Semiautomática} referente à soma dos tempos de execução de forma semiautomática e Esforço (M_A, M_B)_{Manual} referente à soma dos tempos de execução de forma manual, conforme a seguinte fórmula:

$$Esforço = soma(t) \quad (4)$$

onde,

t é o tempo gasto pelo analista para realizar a integração dos modelos.

Hipótese 2. Modelos compostos a partir da integração de dois modelos de entrada podem ser produzidos com erros, tanto sintáticos como semânticos. Esses erros se referem a divergências entre o modelo produzido e o modelo desejado. Desse modo, integrar modelos com menos erros se torna um desafio, uma vez que essa atividade exige atenção e concentração

por parte dos analistas. Com base nisso, esta hipótese avalia se a composição de modelos realizada de forma semiautomática, com base na técnica e na ferramenta propostas, reduz a taxa de erros de composição, ou seja, ajuda os analistas a produzirem modelos mais próximos do desejado.

Hipótese Nula 2, H₂₋₀: A técnica e a ferramenta propostas não reduzem a taxa de erros ao produzir o modelo M_c a partir dos modelos M_A e M_B.

H₂₋₀: Taxa erros (M_A, M_B) Semiautomática \geq Taxa erros (M_A, M_B) Manual

Hipótese Alternativa 2, H₂₋₁: A técnica e a ferramenta propostas reduzem a taxa de erros ao produzir o modelo M_c a partir dos modelos M_A e M_B.

H₂₋₁: Taxa erros (M_A, M_B) Semiautomática $<$ Taxa erros (M_A, M_B) Manual

O cálculo da taxa de erros é realizado pela soma dos erros (sintáticos e semânticos) encontrados nos modelos produzidos pelos analistas, sendo Taxa erros (M_A, M_B)_{Semiautomática} referente à soma dos erros nos modelos produzidos de forma semiautomática e Taxa erros (M_A, M_B)_{Manual} referente à soma dos erros nos modelos produzidos de forma manual, conforme a seguinte fórmula:

$$\text{Taxa erros} = \text{soma}(e) \quad (5)$$

onde,

e é quantidade de erros produzidos pelo analista no modelo composto.

Hipótese 3. Como mencionado anteriormente, analistas podem produzir os modelos compostos com inconsistências. Sabendo-se que o desejado é a integração dos modelos M_A e M_B sem a ocorrência de erros, esta hipótese avalia se a composição realizada de forma semiautomática, com base na técnica e na ferramenta propostas, auxilia os analistas a produzirem modelos livres de erro, ou seja, modelos que não possuem nenhuma inconsistência em comparação com o modelo desejado.

Hipótese Nula 3, H₃₋₀: A técnica e a ferramenta propostas não auxiliam na extinção de erros ao produzir o modelo M_c a partir dos modelos M_A e M_B.

H₃₋₀: Corretude (M_A, M_B) Semiautomática \leq Corretude (M_A, M_B) Manual

Hipótese Alternativa 3, H₃₋₁: A técnica e a ferramenta propostas auxiliam na extinção de erros ao produzir o modelo M_c a partir dos modelos M_A e M_B.

H₃₋₁: Corretude (M_A, M_B) Semiautomática $>$ Corretude (M_A, M_B) Manual

O cálculo da corretude é realizado pela identificação de inconsistência nos modelos produzidos pelos analistas, sendo Corretude (M_A, M_B)_{Semiautomática} referente à contagem de modelos sem erros produzidos de forma semiautomática e Corretude (M_A, M_B)_{Manual} referente à contagem de modelos sem erros produzidos de forma manual, conforme a seguinte fórmula:

$$\text{Taxa erros} = \text{Contagem}(c) \quad (6)$$

onde,

c é quantidade de modelos produzidos sem inconsistências pelo analista.

A Tabela 8 apresenta um resumo das hipóteses nulas e alternativas desta pesquisa, que visam elencar indícios que sejam suficientes para responder à questão de pesquisa deste trabalho. A partir dos testes dessas hipóteses, pretende-se confirmar se a técnica e a ferramenta propostas melhoraram, de fato, o processo de composição de modelos.

Tabela 8: Resumo das hipóteses

Hipótese	Tipo	Representação
H_1	Nula	H_{1-0} : Esforço (M_A, M_B) Semiautomática \geq Esforço (M_A, M_B) Manual
	Alternativa	H_{1-1} : Esforço (M_A, M_B) Semiautomática $<$ Esforço (M_A, M_B) Manual
H_2	Nula	H_{2-0} : Taxa erros (M_A, M_B) Semiautomática \geq Taxa erros (M_A, M_B) Manual
	Alternativa	H_{2-1} : Taxa erros (M_A, M_B) Semiautomática $<$ Taxa erros (M_A, M_B) Manual
H_3	Nula	H_{3-0} : Corretude (M_A, M_B) Semiautomática \leq Corretude (M_A, M_B) Manual
	Alternativa	H_{3-1} : Corretude (M_A, M_B) Semiautomática $>$ Corretude (M_A, M_B) Manual

Fonte: Elaborado pelo autor

6.3 Perfil dos participantes

Para realizar os experimentos, foram convidados 23 participantes com formação superior completa em Ciência da Computação, Sistemas de Informação ou Licenciatura em Informática, dos quais, 15 deles possuem pós-graduação em nível de especialização, conforme está detalhado na Tabela 9, que representa o grau de escolaridade dos participantes. Além disso, 9 participantes estão cursando pós-graduação em nível de mestrado, sendo que 8 em Computação Aplicada e 1 em Educação.

Tabela 9: Grau de escolaridade dos participantes

Escolaridade	Participantes
Especialista	15
Graduado	8

Fonte: Elaborado pelo autor

Desses 23 voluntários, 2 atuam como estudante/pesquisador em nível de mestrado, 11 atuam como desenvolvedor de sistemas, 4 atuam como analista de negócios, 1 atua como arquiteto de sistemas, 2 atuam como administrador de sistemas/dados, 1 como analista de qualidade, 1 como analista de testes, 1 como arquiteto de sistemas e, por último, 1 atua como consultor de informática, conforme descrito na Tabela 10.

A Tabela 11 ilustra faixa etária dos participantes, que compreende 2 entre 20 e 25, 3 entre 26 e 30 anos, 7 entre 31 e 35 anos, 4 entre 36 e 40, 4 entre 41 e 45 anos e 3 com mais de 46 anos de idade.

Tabela 10: Ocupação dos participantes

Ocupação	Participantes
Desenvolvedor de sistemas	11
Analista de negócios	4
Administrador de dados	2
Estudante/pesquisador	2
Analista de testes	1
Arquiteto de sistemas	1
Analista de qualidade	1
Consultor de TI	1

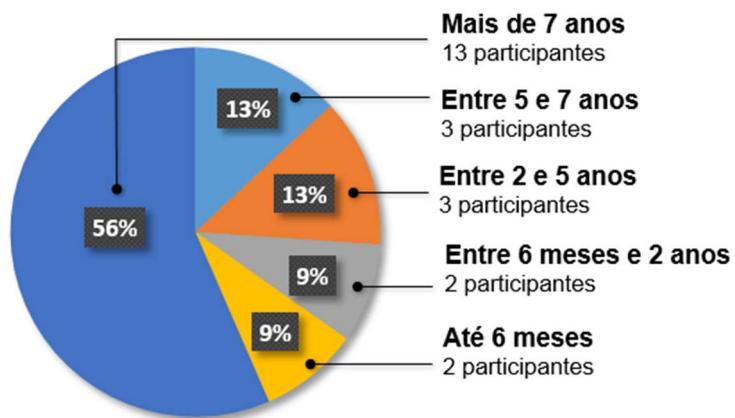
Fonte: Elaborado pelo autor

Tabela 11: Faixa etária dos participantes

Faixa etária	Participantes
Entre 20 e 25 anos	2
Entre 26 e 30 anos	3
Entre 31 e 35 anos	7
Entre 36 e 40 anos	4
Entre 41 e 45 anos	4
Mais de 45 anos	3

Fonte: Elaborado pelo autor

Sobre a experiência em desenvolvimento de software, 13 possuem mais de 7 anos, 3 possuem entre 5 e 7 anos, 3 possuem mais de 2 anos e até 5 anos, 2 possuem mais de 6 meses e até 2 anos e 2 participantes possuem menos de 6 meses, conforme ilustrado na Figura 44.

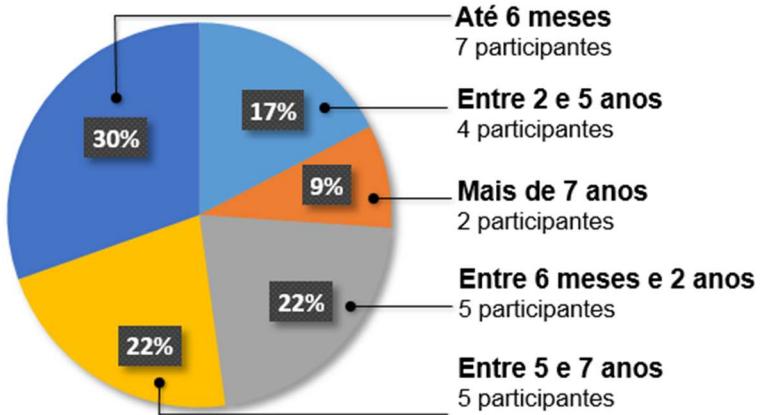
Figura 44: Experiência em desenvolvimento de software

Fonte: Elaborado pelo autor

Em relação à experiência de modelagem de software, 2 possuem mais de 7 anos, 5 possuem mais de 5 anos e até 7 anos, 4 possuem mais de 2 anos e até 5 anos, 5 possuem mais

de 6 meses e até 2 anos e 7 participantes possuem até 6 meses. A Figura 45 apresenta o gráfico dessa amostra.

Figura 45: Experiência em modelagem de software



Fonte: Elaborado pelo autor

Ao analisar o perfil dos participantes dos experimentos, é possível destacar que foram selecionadas pessoas que possuem relação direta com a área de informática (desenvolvimento ou modelagem de software), tanto na formação quanto na profissão. Além disso, identifica-se que os participantes se concentram, na maioria, entre 31 e 45 anos de idade e com mais de 5 anos de experiência em desenvolvimento de software, sendo que quase a metade deles possui mais de 2 anos de experiência em modelagem, caracterizando, assim, que a amostra selecionada possui perfil adequado para a realização dos experimentos.

6.4 Experimento 1 – Composição de modelos sem uso de ferramenta de apoio

Para realizar a integração dos modelos de forma manual, sem uso de ferramenta de apoio, os participantes receberam os artefatos citados na Seção 6.1 – definidos na Tabela 7, um editor de modelos que atende a especificação da notação BPMN 2.0, um rápido treinamento sobre modelagem de processos para nivelar o conhecimento entre eles e um manual BPMN 2.0 para consulta.

O editor BPMN selecionado para este experimento foi o Yaoqiang BPMN Editor¹⁵, desenvolvido pela empresa Blenta, que segue a licença de uso gratuita regida pela *General Public Licence 3.0* (GPL). Esse editor foi escolhido por apresentar suporte completo à versão 2.0 da notação BPMN, possuir licença de uso gratuita, dispensar instalação – uma vez que depende apenas da *Java Virtual Machine* (JVM)¹⁶ instalada no computador e por possuir facilidade de uso para criar, editar e visualizar modelos.

O experimento consistiu em abrir o editor disponibilizado e, com base nos modelos de entrada e nos casos de uso de cada cenário, produzir os modelos compostos (M_{1C}, M_{2C}, ... M_{6C}) que representam a integração dos modelos M_A e M_B de cada cenário, conforme definido nos respectivos casos de uso apresentados na Tabela 7. Destaca-se, neste ponto, que esse editor não

¹⁵ <https://sourceforge.net/projects/bpmn/>

¹⁶ https://www.java.com/pt_BR/

possui qualquer funcionalidade de identificação de similaridade entre os modelos, tampouco capacidade de integração deles, o que faz com que essa atividade seja considerada manual, apesar do uso de um editor de modelos.

Diferentemente de criar modelos a partir de um arquivo vazio, baseando-se apenas nos casos de uso, o experimento permitiu o uso dos arquivos lógicos dos modelos de entrada, possibilitando, assim, copiar elementos dos modelos M_A e M_B para colá-los nos modelos M_C de cada cenário, tornando essa atividade o mais realista possível, como é feito na prática, em que normalmente um analista utiliza os modelos existentes para produzir o modelo composto, aproveitando vários de seus elementos.

Para medir a duração deste experimento e, consequentemente, mensurar o esforço aplicado pelos participantes para integrar os modelos de forma manual, cada participante registrou, em uma ficha de coleta de dados fornecida, o horário de início e término que compreendeu o tempo gasto para realizar a execução de cada cenário. Ressalta-se que o experimento foi realizado sob supervisão de execução dos cenários e de seus respectivos lançamentos na ficha de informações. O detalhamento dos cenários pode ser conferido na Tabela 12, que ilustra a quantidade de elementos, classificada por tipo, de cada cenário deste experimento.

Tabela 12: Caracterização dos modelos utilizados no experimento 1.

Cenário	Elemento	M_{1A}	M_{1B}	M_{1C}
1	Nº Tarefas	6	4	6
	Nº Eventos	2	2	2
	Nº Gateways	2	2	3
2	Nº Tarefas	7	9	11
	Nº Eventos	2	2	2
	Nº Gateways	1	2	3
3	Nº Tarefas	8	8	10
	Nº Eventos	2	2	4
	Nº Gateways	2	2	2
4	Nº Tarefas	9	6	10
	Nº Eventos	2	2	2
	Nº Gateways	2	1	3
5	Nº Tarefas	3	8	8
	Nº Eventos	2	3	3
	Nº Gateways	0	1	2
6	Nº Tarefas	3	5	6
	Nº Eventos	6	4	7
	Nº Gateways	0	1	1

Fonte: Elaborado pelo autor

Cada participante produziu seis modelos, sendo um por cenário. Ao todo, foram produzidos 138 modelos pelos 23 participantes neste experimento, cujos modelos foram analisados e corrigidos visando mensurar o esforço aplicado e quantificar os erros sintáticos e semânticos encontrados. A análise dos resultados do experimento 1 é apresentada a seguir, na Seção 6.3.1.

6.4.1 Tabulação dos resultados do experimento 1

Conforme esclarecido na Seção 6.1, esse experimento acumulou um montante de 138 modelos que foram produzidos através da composição manual pelos participantes, isto é, sem o apoio de uma ferramenta que pudesse identificar as similaridades e/ou integrar os modelos.

Esses modelos foram analisados e corrigidos, visando contabilizar erros sintáticos e semânticos, bem como mensurar o esforço aplicado para produzir os modelos compostos. A Tabela 13 apresenta os resultados individualizados por participante para cada cenário avaliado, onde pode-se visualizar o tempo que cada participante gastou para construir os modelos dos cenários, bem como a quantidade de erros encontrados.

Tabela 13: Resultados individuais por participante – experimento 1

Participante	Cenário 1			Cenário 2			Cenário 3			Cenário 4			Cenário 5			Cenário 6		
	Duração (min.)		Erros															
	Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos	
1	11	1	2	12	1	1	9	1	0	8	0	0	7	0	0	8	0	1
2	23	0	2	14	0	2	3	0	2	4	0	0	5	2	1	5	0	0
3	16	0	0	17	0	0	7	0	0	5	0	0	9	0	0	2	0	0
4	14	0	1	7	0	1	10	0	0	4	0	0	4	0	0	13	0	1
5	14	0	3	8	0	5	6	0	0	3	0	0	10	0	1	7	0	1
6	15	0	2	18	0	4	7	0	1	5	2	4	2	2	1	2	0	1
7	9	0	1	7	0	0	5	2	0	5	0	0	9	0	0	8	0	0
8	18	0	2	14	1	3	33	0	4	6	0	0	14	1	4	5	1	1
9	20	1	0	10	0	0	9	2	0	5	0	0	12	0	1	8	0	0
10	22	0	1	16	0	0	16	0	0	8	0	0	9	0	0	12	0	0
11	16	2	4	20	1	4	9	0	2	12	0	3	10	0	2	6	0	2
12	57	2	3	33	1	1	17	0	1	18	0	1	19	0	0	10	0	0
13	40	0	1	32	1	2	14	0	1	35	1	4	10	0	1	53	1	2
14	24	1	1	21	1	2	16	0	0	17	0	0	9	0	0	19	0	1
15	10	0	2	21	0	1	26	0	1	8	0	2	6	0	2	23	0	3
16	10	0	0	9	0	1	13	0	0	4	0	0	6	0	0	8	0	0
17	12	0	3	9	0	3	8	0	4	10	0	3	7	1	5	16	0	1
18	21	0	1	41	1	2	31	0	4	15	0	3	14	0	3	25	1	2
19	17	0	0	11	0	1	8	0	3	8	0	3	5	0	3	11	0	0
20	11	0	0	10	0	1	7	1	2	3	0	0	7	1	0	14	1	2
21	25	0	0	23	0	0	17	0	1	13	3	0	13	0	0	18	1	0
22	9	1	2	16	0	1	9	1	0	6	0	0	9	0	0	59	1	1
23	19	0	1	10	1	2	8	0	1	7	0	3	9	0	0	10	0	0

Fonte: Elaborado pelo autor

A Tabela 14 apresenta os resultados agrupados por cenário de avaliação, com base nos itens avaliados: média, desvio padrão e mediada, separados por esforço e taxa de erros. O esforço é baseado no tempo de execução em minutos e a taxa de erros na quantidade de erros

sintáticos e semânticos. O item número de avaliações se refere à quantidade de participantes que realizaram o experimento.

Tabela 14: Resultados agrupados por cenário teste

Itens avaliados		Cenário 1	Cenário 2	Cenário 3	Cenário 4	Cenário 5	Cenário 6	Geral
Número de avaliações (participantes)		23	23	23	23	23	23	23
Esforço	Média de duração (min.)	18,8	16,5	12,5	9,1	8,9	14,9	13,4
	Desvio padrão	10,9	9,0	8,0	7,1	3,8	14,3	9,9
	Mediana	16	14	9	7	9	10	10,0
Taxa de erros	Erros sintáticos	8	8	7	6	7	6	42
	Erros semânticos	32	37	27	26	24	19	165
	Total de erros	40	45	34	32	31	25	207
	Média de erros	1,7	2,0	1,5	1,4	1,3	1,1	1,5
	Desvio padrão	1,5	1,6	1,3	1,8	1,7	1,1	1,5
	Mediana	1	2	1	0	1	1	1,0

Fonte: Elaborado pelo autor

Ao realizar a análise descritiva da Tabela 14, identifica-se que os dados não estão normalmente distribuídos, tanto pelos valores de esforço como de erros, dado o alto índice de variância e desvio padrão encontrado na amostra.

6.5 Experimento 2 – Composição de modelos com uso da ferramenta Common

Para realizar a integração dos modelos de forma semiautomática, com o uso da ferramenta Common, os participantes receberam os artefatos citados na Seção 6.1 – definidos na Tabela 8, o mesmo editor de modelos BPMN utilizado no experimento 1 e um rápido treinamento sobre a utilização da ferramenta Common.

O editor BPMN foi disponibilizado para que o participante pudesse editar o modelo produzido antes de encerrar o experimento, devendo contabilizar, também, o tempo extra de edição, se fosse o caso. Essa possibilidade estava disponível caso participante percebesse que, durante o processo de integração dos modelos, algum dos modelos fosse produzido com falhas percebidas por ele. Contudo, nenhum participante utilizou esse recurso.

O experimento consistiu em abrir a ferramenta Common, selecionar o cenário desejado, carregando os modelos de entrada M_A e M_B , e realizar o processo de integração, isto é, produzir o modelo M_C conforme as sugestões que a ferramenta lhe apresentava. Durante a realização do experimento, o tempo de realização de cada cenário foi contabilizado automaticamente pela ferramenta.

Conforme citado na Seção 6.1, apesar de um intervalo de dez meses entre a execução dos dois experimentos, os cenários do segundo experimento foram modificados para garantir a validade da avaliação, conforme pode ser observado na Tabela 15, que representa o detalhamento da Tabela 7.

Tabela 15: Caracterização dos modelos utilizados no experimento 2.

Cenário	Elemento	M1 _A	M1 _B	M1 _C
1	Nº Tarefas	5	6	6
	Nº Eventos	2	2	2
	Nº Gateways	1	1	2
2	Nº Tarefas	8	8	11
	Nº Eventos	2	2	2
	Nº Gateways	2	1	3
3	Nº Tarefas	9	8	11
	Nº Eventos	2	2	2
	Nº Gateways	2	2	4
4	Nº Tarefas	9	8	11
	Nº Eventos	2	2	2
	Nº Gateways	3	1	4
5	Nº Tarefas	4	7	8
	Nº Eventos	2	3	3
	Nº Gateways	0	2	2
6	Nº Tarefas	5	4	6
	Nº Eventos	3	4	5
	Nº Gateways	1	1	2

Fonte: Elaborado pelo autor

Assim como no primeiro experimento, cada participante produziu seis modelos, sendo um por cenário. Ao todo, foram produzidos mais 138 modelos neste segundo experimento, pelos mesmos 23 participantes, cujos modelos foram examinados e corrigidos visando mensurar o esforço aplicado e quantificar os erros sintáticos e semânticos encontrados. A análise dos resultados do experimento 2 é apresenta a seguir, na Seção 6.4.1.

6.5.1 Tabulação dos resultados do experimento 2

Assim como no primeiro experimento, os 138 modelos produzidos neste segundo experimento foram analisados e corrigidos, visando contabilizar os erros, bem como mensurar o esforço aplicado para produzi-los. Os resultados individuais por participante encontram-se na Tabela 17, onde é possível identificar, para cada cenário, o tempo gasto e os erros produzidos de cada um.

Ainda em relação a Tabela 16, percebe-se que não houve a ocorrência de erros sintáticos nos modelos produzidos pelos participantes do experimento 2. Isso se deu porque a ferramenta Common, utilizada como apoio durante o processo de composição, impede a ocorrência desse tipo de erro, uma vez que ela proporciona a integração de elementos de modo sintaticamente correto. Em contrapartida, erros semânticos podem ser produzidos porque dependem da interação do usuário, que decide sobre os conflitos com base nas regras de negócio.

Tabela 16: Resultados individuais por participante – experimento 2

Participante	Cenário 1			Cenário 2			Cenário 3			Cenário 4			Cenário 5			Cenário 6		
	Duração (min.)		Erros															
	Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos		Sintáticos	Semânticos	
1	1,4	0	0	4	0	0	4,8	0	0	6,4	0	1	3,3	0	0	3,2	0	0
2	0,7	0	0	2,7	0	0	1,9	0	0	2,3	0	1	1	0	0	1,1	0	0
3	1	0	0	1,6	0	0	1,7	0	0	1,7	0	0	0,8	0	0	0,8	0	0
4	1,2	0	0	2,6	0	1	1,7	0	0	2,4	0	0	1,4	0	1	0,9	0	0
5	3,4	0	0	2,2	0	0	5,9	0	0	1,9	0	2	0,8	0	0	0,5	0	0
6	1	0	0	2,6	0	0	1,6	0	0	1,8	0	1	1,7	0	0	1,4	0	0
7	1,8	0	0	1,7	0	0	1,5	0	0	2,5	0	0	1,6	0	1	1	0	0
8	3,1	0	0	4,6	0	0	2,8	0	1	3,9	0	1	2,4	0	1	2,6	0	0
9	1,9	0	0	5,2	0	2	3,9	0	0	2,9	0	1	2,3	0	1	2,8	0	0
10	1,8	0	0	2	0	0	2,5	0	0	2,3	0	1	1,5	0	0	1,2	0	0
11	1,5	0	0	2,9	0	0	2,9	0	0	2,7	0	1	1,7	0	0	1,4	0	0
12	2,5	0	0	5,3	0	0	4,4	0	0	4,6	0	0	4,3	0	0	3,1	0	0
13	2	0	0	2,7	0	0	1,9	0	0	1,7	0	1	0,9	0	0	0,9	0	0
14	1,4	0	0	2,7	0	0	2,2	0	0	2,2	0	1	1,3	0	0	1	0	0
15	1,7	0	0	2,9	0	0	2,6	0	0	4,2	0	1	2,3	0	0	1,5	0	0
16	1,5	0	0	4,1	0	0	2,4	0	0	2,3	0	1	1,3	0	0	1,2	0	0
17	1,4	0	0	2,6	0	1	2,5	0	1	2,1	0	1	1,5	0	0	1,6	0	0
18	1,1	0	0	2,2	0	0	2,4	0	1	2,7	0	1	2,3	0	1	1,2	0	0
19	1,3	0	0	2,7	0	0	2	0	0	2,3	0	0	1,5	0	0	1,2	0	0
20	1,1	0	0	2,3	0	0	1,6	0	0	3,2	0	1	1,8	0	0	2	0	0
21	1,1	0	1	2,6	0	0	2,9	0	0	2,7	0	0	1,8	0	1	1,2	0	0
22	1,3	0	0	3,5	0	0	4,8	0	1	4,8	0	0	3,5	0	0	2,3	0	0
23	1	0	0	2,7	0	0	1,9	0	0	3,1	0	0	1,8	0	0	1,3	0	0

Fonte: Elaborado pelo autor

A Tabela 17 apresenta os resultados agrupados por cenário de avaliação, com base nos itens avaliados: média, desvio padrão e mediada, separados por esforço e taxa de erros. O esforço é baseado no tempo de execução em minutos e a taxa de erros na quantidade de erros sintáticos e semânticos. O item número de avaliações se refere à quantidade de participantes que realizaram o experimento.

Ao realizar a análise descritiva desse resultado, identifica-se, também, que os dados não estão normalmente distribuídos, tanto pelos indicadores de esforço como de taxa de erros, dado o alto índice do desvio padrão e da variância em relação à média da amostra.

Tabela 17: Resultados agrupados por cenário teste

Itens avaliados		Cenário 1	Cenário 2	Cenário 3	Cenário 4	Cenário 5	Cenário 6	Geral
Número de avaliações (participantes)	23	23	23	23	23	23	23	23
Esforço	Média de duração (min.)	1,6	3,0	2,7	2,9	1,9	1,5	2,3
	Desvio padrão	0,7	1,0	1,2	1,2	0,9	0,8	1,1
	Mediana	1	3	2	3	2	1	2,0
Corretude	Erros sintáticos	0	0	0	0	0	0	0
	Erros semânticos	1	4	4	16	6	0	31
	Total de erros	1	4	4	16	6	0	31
	Média de erros	0	0,2	0,2	0,7	0,3	0	0,2
	Desvio padrão	0,2	0,5	0,4	0,6	0,4	0	0,4
	Mediana	0	0	0	1	0	0	0

Fonte: Elaborado pelo autor

6.6 Análise comparativa dos experimentos

Haja vista que os dois experimentos foram realizados e avaliados, pretende-se, nesta seção, confrontar os resultados desses experimentos para identificar se a composição realizada de forma semiautomática, com base na técnica e na ferramenta propostas, de fato, proporciona melhoria ao processo de composição de modelos de processos de negócio, sensíveis à notação BPMN 2.0. Esta análise visa elencar subsídios suficientes para responder às questões de pesquisa elaboradas na Seção 1.2, a partir da avaliação das três hipóteses formuladas neste Capítulo.

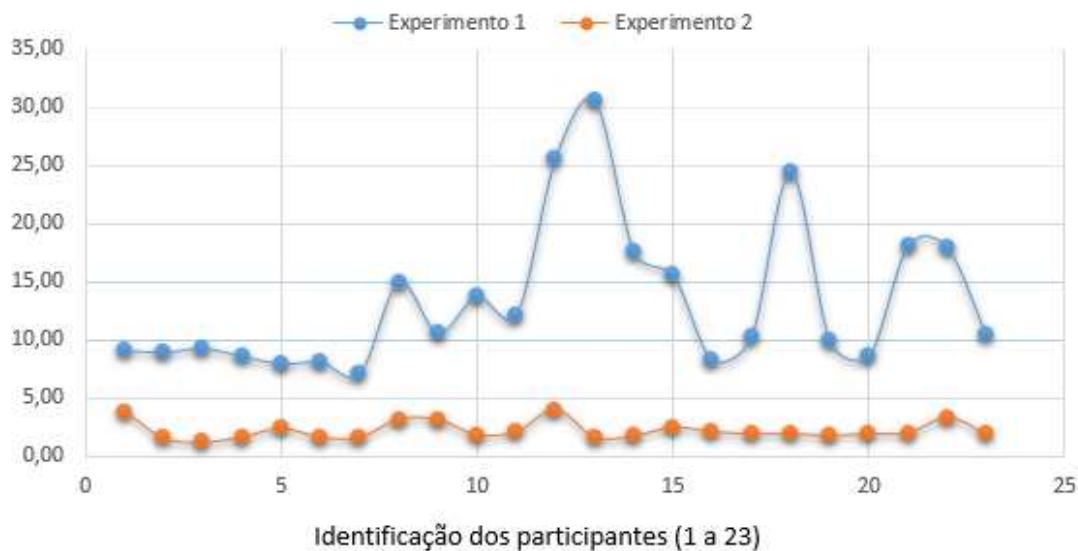
6.6.1 QP-1: avaliação do esforço de composição

Com o objetivo de mensurar o esforço que analistas empenham para integrar modelos de processos de negócio, foram capturados os tempos que os participantes consumiram para realizar os dois experimentos. Esse tempo corresponde ao esforço produzido para compor os seis cenários que os participantes realizaram em cada experimento.

A Figura 46 ilustra o comparativo, em minutos, de cada um dos 23 participantes dos experimentos. Nessa figura é possível visualizar que a média dos tempos gastos pelos participantes no segundo experimento é bastante inferior à média dos tempos gastos pelos mesmos participantes no primeiro experimento.

Na avaliação por participante envolvendo os dois experimentos, visualiza-se que todos os 23 participantes aparentemente consumiram menos tempo para realizar o segundo experimento, destacado na cor laranja no gráfico da Figura 46, sendo que cada ponto na abscissa (x) indica o esforço de cada um dos participantes (1 a 23). Desse modo, conclui-se que os participantes empenderam menos esforço para realizar o segundo experimento em relação ao primeiro.

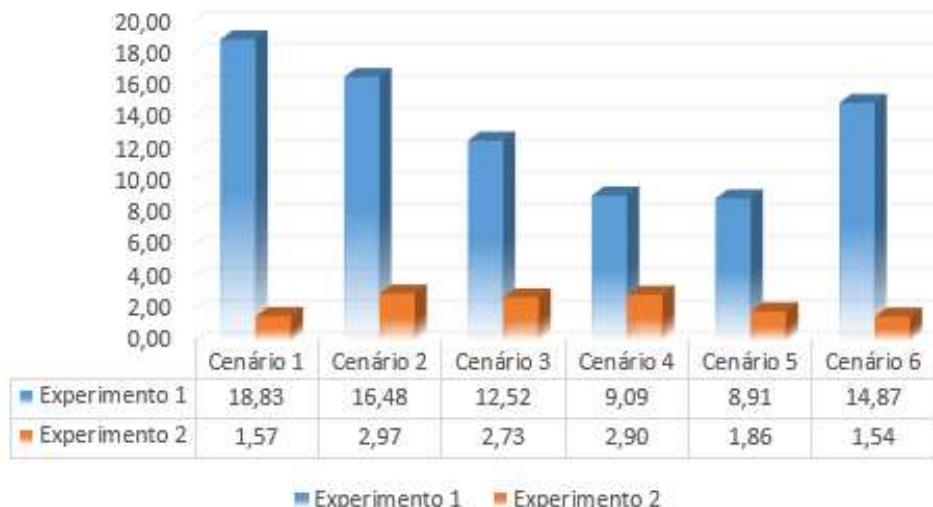
Figura 46: Média de tempo de execução por participante



Fonte: Elaborado pelo autor

Outra análise sobre os mesmos dados foi realizada em relação aos cenários, de modo que os resultados individuais de cada participante foram agrupados visando estabelecer um comparativo entre os dois experimentos. A Figura 47 apresenta a relação entre os experimentos sob a perspectiva dos cenários.

Figura 47: Média de tempo de execução por cenário



Fonte: Elaborado pelo autor

Nessa análise, também é possível visualizar que o esforço empenhado pelos participantes, agrupados por cenário, é aparentemente menor no experimento 2 em relação ao experimento 1. Porém, muito embora os gráficos apresentem números menores no segundo experimento, um teste estatístico foi realizado para verificar se, de fato, a técnica e a ferramenta propostas reduzem o esforço de composição de modelos.

Levando-se em consideração as análises descritivas apresentadas nas Seções 6.3.1 e 6.4.1, que indicam que os levantamentos de tempo realizados nos dois experimentos não estão

normalmente distribuídos, optou-se pela aplicação do teste de duas amostras pareadas de Wilcoxon, voltado para a análise de dados não paramétricos. Segundo Ryan (2011), o teste Wilcoxon é adequado para a investigação de duas amostras, no sentido de identificar se uma população tem valores maiores que a outra, a partir de uma análise estatística de dados cuja distribuição não é conhecida.

Esse teste foi realizado com base nas 138 amostras de tempo de execução de cada experimento, que caracterizam o esforço empregado pelos participantes para realizar a integração dos modelos, comparando os tempos de execução de uma amostra com os da outra. Assim, na avaliação da hipótese 1, conforme descrito na Tabela 18, o teste Wilcoxon apresentou $p\text{-value} < 0,001$, com intervalo de confiança de 95%.

Tabela 18: Resultado do esforço - teste Wilcoxon

Teste Wilcoxon	
Pares (n)	138
Z	-10,2
$p\text{-value}$	< 0,001

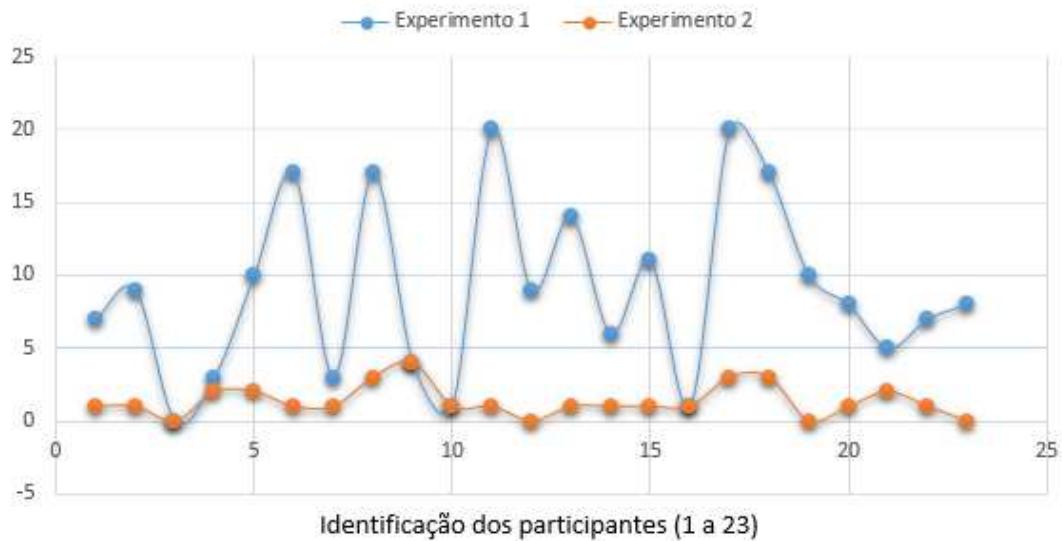
Fonte: Elaborado pelo autor

Sendo $p\text{-value} < 0,05$, há evidências de que a diferença média entre as duas amostras de tempo de execução dos experimentos é significativamente diferente de zero, gerando, assim, indícios suficientes para rejeitar a hipótese nula 1 (H_{1-0}). Portanto, ao apresentar evidências que demonstram que a redução do tempo de execução do segundo experimento em relação ao primeiro é significativa, caracterizar-se que a composição de modelos realizada de maneira semiautomática, com base na técnica e na ferramenta propostas, reduz o esforço que analistas empregam para integrar modelos de processos de negócio sensíveis à notação BPMN 2.0, respondendo assim, a primeira questão de pesquisa deste trabalho.

6.6.2 QP-2: avaliação da taxa de erros

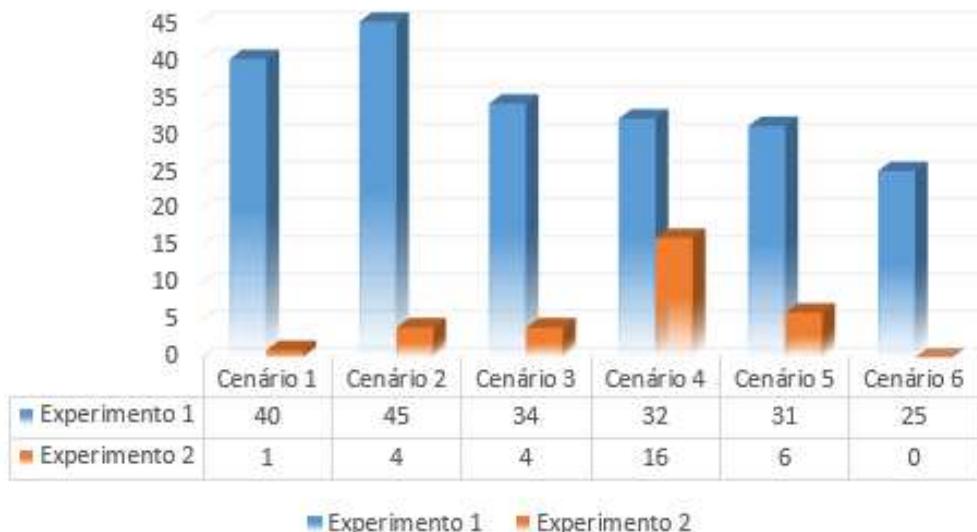
Com o objetivo de mensurar a taxa de erros produzidos por analistas na composição de modelos de processos de negócio, foram quantificados os erros que os participantes apresentaram nos dois experimentos realizados. A Figura 47 ilustra o comparativo do quantitativo de erros de cada um dos 23 participantes dos experimentos. Nessa figura é possível visualizar que o total de erros da maioria dos participantes no segundo experimento é bastante inferior ao total de erros apresentados pelos mesmos participantes no primeiro experimento, exceto em quatro casos em que esse número permaneceu na mesma proporção.

Na avaliação por participante envolvendo os dois experimentos, visualiza-se que a maioria dos participantes aparentemente apresentaram menos erros para realizar o segundo experimento, destacado na cor laranja no gráfico da Figura 48, sendo que cada ponto na abscissa (x) indica a quantidade de erros de cada um dos participantes (1 a 23). Desse modo, conclui-se que os participantes apresentaram menos erros para realizar o segundo experimento em relação ao primeiro.

Figura 48: Taxa de erros por participante

Fonte: Elaborado pelo autor

Além disso, outra análise de erros foi realizada em relação aos mesmos cenários, de modo que os resultados individuais de cada participante foram agrupados visando estabelecer um comparativo entre os dois experimentos. A Figura 49 apresenta a relação entre os experimentos sob a perspectiva dos cenários.

Figura 49: Taxa de erros por cenário

Fonte: Elaborado pelo autor

Nessa análise, também é possível visualizar que os erros apresentados pelos participantes, agrupados por cenário, é aparentemente menor no experimento 2 em relação ao experimento 1. Porém, muito embora os gráficos apresentem números menores no segundo experimento, um teste estatístico foi realizado para verificar se, de fato, a técnica e a ferramenta propostas minimizam a taxa de erros de composição de modelos.

Levando-se em consideração as análises descritivas apresentadas nas Seções 6.3.1 e 6.4.1, que indicam que os levantamentos de erros realizados nos dois experimentos não estão normalmente distribuídos, optou-se também pela aplicação do teste de duas amostras pareadas de Wilcoxon.

Esse teste foi realizado com base nas 138 amostras de erros de cada experimento, que totalizam os erros apresentados pelos participantes para realizar a integração dos modelos, comparando a taxa de erros de uma amostra com a outra. Assim, na avaliação da hipótese 2, conforme descrito na Tabela 19, o teste Wilcoxon apresentou $p\text{-value} < 0,001$, com intervalo de confiança de 95%.

Tabela 19: Resultado da taxa de erros - teste Wilcoxon

Teste Wilcoxon	
Pares (n)	138
Z	-7,2
$p\text{-value}$	< 0,001

Fonte: Elaborado pelo autor

Sendo $p\text{-value} < 0,05$, há evidências de que a diferença média entre as duas amostras de taxas de erros dos experimentos é significativamente diferente de zero. Assim, há indícios suficientes para rejeitar a hipótese nula 2 (H_{2-0}). Portanto, ao apresentar evidências que demonstram que a redução da taxa de erros do segundo experimento em relação ao primeiro é significativa, caracterizar-se que a composição de modelos realizada de maneira semiautomática, com base na técnica e na ferramenta propostas, reduz a taxa de erros dos modelos produzidos a partir da composição de modelos de processos de negócio sensíveis à notação BPMN 2.0, respondendo assim a segunda questão de pesquisa deste trabalho.

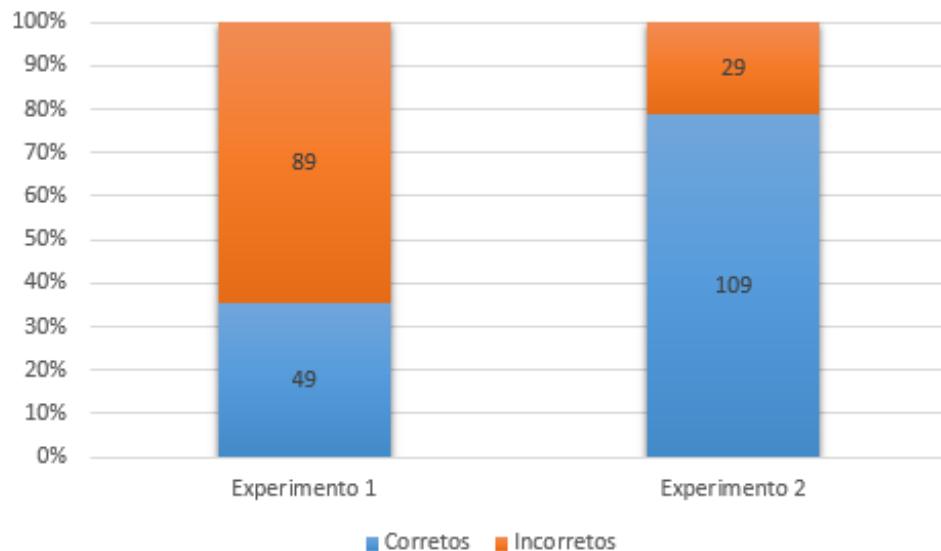
6.6.3 QP-3: avaliação da corretude

Investigar se os modelos produzidos foram corretamente integrados é um fator decisivo para a qualidade do processo de composição, uma vez que as inconsistências existentes prejudicam a compreensibilidade dos fluxos dos modelos.

Nesse sentido, para comparar a corretude dos modelos produzidos nos experimentos 1 e 2 foram contabilizados, para cada modelo, dois indicadores que representam se o modelo está ou não correto, sendo que esses indicadores foram utilizados para realizar o teste McNemar.

Conforme Sheskin (2009), o teste NcNemar é ideal para avaliar se há diferença entre duas amostras pareadas após a aplicação de determinada ação, como, por exemplo, o uso da ferramenta Common na realização do experimento 2. Dessa forma, essa avaliação leva em consideração os 138 pares de modelos dos dois experimentos para determinar se houve aumento da corretude nos modelos produzidos, isto é, se mais modelos foram produzidos sem inconsistências no experimento 2 em relação ao experimento 1.

A Figura 49 ilustra o comparativo da corretude dos dois experimentos, onde é possível identificar no primeiro experimento que 50 modelos foram produzidos corretamente e 89 modelos foram gerados incorretamente, ou seja, apresentam algum erro sintático ou semântico. Já no experimento 2, 109 modelos foram produzidos corretamente contra 29 modelos gerados incorretamente.

Figura 50: Corretude por experimento

Fonte: Elaborado pelo autor

A Tabela 20 apresenta os indicadores utilizados para a aplicação do teste McNemar, conforme a quantidade de modelos corretos e incorretos dos 138 pares de modelos avaliados. Essa tabela apresenta a frequência de duas linhas por duas colunas, que representa o primeiro e o segundo conjunto de resultados dos mesmos participantes. Com isso, cada par foi relacionado para preencher a respectiva tabela, conforme segue: 35 correto x correto, 74 correto x incorreto, 14 incorreto x correto e 15 incorreto x incorreto, totalizando os mesmos valores apresentados na Figura 49.

Tabela 20: Resultado da corretude - teste McNemar

	Correto	Incorreto	Total
Correto	35	14	49
Incorreto	74	15	89
Total	109	29	138

Fonte: Elaborado pelo autor

Com intervalo de confiança de 95%, o resultado do teste McNemar apresentou *p-value* menor que 0,001, indicando que as proporções não são as mesmas, rejeitando assim a hipótese nula 3 (H_{3-0}). Portanto, ao apresentar evidências que demonstram que o aumento da corretude do segundo experimento em relação ao primeiro é significativo, caracterizar-se que a composição de modelos realizada de maneira semiautomática, com base na técnica e na ferramenta propostas, aumenta a corretude dos modelos produzidos a partir da composição de modelos de processos de negócio sensíveis à notação BPMN 2.0, respondendo assim a segunda questão de pesquisa deste trabalho.

6.6.4 Avaliação final

Os experimentos apresentados revelaram uma questão importante no contexto da modelagem de processos de negócio, que diz respeito à mensuração do esforço que analistas

empenham e à quantificação de erros que são encontrados quando se realiza composição de modelos, tanto de forma manual como de forma semiautomática - com o apoio de uma ferramenta de composição.

Com base nos testes estatísticos realizados nos dois experimentos, foi possível avaliar as hipóteses testadas nas Seções 6.6.1, 6.6.2 e 6.6.3, que se referem ao esforço, à taxa de erros e à corretude, relacionados às questões de pesquisa elaboradas na Seção 1.2. Os resultados mostraram sinais suficientes para rejeitar as três hipóteses nulas formuladas, no contexto apresentado. Assim sendo, ao levantar indícios suficientes de que as hipóteses alternativas podem ser verdadeiras, o resultado dos experimentos contribui para o atendimento do primeiro e segundo objetivos específicos, citados na Seção 1.3, desta pesquisa.

Além disso, o resultado dos experimentos gerou uma massa de informações rica até então desconhecida, uma vez que não foram encontrados estudos que apresentam dados comparativos de esforço e de taxas de erro na realização de integração de modelos executada de forma manual e semiautomática. Dessa forma, o conhecimento empírico obtido pela realização dos experimentos contribui para o atendimento do terceiro e quarto objetivos específicos, citados na Seção 1.3 deste trabalho.

Ao atingir todos os objetivos específicos formulados nesta pesquisa, que dizem respeito à proposta de uma técnica de composição, ao desenvolvimento de uma ferramenta que implemente a técnica proposta, à realização de experimentos de composição para validar essa técnica e à produção de conhecimento empírico sobre composição de modelos de processos de negócio, comprova-se que o objetivo principal do trabalho foi alcançado.

Portanto, com comprovação de que os objetivos definidos neste trabalho foram alcançados, afirma-se que a técnica e a ferramenta, propostas nesta pesquisa, de fato, reduzem o esforço que analistas empenham para realizar a integração de modelos, minimizam a taxa de erros sintáticos e semânticos e colaboram para o aumento da corretude nos modelos produzidos, respondendo de forma satisfatória às questões de pesquisa elaboradas neste trabalho.

6.7 Ameaças à validade do estudo

Quando se trabalha com experimentos surge a necessidade de verificar o quanto válido são os métodos e os participantes escolhidos, a forma como foram aplicados e avaliados, e os resultados encontrados. Segundo Wohlin et al. (2012), essa verificação se dá através da avaliação de quatro tipos de validade: validade de construção, validade interna, validade externa e validade de conclusão, que visam reduzir as ameaças ao estudo.

A validade de construção leva em consideração a relação entre a teoria e a observação, isto é, se os resultados obtidos através dos questionários ou experimentos possuem relação com as expectativas da teoria estudada (WOHLIN et al., 2012). Tendo isso em evidência, os experimentos realizados nesta pesquisa foram planejados para mensurar o esforço de composição, bem como quantificar erros nos modelos produzidos pelos participantes, sendo essa estratégia já conhecida na literatura como, por exemplo, no trabalho de Farias et al. (2014). Além disso, os procedimentos de execução e correção dos experimentos foram cuidadosamente planejados, adotando boas práticas de quantificação definidas em Wohlin et al. (2012).

A validade interna considera a causalidade entre o relacionamento observado e o resultado obtido, como forma de invalidar a pesquisa por gerar uma conclusão causal entre o tratamento e os resultados (SHADISH et al., 2001; WOHLIN et al., 2012). Nesta pesquisa, 23 participantes produziram 276 modelos para formar uma boa massa de dados para análise, sendo que ambos experimentos foram realizados pelos mesmos participantes com um intervalo de 10

meses entre o primeiro e o segundo experimento, caracterizando assim a existência de um grupo de controle. Além disso, os modelos sofreram pequenas alterações no segundo experimento quanto ao posicionamento, quantidade e descrição dos elementos, com o objetivo de evitar o aprendizado dos participantes, conforme esclarecido na Seção 6.1.

A validade externa leva em consideração as condições que permitem generalizar os resultados do experimento para a prática industrial ou o mais realista possível (MITCHEL e JOLLEY; WOHLIN et al., 2012). Nesse contexto, foram selecionados participantes que possuem formação adequada para a prática da atividade relacionada aos experimentos, isto é, todos possuem formação superior em informática e alguns, inclusive, com pós-graduação na área, além de boa experiência na área de desenvolvimento ou modelagem de software, conforme detalhado na Seção 6.3. Além disso, as ferramentas e equipamentos utilizados nos experimentos são equivalentes com a prática corporativa, uma vez que os experimentos foram realizados no laboratório da universidade e no local de trabalho dos participantes, proporcionando assim maior familiaridade com a execução dos experimentos.

A validade de conclusão considera a capacidade de obter a conclusão correta com base no resultado do experimento, através da escolha certa dos métodos estatísticos levando em consideração itens como o tamanho da amostra e a confiabilidade das medidas (WOHLIN et al., 2012). Nesse critério, os experimentos produziram uma amostra de 276 modelos, cuja análise estatística descritiva identificou que os dados não estavam normalmente distribuídos, decidindo-se então aplicar os testes não paramétricos Wilcoxon, para as medidas de esforço e taxa de erros, e McNemar para a medida de corretude, com intervalos de confiança de 95%, conforme detalhado na Seção 6.6.1, Seção 6.6.2 e Seção 6.6.3, respectivamente.

Por fim, apesar dos cuidados tomados em relação aos quatro tipos de validade, citados nesta Seção, não se pode afirmar que os resultados encontrados podem ser generalizados para modelos maiores, uma vez que os modelos testados são considerados pequenos, conforme detalhado nas Tabelas 6 e 7.

7 CONCLUSÃO

Este trabalho apresentou uma técnica de composição de modelos de processos de negócio baseada nas limitações das referências mais importantes relacionadas ao tema da pesquisa. Inicialmente foram levantadas questões sobre as técnicas de composição e identificação de similaridades entre os elementos de modelos BPM, que buscam investigar como os problemas são resolvidos na prática e o que o estado da arte já produziu nessa área. Esse levantamento demonstrou que não foram identificadas ferramentas de composição de modelos de processos de negócio disponíveis na indústria e que os mecanismos produzidos pela academia podem ser evoluídos.

Uma avaliação foi realizada com analistas que resolveram seis cenários de composição de modelos de processos de negócio sem apoio de uma ferramenta que os auxiliasse na composição, onde foram contabilizados o esforço e as erros produzidos. O resultado dessa avaliação mostrou que compor modelos BPMN é uma atividade difícil, onerosa e propensa a erros, tanto sintáticos como semânticos.

Uma ferramenta de composição de modelos de processos de negócio foi desenvolvida para pôr em prática a técnica proposta nesta pesquisa, cujo objetivo estava em verificar se, de fato, essa técnica melhora a precisão e reduz o esforço de composição. Para avaliar o desempenho da técnica, um novo experimento foi realizado com os mesmos participantes do primeiro experimento.

Testes estatísticos foram aplicados para confrontar os dados dos experimentos, sendo que esses resultados mostraram que os participantes conseguiram reduzir o tempo médio de composição de 13,4 minutos para 2,3 minutos, diminuição da taxa de erros em 567% ao reduzir de 207 para 31 erros e aumento da corretude em 122% ao produzir 60 modelos ideais (desejados) a mais no segundo experimento, demonstrando assim que a técnica proposta pode proporcionar diminuição de custos e aumento de produtividade.

Com a demonstração da redução de esforço e da taxa de erros e do aumento da corretude na realização de composição de modelos de processos de negócio, através das avaliações realizadas, conclui-se que a ferramenta Common, implementada com base na técnica proposta neste trabalho, responde às questões de pesquisa levantadas neste trabalho.

7.1 Contribuições

Esta pesquisa propôs uma técnica de composição de modelos de processos de negócio que comprehende um conjunto composto das melhores técnicas disponíveis na literatura, bem como uma ferramenta que permite ao analista interagir durante a integração dos modelos sensíveis à notação BPMN 2.0. Com essa ferramenta, analistas conseguem integrar modelos com redução de esforço e minimização de erros, proporcionando redução de custos e ganho de produtividade, estabelecendo assim a primeira contribuição científica deste trabalho.

A segunda contribuição científica deste trabalho é a produção de conhecimento empírico sobre composição de modelos de processos de negócio, apresentada pelos estudos experimentais de composição realizados tanto de forma manual como de forma semiautomática, que proporcionaram mensurar o esforço e a taxa de erros que analistas empenham para compor modelos.

Por último, através de um estudo comparativo, entre os trabalhos relacionados, construiu-se uma base de conhecimento referente às técnicas de composição de modelos de processos de negócio mais usadas na literatura. Este estudo reuniu em um único trabalho as

qualidades e deficiências de cada mecanismo de composição analisado, tornando-se assim um referencial de pesquisa sobre esse tema.

7.2 Limitações e trabalhos futuros

Neste trabalho foram levantadas questões de composição de modelos de processos de negócio que foram tratadas ao longo de seu desenvolvimento. Embora as contribuições desta pesquisa sejam perceptíveis, a abordagem precisa ser melhorada, uma vez que o tempo de desenvolvimento deste trabalho não é suficiente para abranger todas as possibilidades existentes. Nesse sentido, são apresentadas a seguir as principais limitações identificadas, bem como as sugestões de trabalhos futuros que atacam essas fraquezas.

- **Estender a ferramenta Common.** A ferramenta proposta neste trabalho atua de forma isolada dos editores de modelagem de processos de negócio, de modo que a composição é realizada fora do ambiente de modelagem. Isso impede que o modelo que está sendo construído possa ser editado após a composição na própria ferramenta. Nesse caso, a sugestão de melhoria apontada se baseia na extensão da ferramenta Common em um *plug-in* que seja compatível com os principais editores *open source* de modelagem BPMN ou que a própria ferramenta Common permita editar o modelo produzido, tornando-se um editor de modelos completo.
- **Atender a especificação BPMN 2.0 por completo.** Apesar de compreender os principais elementos da notação BPMN 2.0, nem todos os elementos disponíveis na especificação são atendidos pela ferramenta proposta. Nesse sentido, sugere-se como trabalho futuro o atendimento por completo dos elementos BPMN 2.0, possibilitando que a ferramenta Common possa realizar a integração de qualquer modelo desenhado nessa notação, mesmo os mais complexos ou que usem elementos pouco comuns.
- **Encorpar os modelos dos cenários de avaliação.** Apesar de modelos de processos de negócio serem normalmente pequenos, principalmente porque descrevem fluxos em alto nível, os modelos utilizados nos cenários apresentados possuem quantidade reduzida de elementos e pouca riqueza de símbolos. Portanto, outra sugestão de trabalho futuro indica o uso de modelos semanticamente mais ricos, que contenham maior variedade de símbolos e quantidade de elementos para, com isso, também realizar testes de escalabilidade e desempenho de composição.

REFERÊNCIAS

- ARMAS-CERVANTES, Abel; BALDAN, Paolo; DUMAS, Marlon, GARCÍA-BAÑUELOS, Luciano. BP-Diff: A Tool for Behavioral Comparison of Business Process Models. In: Proceedings of the BPM Demo Sessions, 2014, Eindhoven. **Anais...** Eindhoven: CEUR, 2014.
- BALDAM, R.; VALLE, R.; ROZENFELD, H. Gerenciamento de Processos de Negócio BPM: Uma Referência para Implantação Prática. Rio de Janeiro: Elsevier, 2014.
- BALTZAN, P.; PHILLIPS, A. Sistemas de Informação: A Importância e as Responsabilidades do Pessoal de TI nas Tomadas de Decisões. Porto Alegre: Bookman, 2012.
- BARROS, B. T. Fusões, aquisições & parcerias. São Paulo: Atlas, 2001.
- BECKER, Michael; LAUE, Ralf. A Comparative Survey of Business Process Similarity Measures. **Journal Computers in Industry**, Amsterdam, v. 63(2), p. 148-167, 2012.
- BEZERRA, E. Princípios de Análise e Projeto de Sistemas com UML. Rio de Janeiro: Elsevier, 3 ed., 2015.
- BIANCHIN, L. A.; WICKBOLDT, J. A.; GRANVILLE, L. Z.; GASPARY, L. P. Similarity Metric for Risk Assessment in IT Change Plans. In: Proceedings of the 6th International Conference on Network and Service Management, 2010, Niagara. **Anais...** Niagara: CNSM, IEEE, 2010, p. 25–32.
- BIZAGI. BIZAGI Modeler. Disponível em: <<http://www.bizagi.com/en/products/bpm-suite/modeler>>. Acesso em: 1 nov. 2015.
- BJURSELL, Cecilia. Cultural Divergence in Merging Family Businesses. **Journal of Family Business Strategy**, v. 2(2), p. 69-77, 2011.
- BPM-CBOK. Guia para o Gerenciamento de Processos de Negócio - Corpo Comum de Conhecimento. v. 3.0, 1 ed., 2013.
- BRANCO, Moises A. C. **Managing Consistency of Business Process Models across Abstraction Levels**. 2014. 155 f. Teste (Doctor of Philosophy in Electrical and Computer Engineering). University of Waterloo, Ontario, Canada, 1014.
- BRAGHETTO, Kelly Rosa. **Técnicas de Modelagem para a Análise de Desempenho de Processos de Negócio**. 2011. 151 f. Teste (Doutorado em Ciências) – Instituto de Matemática e Estatística. Universidade de São Paulo, São Paulo, SP, 2011.
- BROCKE, J.; ROSEMANN, M. Manual de BPM: Gestão de Processos de Negócio. Porto Alegre: Bookman, 2013.
- BROCKE, J.; ROSEMANN, M. Handbook on Business Process Management 1: Introduction, Methods and Information Systems. Berlin Heidelberg: Springer-Verlag, 2 ed., 2015.
- BRUNET, Greg; CHECHIK, Marsha; EASTERBROOK, Steve; NEJATI, Shiva; NIU, Nan; MEHRDAD, Sabetzadeh. A Manifesto for Model Merging. In: Proceedings of the 2006

- International Workshop on Global Integrated Model Management, 2006, Shanghai. **Anais...** Shanghai: GAMMA, ACM, 2006, p. 5-12.
- BUSCHMANN, F.; MEUNIER, R.; ROHNERT, H.; SOMMERLAD, P.; STAL, M. Pattern-Oriented Software Architecture Volume 1: A System of Patterns. New York: Wiley, 1996.
- CAMPOS, A. Modelagem de Processos com BPMN. Rio de Janeiro: Brasport, 2 ed., 2014.
- CHINOSI, M.; TROMBETTA, A. BPMN: Introduction to the Standard. **Journal Computer Standards & Interfaces**, v. 34(1), p. 124-134, 2012.
- CLEMENTE, M. N.; GREENSPAN, D. S. Winning at Mergers and Acquisitions: The Guide to Market-Focused Planning and Integration. New York: John Wiley & Sons, 1998.
- CORMEN, H. Thomas; LEISERSON, E. Charles; RIVEST, Ronald L.; STEIN, Clifford. Introduction to Algorithms. London: MIT Press, 3 ed., 2009.
- COSTA, Eduard M. M. Redes de Petri e Aplicações aos Sistemas a Eventos Discretos. São Paulo: LTC, 2011.
- DAL FORNO, Ana J.; PEREIRA, Fernando A.; FORCELLINI, Fernando A.; KIPPER, Liane M. Value Stream Mapping: A Study about the Problems and Challenges Found in Literature from the Past 15 Years about Application of lean Tools. **The International Journal of Advanced Manufacturing Technology**, v. 72(5), p. 779-790, 2014.
- DAVIS, R.; BRABÄNDER, E. Aris Design Platform. Getting Started with BPM. London: Springer, 2007.
- DELGADO, A.; WEBER, B.; RUIZ, F. Continuous Improvement of Business Process Realized by Services Based on Execution Measurement. **Journal Computer and Information Science**, v. 275, p. 64-81, 2013.
- DIAS-DA-SILVA, Bento C.; SOJKA, P.; CHOI, K. S.; FELLBAUM, C.; VOSSEN, P. Wordnet.Br: An Exercise of Human Language Technology Research. In: Proceedings of the 3th International Wordnet Conference, 2005, Masarykova. **Anais...** Masarykova: BRNO, 2005, p. 301-303.
- DIJKMAN, Remco; DUMAS, Marlon; GARCÍA-BAÑUELOS, L. Graph Matching Algorithms for Business Process Model Similarity Search. **Lecture Notes in Computer Science**, v. 5701, p. 48-63, 2009.
- DIJKMAN, Remco; DUMAS, Marlon; DONGEN, Boudewijn; KÄÄRIK, Reina; MENDLING, Jan. Similarity of Business Process Models: Metrics and Evaluation. **Journal Information Systems**, v. 36(2), p. 498-516, 2011.
- DIJKMAN, Remco; LA ROSA, Marcello; REIJERS, Hajo. Managing Large Collections of Business Process Models - Current Techniques and Challenges. **Journal Computers in Industry**, vol. 63(2), p. 91-97, 2012.
- DRISCOLL, P. Decision Making in Systems Engineering and Management. New Jersey: Wiley, 2 ed., 2011.

ECLIPSE. Eclipse BPMN2 Modeler. Disponível em: <<https://www.eclipse.org/bpmn2-modeler/>>. Acesso em: 15 out. 2015.

EPING. Padrões de Interoperabilidade de Governo Eletrônico, 2015. Disponível em: <<http://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padroes-de-interoperabilidade>>. Acesso em: 11 out. 2015.

FARIAS, Kleinner; GARCIA, A.; WHITTLE, J.; CHAVEZ, C.; LUCENA C. Evaluating the Effort of Composing Design Models: A Controlled Experiment. **Journal Software & Systems Modeling**, v. 14(4), p. 1349-1365, 2014.

FAZAL-BAQAIE, M. **Structural Matching of Process Models for Change Detection**. 2009. 125 f. Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-Graduação em Ciência da Computação, Engenharia Elétrica e Matemática. Universidade de Paderborn, Paderborn, Alemanha, 2009.

FERNANDES, A.; ABREU, V. Implantando a Governança de TI: Da Estratégia à Gestão dos Processos e Serviços. Rio de Janeiro: Brasport, 4 ed., 2014.

GENON, N.; HEYMANS, P.; AMYOT, D. Analyzing the Cognitive Effectiveness of the BPMN 2.0 Visual Notation. **Lecture Notes in Computer Science**, v. 6563, p. 377-396, 2011.

GERTH, Christian; KÜSTER, Jochen M.; LUCKEY, Markus; ENGELS, Gregor. Detection and Resolution of Conflicting Change Operations in Version Management of Process Models. **Journal Software & Systems Modeling**, v. 12(3), p. 517-535, 2011.

GIL, Antônio C. Como Elaborar Projetos de Pesquisa. São Paulo: Atlas, 4 ed., 2002.

GOTTSCHALK, F.; VAN DER AALST, Will M.P.; JANSEN-VULLERS, M. Merging Event-driven Process Chains. **Lecture Notes in Computer Science**, v. 5331, p. 418-426, 2008.

GUERRINI, F.; ESCRIVÃO E.; CAZARINI, E.; PADUA, S. Modelagem da Organização: Uma Visão Integrada. Porto Alegre: Bookman, 2014.

HALLERBACH, A.; BAUER, T.; REICHERT, M. Issues in Modeling Process Variants with Provop. **Lecture Notes in Business Information Processing**, v. 17, p. 56-67, 2008.

HERNANDEZ-MATIAS, J.; VIZAN, A.; HIDALGO, A. Evaluation of Techniques for Manufacturing Process Analysis. **Journal of Intelligent Manufacturing**, v. 17(5), p. 571-583, 2006.

HITPASS, B. Business Process Management (BPM): Concepts and How Do Apply and Integrate it with IT. Santiago de Chile: BHH, 2014.

IBM. Business Process Manager (IBM). Disponível em: <<http://www-03.ibm.com/software/products/pt/business-process-manager-standard>>. Acesso em: 30 out. 2015.

- IVANOV, Sergey Y.; KALENKOVA, Anna A.; VAN DER AALST, Will M.P. BPMNDiffViz: A Tool for BPMN Models Comparison. In: Proceedings of the BPM Demo Sessions, 2015, Innsbruck. **Anais...** Innsbruck: CEUR, 2015.
- KO, R.; LEE, S.; LEE, E. Business Process Management (BPM) Standards: A Survey. **Business Process Management Journal**, v. 15, p. 744-791, 2009.
- LA ROSA, Marcello; DUMAS, Marlon; UBA, Reina; DIJKMAN, Remco. Merging Business Process Models. **Lecture Notes in Computer Science**, v. 6426, p. 96-113, 2010.
- LA ROSA, M.; DUMAS, M.; UBA, U.; & DIJKMAN, R. Business Process Model Merging: An Approach to Business Process Consolidation. **Journal ACM Transactions on Software Engineering and Methodology**, v. 22(2), art. 11, 2013.
- LEVENSHTEIN, Vladimir I. Binary Code Capable of Correcting Deletions, Insertions and Reversals. **Soviet Physics Doklady**, v. 10, p. 707-710, 1966.
- LEVENSHTEIN. Levenshtein Algorithm Implementation. Disponível em <<https://commons.apache.org/sandbox/commons-text/jacoco/org.apache.commons.text.similarity/LevenshteinDistance.java.html>>. Acesso em 18 jan. 2016.
- LOCHER, D. Value Stream Mapping for Lean Development: A How-to Guide for Streamlining Time to Market. New York: CRC Press, 2008.
- LUCINDA, M. Qualidade: Fundamentos e Práticas para Cursos de Graduação. Rio de Janeiro: Brasport, 2010.
- MARCONI, M. A.; LAKATOS, E. M. Fundamentos de Metodologia Científica. São Paulo: Atlas, 5 ed., 2003.
- MAZIERO, Erick. G.; PARDO, Thiago. A. DI FELIPPO, Ariani. DIAS-DA-SILVA, Bento C. **A Base de Dados Lexical e a Interface Web do TeP 2.0 - Thesaurus Eletrônico para o Português do Brasil**. In: VI Workshop em Tecnologia da Informação e da Linguagem Humana (TIL), p. 390-392, 2008.
- MENDLING, Jan; NÜTTGENS, Markus. EPC Markup Language (EPML) = AN XML-Based Interchange Format for Event-driven Process Chains (EPC). **Information Systems and e-Business Management**, v. 4(3), p. 245-263, 2006.
- MENZEL, Christopher; MAYER, Richard. The IDEF Family of Languages. **International Handbooks on Information Systems**, p. 215-249, 2006.
- MILI, H.; TREMBLAY, G.; JAOUDE, G.; LEFEBVRE, E.; ELABED, L.; BOUSSAIDI, G. Business Process Modeling Languages: Sorting Through the Alphabet Soup. **Journal ACM Computing Surveys**, v. 43(1), art. 4, 2010.
- MITCHEL, Mark L.; JOLLEY, Janina M. Research Design Explained. Belmont: Wadsworth Publishing, 8 ed., 2012.

MURATA, T. Petri Nets: Properties, Analysis and Applications. **Proceedings of the IEEE**, v. 55(4), p. 541-580, 1989.

NIEMANN, Michael; SIEBENHAAR, Melanie; SCHULTE, Stefan; STEINMETZ, Ralf. Comparison and Retrieval of Process Models Using Related Cluster Pairs. **Computers in Industry**, v. 63, p. 168-180, 2012.

OHLSSON, J.; HÄNDEL, P.; HAN, S.; WELCH, R. Process Innovation with Disruptive Technology in Auto Insurance: Lessons Learned from a Smartphone-Based Insurance Telematics Initiative. **Management for Professional**, p. 85-101, 2014.

OLIVEIRA, V. Machado; ROMAN, N. Trevisan. JWN-Br – Uma API Java para a WordNet.Br. In: 9º Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, 2013, Fortaleza. **Anais...** Fortaleza: STIL, p. 153-157, 2013.

OMG-BPDM. Formal Specifications: Business Process Definition Metamodel, v. 1.0, 2008.

OMG-BPMN. Formal Specifications: Business Process Model and Notation, v. 2.0.2, 2013.

OMG-UML. Formal Specifications: Unified Modeling Language, v 2.5, 2015.

ORACLE. Oracle Business Process Management. Disponível em:
[<http://www.oracle.com/br/technologies/bpm/overview/index.html>](http://www.oracle.com/br/technologies/bpm/overview/index.html). Acesso em: 30 Out. 2015.

PEDERSEN, T.; DHAN, S. P.; MICHELIZZI, J. WordNet::Similarity - Measuring the Relatedness of Concepts. In: American Association for Artificial Intelligence, 2004. **Anais...** AAAI, p. 1024–1025, 2004.

PILLAT, Raquel M.; OLIVEIRA, Toacy C.; ALENCAR, Paulo C.; COWAN, Donal D. BPMNt: A BPMN extension for specifying software process tailoring. **Journal Information and Software Technology**, v. 57, p. 95-115, 2015.

PROM. Process Mining (ProM). Disponível em: <<http://www.processmining.org/prom/start>>. Acesso em: 01 nov. 2015.

RASMUSSEN, W. U. Aquisições: Fusões & Incorporações Empresariais. São Paulo: Aduaneiras, 1989.

RAYN, Thomas. Estatística Moderna para Engenharia. Rio de Janeiro: Elsevier, 2011.

RECKER, Jan; SAFRUDIN, Norizan; ROSEMAN, Michael. How Novices Design Business Processes. **Journal of Information Systems**, v. 37, p. 557-573, 2012.

ROZMAN, T.; POLANČÍČ, G; HORVAT, R. Analysis of Most Common Process Modelling Mistakes in BPMN Process Models. In: FISCHER, Layna. 2008 BPM & Workflow Handbook: Spotlight on Human-Centric BPM. **Anais...** Florida: Future Strategies, Inc, 2008, p. 293-306.

SANTOS, Rodrigo C. **Um Modelo de Apoio à Tomada de Decisão para Análise de Similaridade e Unificação de Processos de Negócio**. 2014. 135 f. Tese (Doutorado em

Engenharia de Sistemas e Computação) – Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa em Engenharia. Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 2014.

SCHWEGMANN, Ansgar; LASKE, Michael. As-is Modeling and Process Analysis. **Journal Process Management**, p. 107-133, 2003.

SHADISH, William R.; COOK, Thomas D.; CAMPBELL, Donald T. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin, 2 ed., 2001.

SHESKIN, David J. Handbook of Parametric and Nonparametric Statistical Procedures. New York: Chapman & Hall/CRC, 3 ed., 2004.

SUN, Shuang; KUMAR, Akhil; YEN, John. Merging workflows: A New Perspective on Connecting Business Processes. **Journal Decision Support Systems**, v. 42, p. 844-858, 2006.

VAN DER AALST, Will M. P. Business Process Management Demystified: A Tutorial on Models, Systems and Standards for Workflow Management. **Lecture Notes in Computer Science**, v. 3098, p. 1-65, 2004.

VIDAL, P. G., MARTIN, I. H., RAGAZZI, P. G. Inovação de Empresas em Processo de Fusão: Um Estudo das Empresas de Serviços e Indústria. **Ciências da Administração**, v. 16(39), p. 27-48, 2014.

VISIO. Microsoft BPMN Visio Modeler. Disponível em: <<https://products.office.com/pt-br/visio/flowchart-software>>. Acesso em: 31 Out. 2015.

WEBER, B.; REICHERT, M.; MENDLING, J.; REIJERS, H. Refactoring Large Process Model Repositories. **Journal Computers in Industry**, v. 62(5), p. 467-486, 2011.

WHITE, S.; MIERS, D. BPMN Modeling and Reference Guide: Understanding and Using BPMN. Lighthouse Point: Future Strategies, 2008.

WOHLIN, Claes; RUNESON, Per; HÖST, Martin; OHLSSON, Magnus; REGNELL, Björn; WESSLÉN, Anders. Experimentation in Software Engineering. New York: Springer, 2012.

WONG, P. Compositional Development of BPMN. **Lecture Notes in Computer Science**, v. 8088, p. 97-112, 2013.

WONG, P.; GIBBONS, J. Formalisations and Applicatins of BPMN. **Journal Science of Computer Programming**, v. 76(8), p. 633-650, 2009.

WFMC-XPDL. XML Process Definition Language. v. 2.2. Disponível em: <<http://www.wfmc.org/standards/xpdl>>. Acesso em: 9 Nov. 2015.

YAN, Zhiqiang; DIJKMAN, Remco; GREFEN, Paul. Fast Business Process Similarity Search with Feature-Based Similarity Estimation. **Lecture Notes in Computer Science**, v. 6426, p. 60-77, 2010.

ZEMNI, Mohamed A.; MAMMAR, Amel; HADJ-ALOUANE, Nejib Ben. An Automatized Approach for Merging Business Process Fragments. **Journal Computers in Industry**, v. 82, p. 104-118, 2016.

ZHU, Xinwei; RECKER, Jan; ZHU Guobin; SANTORO, Flávia Maria. Exploring Location-Dependency in Process Modeling. **Business Management Journal**, v. 20, p. 794-815, 2014.