

Identificando Fontes de Dados em Modelos de Processos de Negócio com base em Elementos de BPMN

Alternative Title: Identifying Data Sources in Business Process Models based on BPMN Elements

Marcelo Balbinot
Programa de Pós-Graduação
em Computação
Instituto de Informática
Universidade Federal do Rio
Grande do Sul
Porto Alegre – RS – Brasil
mbalbinot@inf.ufrgs.br

Lucineia Heloisa Thom
Programa de Pós-Graduação
em Computação
Instituto de Informática
Universidade Federal do Rio
Grande do Sul
Porto Alegre – RS – Brasil
lucineia@inf.ufrgs.br

Marcelo Fantinato
Programa de Pós-Graduação
em Sistema de informação
Escola de Artes, Ciências e
Humanidades
Universidade de São Paulo
São Paulo – SP – Brasil
m.fantinato@usp.br

RESUMO

A Notação e Modelo de Processo de Negócio (BPMN - *Business Process Model and Notation*) fornece um mecanismo para apresentar informações sobre o fluxo de dados de um processo por meio de artefatos, tais como objetos, associações e repositórios de dados. No entanto, esses elementos apresentam poder de expressão limitado, principalmente quando se referem ao mapeamento das fontes de dados (ex.: serviço web) em um modelo de processo. Neste contexto, o presente trabalho propõe uma abordagem para associar fontes de dados a elementos de BPMN (ex.: tarefa de serviço). Tal abordagem deve servir como um guia para profissionais de gerenciamento de processos de negócio quanto à definição, na etapa de modelagem, da fonte de dados mais apropriada conforme os elementos de BPMN existentes no modelo de processo. Visando demonstrar os resultados, foram evidenciadas cinco correlações, que são chamadas de “definições”, seguidas por sua descrição textual e um exemplo de uso. Essas definições foram validadas por meio de uma pesquisa de opinião, cujos resultados são apresentados e discutidos neste artigo.

Palavras-Chave

Gerenciamento de Processos de Negócio, BPMN, Modelagem de Processos, Fontes de Dados.

ABSTRACT

Business Process Model and Notation (BPMN) provides a mechanism to display information about the data flow of a process through artifacts such as objects, associations and

data stores. However, these elements provide a limited capacity, especially when referring to the mapping of the data sources (e.g. web service) in a process model. In this context, this paper proposes an approach to associate data sources to BPMN elements (e.g. service task). Such approach aims to serve as a guide to business process management professionals in order to define, in design time, the most appropriate data source depending on the BPMN elements presented in a given process model. In order to demonstrate the results, five correlations were evidenced, which are called “definitions”, followed by their textual description and an example of use. These definitions were validated through a survey which have their result also presented and discussed during this research.

CCS Concepts

•Applied computing → Business process modeling;
•Information systems → Data exchange; •Theory of computation → Data integration;

Keywords

Business Process Management, BPMN, Process Design, Data Sources.

1. INTRODUÇÃO

A Notação e Modelo de Processos de Negócio (BPMN) usa um conjunto específico de elementos gráficos para descrever o processo e como ele é executado. Os principais elementos de BPMN são: tarefas, eventos, desvios e fluxos de sequência. São esses elementos que definem a estrutura básica e o comportamento de um processo de negócio [20].

Por meio de artefatos, BPMN provê um mecanismo para apresentar informações adicionais sobre o processo, porém, essas informações não impactam diretamente no fluxo do processo [20]. Exemplos desses artefatos incluem: objetos de dados, grupos e anotações de texto.

Os dados e documentos manipulados pelos processos são representados pelos objetos de dados, que representam o fluxo de informações de entrada e saída de uma tarefa [5]. Outra forma de representar dados no processo é via repositório

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SBSI 2017 June 5th – 8th, 2017, Lavras, Minas Gerais, Brazil

Copyright SBC 2017.

rio de dados, que representa um local para armazenamento de objetos de dados e que são persistidos durante a execução da instância do processo [5].

BPMN apresenta poder de expressão limitado quanto ao mapeamento de fontes de dados usados pelos processos [10]. Porém, sua semântica é insuficiente para expressar as dependências de dados em um modelo de processo [11]. Logo, a notação prevê a existência de dados, mas pressupõe que eles estejam em uma base de dados compartilhada, ou que sejam internos do processo. Contudo, a realidade é que os dados estão tipicamente espalhados em diferentes locais, sendo necessários esforços para que possam ser consumidos [17].

Além disso, em organizações do ramo financeiro, por exemplo, existe a necessidade diária de troca de informações com órgãos reguladores. Essa comunicação pode ocorrer de várias formas, desde a troca de arquivos até o acesso a serviços web, fazendo com que muitas das rotinas de acesso a dados devam ser adaptadas a essa necessidade.

Ainda, os Sistemas de Gerenciamento de Processos de Negócio (BPMS - *Business Process Management Systems*) muitas vezes são utilizados como uma plataforma de integração para combinar aplicações pré-existentes ou legadas [14]. No entanto, experiências práticas demonstram que o reúso de aplicações como sendo componentes de um processo de negócio não é uma tarefa fácil e nem completamente apoiada pelas ferramentas e métodos existentes [11, 14].

Consequentemente, a necessidade de as organizações em representar as fontes de dados, em seus processos de negócio, acaba por não ser atendida em vista de limitações de BPMN. Isso faz com que as organizações continuem dependentes de documentação externa ao processo ou de analistas e especialistas de domínio [14].

Neste contexto, alguns trabalhos citam as limitações de BPMN em relação ao mapeamento de dados [10, 11, 14, 15]. Entretanto, esses trabalhos focam no mapeamento de dados a fim de estruturá-los de forma semelhante a um diagrama entidade-relacionamento, mas mantém a dependência do processo a uma base de dados compartilhada, não abrangendo outras fontes de dados.

Por esses motivos, este trabalho propõe uma abordagem para relacionar fontes de dados (ex.: serviço web) a elementos de BPMN (ex.: tarefa de serviço) por meio de sua semântica. Essa abordagem visa servir como um guia para os profissionais de gerenciamento de processos de negócio na definição da fonte de dados mais apropriada de acordo com os elementos de BPMN presentes no modelo de processo, na fase de modelagem.

Essa abordagem torna-se útil na fase de modelagem, principalmente para as empresas que não automatizam seus processos, desejando documentá-los com uma maior riqueza de detalhes. Além disso, esse mapeamento poderá auxiliar na fase de implementação por abordar as fontes de dados que o processo a ser implementado deverá contemplar.

As hipóteses que motivam o desenvolvimento deste trabalho são: **(I)** é possível associar fontes de dados a elementos de BPMN a partir de suas semânticas; e, **(II)** usar essa associação como ponto de partida na definição da fonte de dados relacionada a um determinado conjunto de elementos de BPMN em tempo de modelagem.

Essas hipóteses foram avaliadas por análise da literatura referente às fontes de dados e aos elementos de BPMN, visando encontrar correlações entre suas definições. Posteriormente, foi feita a validação por meio de pesquisa de opinião,

aplicada a grupos de pessoas com conhecimento sobre gerenciamento de processos de negócios, BPMN e fontes de dados.

Este artigo está estruturado da seguinte forma: na seção 2 são apresentados os trabalhos relacionados, bem como a motivação que originou a pesquisa; a seção 3 traz o referencial teórico, relacionado aos conceitos usados neste trabalho; a seção 4 apresenta as definições propostas neste trabalho; a seção 5 apresenta os resultados obtidos por este trabalho; e, finalmente, a seção 6 contempla as conclusões e os trabalhos futuros referentes a este trabalho.

2. TRABALHOS RELACIONADOS E MOTIVAÇÃO

O uso de dados nas tarefas de um fluxo é considerado pela BPMN, porém, nem todo dado consumido em um processo de negócio tem origem interna no processo, sendo necessário identificá-lo entre dado interno ou externo ao fluxo [14]. Em se tratando de modelagem de dados, o artigo em questão propõe que cada tarefa tenha um esquema de entrada e de saída de dados, possibilitando assim, a modelagem e validação do fluxo de dados.

Um outro trabalho também cita a limitação sobre o mapeamento de dados de BPMN, propondo uma nova notação, chamada de BPDMN (*Business Process and Data Modeling Notation*) [10]. Nessa notação, os objetos de dados de BPMN são estendidos para se tornarem estruturados, permitindo a representação de diversos níveis de abstração de dados, como no caso dos repositórios de dados, que passam a possuir uma estrutura complexa, similar a um modelo entidade-relacionamento.

Já em [11] os objetos de dados são estendidos para representarem, por meio de anotações, uma analogia a bases de dados, permitindo definir relacionamento $1:n$ ou $m:n$ com outros objetos, além de conceitos como chave primária e chave estrangeira. Além disso, introduz uma abordagem para derivar consultas SQL dos modelos de processo gerados por meio da extensão.

Os trabalhos relacionados, reportados nesta seção, estão focados na necessidade de acessar dados que estão armazenados em bases de dados, às quais os processos possuem acesso direto. Porém, nas organizações, tal acesso nem sempre é possível devido à existência de programas legados, aplicações terceirizadas ou regras de isolamento de dados e aplicações. Assim, há necessidade de as aplicações possuírem integração com múltiplas fontes de dados, tornando-a um desafio para os sistemas atuais [2, 9].

Outro problema é a necessidade de documentação do processo em conjunto com suas fontes de dados, no intuito de manter uma base de conhecimento atualizada e confiável. Essa documentação permite mensurar o impacto propagado ao se decidir fazer uma alteração em um dos processos ou fontes de dados em questão.

Sendo assim, a possibilidade de relacionar fontes de dados a elementos de BPMN se torna de grande valia para as organizações, fazendo com que essa dependência consiga ser induzida diretamente por meio do modelo de processo, evitando assim, a dependência de documentação externa ou de analistas e especialistas de domínio [14].

3. FUNDAMENTAÇÃO TEÓRICA

Esta seção serve como embasamento teórico para o trabalho aqui apresentado. Nele se busca contextualizar os prin-

principais conceitos necessários para a compreensão da abordagem proposta, tais como: fontes de dados e elementos de BPMN. Tais conceitos são usados posteriormente para, via análise semântica, evidenciar o relacionamento entre fontes de dados e elementos de BPMN.

3.1 Fontes de Dados

Autores como [6, 7] evidenciam o uso de algumas fontes de dados como os principais estilos de integração entre aplicações. Essas fontes estão descritas a seguir, sendo elas: *Base de Dados Compartilhada*, *Transferência de Arquivos*, *Chamada de Procedimento Remoto* e *Serviço Web*.

Com a intenção de avaliar o uso na prática das fontes de dados citadas como as mais recorrentes na literatura, optou-se por fazer um estudo de caso de uma organização do ramo financeiro. Para isso, foram analisados os modelos de processos de negócio da organização, além da coleta de evidências por meio da documentação anexa aos processos.

Algumas fontes de dados podem ser consideradas controversas neste artigo (ex. invocação remota e serviços web), porque elas se referem a chamadas de métodos, encapsulados a partir de uma aplicação externa, e não necessariamente precisam ter uma entrada ou saída de dados. No entanto, neste artigo são consideradas como fontes de dados porque fornecem a funcionalidade necessária para alcançar dados que estão fora do contexto do processo.

3.1.1 Base de Dados Compartilhada

Base de dados compartilhada, conforme [4, 8, 16], é um repositório de dados centralizado que pode ser acessado por vários sistemas ou componentes, uma vez que previamente permitido. Por acessarem aos mesmos dados, é exigido que os diferentes elementos do sistema concordem com um modelo de dados comum [7].

O principal uso desse modelo é quando existem aplicações ou componentes que têm dados em comum, usando a ideia de uma única base de dados compartilhada e centralizada, permitindo a otimização do uso de dados e o intercâmbio de informações [5]. Essa fonte de dados é frequentemente adotada para facilitar a integração desses aplicativos e componentes, pois é um modelo de fácil implementação nas aplicações, visando evitar a duplicação de dados entre elas.

No entanto, um problema com a base de dados compartilhada é a dificuldade de definir e impor limites claros entre os sistemas [7]. Também há um problema com o esquema de dados que, uma vez definido, se torna difícil de modificar, pois pode quebrar a integridade com os outros sistemas, fazendo com que sejam criadas variadas adaptações no esquema para apoiar todos os sistemas integrados.

3.1.2 Transferência de Arquivos

O acesso a dados com base na transferência de arquivos ocorre quando uma aplicação grava um arquivo, que, posteriormente será lido por outra [7, 16]. As aplicações precisam estar de acordo sobre o nome dos arquivos, o diretório em que são colocados, o formato e as regras sobre como criar, ler, atualizar e, se necessário, excluir os arquivos.

Ainda, segundo [7], é necessário um certo esforço para produzir e processar um arquivo, não sendo uma boa prática usar esse tipo de comunicação com frequência. Geralmente é usado de acordo com uma diretriz de negócio, que define um ciclo para a sua geração: diário, mensal, etc.

Os formatos típicos de arquivos usados, segundo [6], in-

cluem DBF (*dBASE database file*), XML (*eXtensible Markup Language*) e TXT (*plain text*). Sistemas *Mainframe* normalmente usam arquivos de dados baseados nos formatos de sistema de arquivos do *COBOL*, e os sistemas *Unix* usam arquivos baseados em texto [7].

Quanto a sua geração, o arquivo de dados pode ter origem na mesma rede em que se encontra o programa consumidor, ou, pode ser gerado e transmitido em seguida. Neste caso, algum aplicativo deve assumir a responsabilidade de transferir o arquivo do local que foi gerado para um local previamente acordado, usando FTP (*File Transfer Protocol*) ou outro protocolo de transferência de arquivos [7].

3.1.3 Chamada de Procedimento Remoto

Existem 2 paradigmas relacionados à chamada de procedimentos remotos: Chamada de Procedimento Remoto (RPC - *Remote Procedure Call*), e; Invocação de Método Remoto (RMI - *Remote Method Invocation*). O RPC estende a abstração de programação comum da chamada de procedimento para ambientes distribuídos, permitindo que um processo chame um procedimento em um nó remoto como se fosse local [1, 3]. RMI fornece a mesma funcionalidade que o RPC, no entanto, com vantagens adicionais no uso de conceitos de programação orientada a objetos em sistemas distribuídos e permitindo o uso de referências de objetos como parâmetros em invocações remotas [19, 3].

Essa forma de acesso ocorre quando uma aplicação expõe algumas de suas funcionalidades para que possam ser acessadas remotamente por outros aplicativos como uma rotina remota [7]. A comunicação ocorre em tempo real e de forma síncrona. Quando o procedimento é chamado, o ambiente chamador é suspenso, os parâmetros são passados via rede para o ambiente onde se encontra a aplicação que será executada, e o procedimento, então, executado [1].

A chamada de procedimento remoto atua por meio do princípio de encapsulamento [7], mantendo os dados e regras de negócio sob responsabilidade da aplicação proprietária, garantindo assim, a integridade dos dados. Dessa forma, qualquer acesso a dados deve passar pela aplicação proprietária via invocação remota.

3.1.4 Serviços Web

Serviços web (*web services*) fornecem um meio padrão de interoperabilidade entre diferentes aplicações e *software*, executando em uma variedade de plataformas e *frameworks*. As operações de um serviço web podem ser fornecidas por uma variedade de recursos diferentes, por exemplo, aplicativos, objetos ou bancos de dados [3].

Um serviço web pode ser gerenciado por um servidor web, páginas da web ou pode ser um serviço inteiramente separado [3]. Os serviços web têm baixo acoplamento entre as aplicações, no entanto, fornecem serviços de alto valor agregado [18].

Os serviços web são caracterizados pela sua grande interoperabilidade e extensibilidade devido ao uso de XML para a representação de dados externos e ao empacotamento de mensagens trocadas entre clientes e serviços web [18, 3]. O protocolo SOAP (*Simple Object Access Protocol*) especifica as regras de uso de XML para empacotar as mensagens e passá-las via HTTP ou outro protocolo.

Geralmente, um serviço web fornece uma descrição de serviço via WSDL (*Web Service Description Language*), que inclui uma definição de interface e outras informações, como

a URL do servidor. Essa descrição é usada como base para um entendimento comum entre cliente e servidor quanto ao serviço oferecido [3, 18].

3.2 Elementos de BPMN

BPMN usa um conjunto específico de elementos gráficos para descrever o processo e como ele é executado. Seus principais elementos são: tarefas, eventos, desvios e fluxos de seqüência. Esses elementos definem a estrutura básica e o comportamento de um processo de negócio [20].

A partir da especificação formal de BPMN [12], foi realizada a análise semântica, visando identificar a correlação existente entre os elementos da notação, em conjunto com as fontes de dados referenciadas na seção 3.1. A partir dessa análise, foram selecionados os elementos de BPMN apresentados a seguir, os quais são considerados na abordagem proposta neste artigo, cuja Figura 1 demonstra a representação gráfica desses elementos.

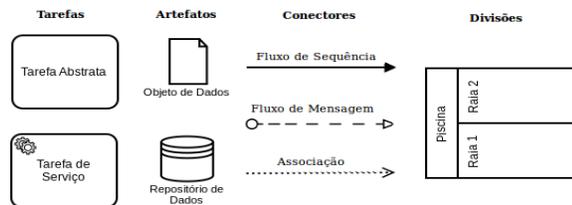


Figure 1: Elementos de BPMN

3.2.1 Tarefas

Tarefa (*Task*) é um passo atômico dentro de um fluxo de processo [12]. Uma tarefa é usada quando o trabalho dentro de um processo não pode ser desmembrado a um nível menor de detalhes [5, 13]. Geralmente, usuários finais ou aplicações são usadas para executar a tarefa.

Existem diferentes tipos de tarefas identificadas no âmbito de BPMN para separar os tipos de comportamento que as tarefas podem representar. Quando não é especificado o tipo de uma tarefa, ela é chamada de **Tarefa Abstrata** (*Abstract Task*) [12].

A **Tarefa de Serviço** (*Service Task*) faz uso de algum tipo de serviço para ser executada. Dessa forma, faz-se necessário especificar na tarefa como se comunicar com o aplicativo externo que irá fazer essa execução [5].

Seja um sistema complexo ou um aplicativo simples, na perspectiva do BPMS é necessário apenas que o aplicativo externo forneça uma interface de serviço que possa ser usada pela tarefa (ex.: um serviço web ou uma aplicação automatizada) [12].

A operação chamada pela tarefa de serviço pode ser síncrona ou assíncrona [12]. Em uma operação síncrona, o serviço espera uma mensagem de requisição e responde com uma mensagem de resposta, uma vez concluída a operação [5]. Em uma operação assíncrona, o serviço espera uma mensagem de requisição, porém, não enviará uma mensagem de resposta.

Por padrão, BPMN usa a tecnologia de serviços web para implementar interfaces de serviço. Mas, outras implementações são possíveis (ex.: Java RMI) [5].

3.2.2 Artefatos

Um requisito tradicional da modelagem de processos é ser

capaz de modelar os itens (físicos ou de informação) que são criados, manipulados e usados durante a execução do processo [12]. Um aspecto importante disso é a habilidade de capturar a estrutura de dados e consultar ou manipular essa estrutura.

Os **Objetos de Dados** (*Data Objects*) representam o fluxo de informação, entrando e saindo das tarefas. Essa informação pode ser física (ex.: um pedido ou uma carta), ou um artefato eletrônico (ex.: um e-mail ou arquivo) [5]. Ele é vinculado ao processo por meio de setas abertas e pontilhadas, cuja direção da seta indica se a tarefa está lendo ou gravando o conteúdo do objeto de dados.

Uma notação abreviada para a passagem de um objeto de dados de uma tarefa para a outra consiste em ligar o objeto de dados diretamente no fluxo de seqüência entre as duas tarefas. Essa ligação é feita por meio de uma associação não-dirigida, representada por uma linha tracejada sem setas [5].

O objeto de dados também é usado para representar o estado do dado, porém, essa representação é opcional. A representação do estado do dado é feita colocando o estado entre colchetes, tal como segue: “Data object [state]”.

Quanto ao ciclo de vida, os objetos de dados estão vinculados à instância de seu processo pai, existindo apenas durante a execução dessa instância.

O **Repositório de Dados** (*Data Store*) é um local que contém objetos de dados que necessitam ser persistidos para além do escopo do processo, em, por exemplo, uma base de dados ou um diretório [5]. Similar aos objetos de dados, é ligado às tarefas do processo via associações de dados.

De acordo com [5], os repositórios de dados não são diretamente interpretados por um BPMS, porém o BPMS assume a existência de um serviço dedicado que pode acessar esses repositórios de dados. Assim, o BPMS se conectará a esse serviço ao invés de diretamente com os repositórios de dados.

3.2.3 Conectores

O **Fluxo de Seqüência** (*Sequence Flow*) é usado para indicar a ordem em que as atividades são executadas em um processo [12]. Para que a comunicação ocorra por meio do fluxo de seqüência, as tarefas devem estar em uma mesma piscina, podendo ou não estarem separadas por raias [11].

O **Fluxo de Mensagem** (*Message Flow*) é usado para apresentar o fluxo de mensagens entre dois participantes de um processo [5, 11, 12]. A seta indica a direção do fluxo da informação entre os dois participantes.

A **Associação** (*Association*) é usada para conectar informações e artefatos utilizados no modelo do processo [12]. A seta indica a direção do fluxo de informações, representando a entrada ou saída de objetos de dados das atividades [5].

3.2.4 Divisões

As divisões são usadas para identificar os responsáveis por executar cada atividade de um processo de negócio.

A **Piscina** (*Pool*) é a representação gráfica de um participante do processo [12].

A **Raia** (*Lane*) é uma sub-partição dentro de uma piscina, usada para organizar e categorizar as atividades [12]. Uma piscina comporta tantas raias quanto forem necessárias para o processo.

Não há uma limitação de tipo de participante a ser modelado por piscina ou raia. Segundo [5, 11], um uso típico da piscina é para representar uma organização, uma vez que a raia representaria um departamento ou um sistema.

4. RELACIONANDO FONTES DE DADOS A ELEMENTOS DE BPMN

O mapeamento de fontes de dados em modelos de processo de negócio a partir de elementos de BPMN refere-se à correlação existente entre a semântica dos tipos de tarefa e artefatos em conjunto com as fontes de dados. Neste estudo, foram evidenciadas cinco correlações, as quais são chamadas de *definições*, e apresentadas nas subseções abaixo.

Tais definições foram derivadas tendo como base, inicialmente, pesquisa bibliográfica, buscando coletar informações sobre as fontes de dados e os elementos de BPMN. Foi evidenciada a existência das fontes de dados coletadas na literatura, tendo como ponto de partida o estudo de caso de uma instituição do ramo financeiro. A partir disso, foi realizada a análise semântica das relações entre o conjunto de fontes de dados e os elementos de BPMN, resultando nas definições apresentadas nesta seção.

Para cada definição, foram elencados dados com o intuito de explicar a relação entre a fonte de dados e os elementos de BPMN relacionados. Esses dados estão subdivididos em: “*Descrição*”, que explica o fluxo de elementos de BPMN no modelo do processo; “*Fonte de dados*”, referenciando a fonte de dados de acordo com a seção 3.1, e; “*Evidências*”, que especifica, via definição dos elementos de BPMN, qual a relação semântica que possui com a fonte de dados.

Essas definições são seguidas também por um exemplo, que consiste em um fragmento de processo, elaborado por meio da notação de BPMN. Esse fragmento de processo apresentado visa complementar as informações relacionadas no campo descrição, a fim de facilitar a compreensão.

4.1 Definição 1 - Dados Internos

Descrição: comunicação entre duas tarefas (T1 e T2), ligadas por um fluxo de sequência e sem indicação de uso de dados externos, como representado na Figura 2.

Fonte de dados: dados internos.

Evidências: de acordo com BPMN, todas as tarefas têm, em sua especificação, entrada e saída de dados [12]. Essas entrada e saída de dados podem ser exibida no modelo de processo, no entanto, ela também pode ser mantida oculta.

Como apresentado na Figura 2, os dados manipulados pelas tarefas T1 e T2 não são apresentados no modelo de processo. Isso dá um forte indício de que os dados que estão sendo manipulados pelas tarefas em questão são internos ao processo.

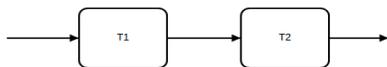


Figure 2: Definição 1 - dados internos

4.2 Definição 2 - Transferência de Arquivos

Descrição: comunicação entre tarefas (T1 e T2) onde há, explicitamente, o envio ou recepção de um objeto de dados, como representado na Figura 3.

Fonte de dados: transferência de arquivos.

Evidências: de acordo com [5], os objetos de dados representam informações que fluem dentro e fora das atividades. Essas informações podem ser físicas (ex.: um documento) ou eletrônica (ex.: um arquivo).

Conforme apresentado na Figura 3, os dados manipulados pelas tarefas T1 e T2 são transferidos por meio de um ob-

jeito de dados. Isso dá uma forte indicação de que os dados manipulados pelas tarefas em questão são provenientes de uma transferência de arquivos.

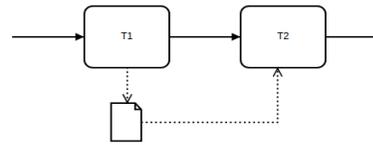


Figure 3: Definição 2 - transferência de arquivos

4.3 Definição 3 - Base de Dados Compartilhada

Descrição: comunicação entre tarefas (T1 e T2) onde há, explicitamente, dados sendo enviados ou recebidos via repositório de dados, como representado em Figura 4.

Fonte de dados: base de dados compartilhada.

Evidências: de acordo com [5], um repositório de dados é um local que contém objetos de dados (ex.: uma base de dados). As tarefas do processo podem ler ou gravar objetos de dados de um repositório de dados, fazendo com que esses dados sejam persistidos para além da duração da instância de processo.

Conforme apresentado na Figura 4, os dados manipulados pelas tarefas T1 e T2 são transferidos para um repositório de dados. Isso dá uma forte indicação de que a fonte de dados das tarefas em questão é uma base de dados compartilhada.

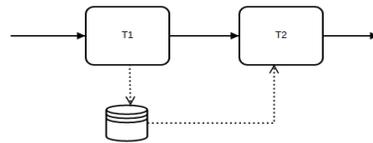


Figure 4: Definição 3 - base de dados compartilhada

4.4 Definição 4 - Chamada de Procedimento Remoto

Descrição: tarefas de serviço (T1 e T2) sendo executadas em 2 raias diferentes, o que pode representar sistemas distintos, como apresentado na Figura 5.

Fonte de dados: chamada de procedimento remoto.

Evidências: como tarefas de serviço são chamadas em diferentes raias da mesma piscina, é possível dizer que os dados entre as tarefas estão sendo passados via troca de mensagens, sem compartilhar um repositório de dados. BPMN não especifica o uso das raias, porém, cita que a distinção entre raias pode indicar que a comunicação ocorre entre diferentes participantes ou sistemas dentro da mesma organização [12], como apresentado na Figura 5, dando-nos uma forte indicação de uma invocação remota.

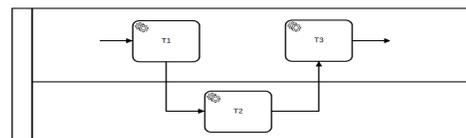


Figure 5: Definição 4 - chamada de procedimento remoto

Essa evidência também é complementada pela definição da tarefa de serviço que, de acordo com o [12], é uma tarefa que fornece algum tipo de serviço, seja um serviço da web ou uma aplicação automatizada. Outro complemento à definição é que as tarefas estão conectadas por meio de um fluxo de sequência, o que faz com que o controle seja passado para a próxima tarefa e atenda à definição da chamada remota, de acordo com [1].

4.5 Definição 5 - Serviço Web

Descrição: tarefas de serviço (T1 e T2) sendo executadas em 2 piscinas distintas, como apresentado na Figura 6.

Fonte de dados: serviço web.

Evidências: como tarefas de serviço são chamadas em piscinas diferentes, é possível dizer que os dados entre as tarefas estão sendo passados via troca de mensagens, sem o compartilhamento de um repositório de dados. Essa evidência também é complementada pela definição da tarefa de serviço que, de acordo com o [12], é uma tarefa que fornece algum tipo de serviço, seja um serviço da web ou uma aplicação automatizada. Outro complemento à definição é que as tarefas estão conectadas por meio de um fluxo de mensagens, que não passa o controle para a tarefa chamada, atendendo à definição de serviço da web.

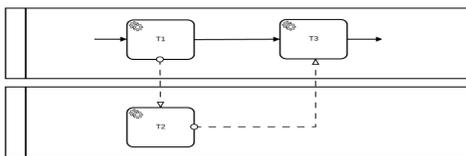


Figure 6: Definição 5 - serviço web

5. AVALIAÇÃO DOS RESULTADOS

Com o intuito de validar as definições propostas e as fontes de dados discutidas neste artigo, foi elaborada uma pesquisa de opinião¹, que foi disponibilizada a um público-alvo com conhecimentos em Gerenciamento de Processos de Negócios e BPMN. A disponibilização da pesquisa ocorreu via redes sociais, alunos da disciplina de Gerenciamento de Processos de Negócio e contatos na área.

Na pesquisa de opinião em questão, os entrevistados foram questionados sobre seu conhecimento nos temas relacionados ao trabalho. Questionou-se sobre o uso das fontes de dados abordadas na pesquisa e a validação dos fragmentos de processo apresentados na seção 4.

A pesquisa foi respondida por 23 pessoas, subdivididas entre alunos, professores, especialistas em software, especialistas em processos e outros. Todas elas demonstraram conhecimento no gerenciamento de processos, no entanto, cinco (21,7%) delas não possuem conhecimento BPMN, como pode ser visto na Figura 7.

De acordo com a Figura 7, há três pessoas classificadas em “Outros”. Isso ocorreu porque os entrevistados não se enquadraram nas categorias existentes, especificando a função correta no campo previsto para isso, sendo eles: um Coordenador de TI; um Analista de Serviços de TI; e um Gerente de TI.

Nas seções a seguir, as respostas aos tópicos discutidos neste artigo são avaliadas. A seção 5.1 contém os resultados

¹Disponível em: <https://goo.gl/forms/8AbOhfpmjgJF8Klu2>

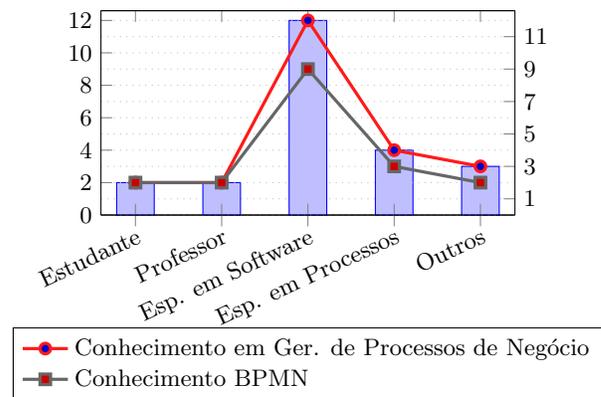


Figure 7: Público alvo da pesquisa de opinião

relativos às fontes de dados contempladas pela pesquisa e, na seção 5.2, a avaliação das definições propostas.

5.1 Fontes de Dados

A partir das fontes de dados apresentadas na seção 3.1, foi questionado, na pesquisa, se elas realmente são usadas pelos entrevistados, usando uma questão de múltipla escolha com as fontes de dados abordadas neste artigo. Além disso, também foi questionado, por meio de um campo aberto, se os entrevistados usavam outras fontes de dados.

O resultado dessa pergunta pode ser visto na Figura 8, que apresenta a porcentagem de uso de cada fonte de dados para a amostra coletada. Dessa forma, é possível constatar que as fontes de dados escolhidas por este trabalho foram validadas pelos participantes.

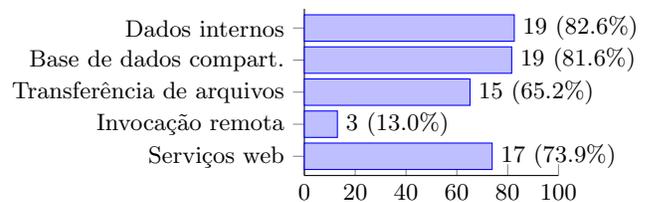


Figure 8: Avaliação do uso das fontes de dados

No gráfico apresentado, pode-se verificar que todos os entrevistados usam, ou têm conhecimento, pelo menos em uma das fontes de dados abordada no trabalho. Além disso, nenhum dos entrevistados respondeu outras fontes de dados. Logo, é possível afirmar que as fontes de dados escolhidas são as mais usadas para integrar sistemas, de acordo com os participantes da pesquisa.

5.2 Definições

Para validar as definições propostas no artigo, os fragmentos de processos apresentados na seção 4 foram exibidos aos entrevistados juntamente com uma questão de múltipla escolha, que apresentava as fontes de dados discutidas neste trabalho.

Para cada fragmento de processo, foi solicitado aos entrevistados que, de acordo com seu conhecimento em modelagem de processos de negócios, indicassem quais fontes de dados o fragmento lhes dava evidências a pertencer. Abaixo discutiremos o resultado de cada um dos fragmentos apresentados na seção 4.

Referente à **Definição 1**, apresentada na seção 4.1, obtivemos 73.9% das respostas indicando que ela representa o uso de dados internos. Esse resultado, juntamente com as demais respostas são apresentados na Figura 9.

Além disso, através do campo “Justificativa”, disponibilizado na pesquisa para cada um dos fragmentos apresentados, os respondentes citam que o fragmento não possui indicação de uso de dados externos e que as tarefas devem estar usando a mesma fonte de dados, já que esses dados estão trafegando entre as tarefas via fluxo de sequência. De acordo com isso, podemos inferir que a Definição 1 foi validada pelos participantes.

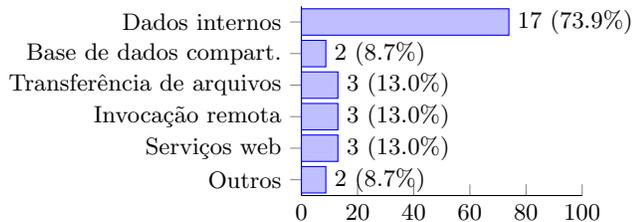


Figure 9: Avaliação da Definição 1

Em relação à **Definição 2**, apresentada na seção 4.2, recebemos como resultado 78.3% dos respondentes indicando o uso de transferência de arquivos. Esse resultado, juntamente com as demais respostas são apresentados na Figura 10.

Além disso, a amostra justificou a resposta indicando que o objeto de dados passado entre as tarefas indica a transferência de arquivos, o qual foi gerado na tarefa T1 e consumido na tarefa T2, conforme apresentado na Figura 3. De acordo com isso, pode-se afirmar que a Definição 2 foi validada pelos participantes.

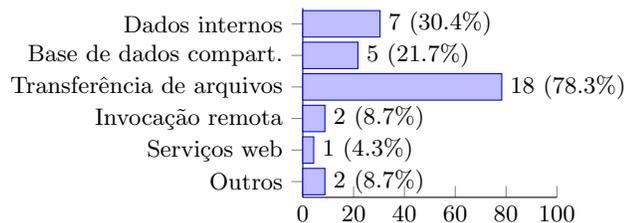


Figure 10: Avaliação da Definição 2

Quanto à **Definição 3**, apresentada na seção 4.3, obtivemos 91.3% das respostas indicando o uso de base de dados compartilhada. Esse resultado, juntamente com as demais respostas são apresentados na Figura 11.

Como justificativa, os respondentes citaram o uso de alguma base de dados acessível a ambas as tarefas. De acordo com isso, temos a indicação de que a Definição 3 está validada pelos participantes.

Em relação à **Definição 4**, apresentada na seção 4.4, 56.5% dos respondentes dizem que ela representa o uso de invocação remota, como apresentado na Figura 12. Porém, para 69.6% dos respondentes o fragmento representa o uso de um serviços web.

Como justificativa, os respondentes citaram, em sua maioria, o uso de algum serviço para acesso aos dados, não especificando o tipo de serviço usado. Esse resultado ocorreu pela interpretação isolada da tarefa de serviço.

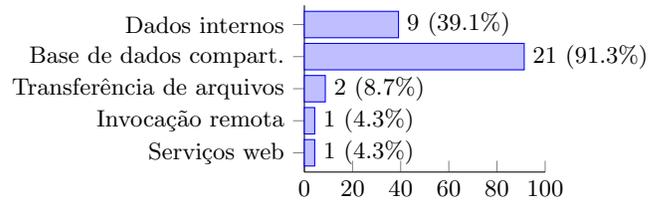


Figure 11: Avaliação da Definição 3

Assim, pode-se afirmar que a Definição 4 não foi validada pelos participantes, uma vez que a definição e seu resultado são discutidos posteriormente na seção 5.3.

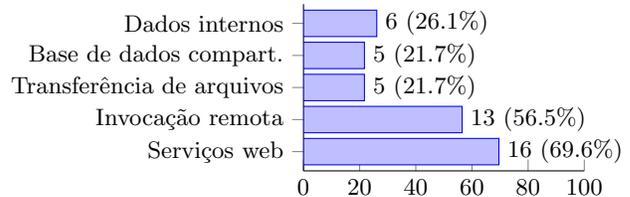


Figure 12: Avaliação da Definição 4

No que se refere à **Definição 5**, apresentada na seção 4.5, obtivemos 73.9% dos respondentes dizendo que representa o uso de serviços web, como apresentado na Figura 13. Entretanto, para 56.5% dos respondentes o fragmento representa o uso de um invocação remota.

Como justificativa, recebemos resultados muito semelhante aos obtidos na Definição 4, onde os respondentes indicam o uso de algum tipo de serviço. Porém, não temos uma margem significativa entre as duas opções mais indicadas pelos participantes, o que indica que a Definição 5 não foi validada por estes, sendo que a definição e seu resultado são discutidos posteriormente na seção 5.3.

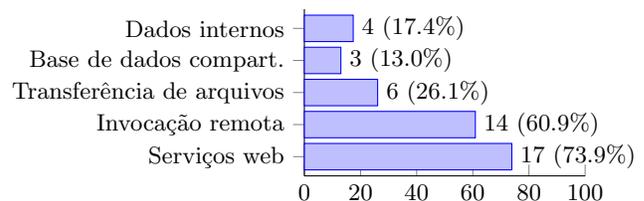


Figure 13: Avaliação da Definição 5

5.3 Discussão

O principal objetivo deste trabalho é relacionar as fontes de dados aos elementos de BPMN com base na sua semântica. Para que isso seja possível, cinco definições foram apresentadas no decorrer deste trabalho.

Para tornar essas definições um ponto de partida para profissionais que trabalham com Gerenciamento de Processos de Negócio, foi elaborada uma representação gráfica (ex.: Figura 2) contendo um fragmento de um processo para facilitar sua compreensão. A partir da elaboração dessas definições gráficas e da base teórica referente às fontes de dados, essas foram apresentadas a uma amostra para que pudessem validar as definições.

As definições 1, 2 e 3, que tratam de dados internos, transferência de arquivos e base de dados compartilhada, respecti-

vamente, foram validadas conforme esperado pela pesquisa. No entanto, as definições 4 e 5, que tratam da invocação remota e serviços web, respectivamente, apresentaram resultados muito semelhantes, indicando o uso de serviços web em ambos os fragmentos de processo.

Esses resultados podem ter sido causados pela definição de tarefa de serviço, que aponta para o uso tanto de serviços web, quanto de invocação remota. Além disso, outro fator que pode ter influenciado esse resultado foi o número reduzido de pessoas usando invocação remota, representado apenas por 13.0% da amostra.

6. CONCLUSÕES E TRABALHOS FUTUROS

Neste artigo foram apresentadas evidências de relações semânticas entre os elementos de BPMN com as fontes de dados obtidas a partir da literatura. Isso tornou possível a elaboração de um conjunto de definições, que tem por objetivo ser o ponto de partida na definição de fontes de dados a ser associada à determinado conjunto de elementos de BPMN em tempo de modelagem.

Por meio das definições apresentadas, é possível identificar a fonte de dados mais adequada para um determinado conjunto de elementos de um processo de negócios, analisando os elementos do modelo. A mesma possibilidade existe quando se deseja mapear, no modelo de processo de negócio, uma fonte de dados que já está sendo usada.

Como trabalhos futuros, pretende-se aprofundar as definições sobre invocação remota e serviços web, na tentativa de encontrar uma clara distinção entre os fragmentos de processos referentes a essas fontes de dados, conforme discutido na seção 5.3. Além disso, pretende-se expandir o conjunto de resultados com a identificação de novas definições.

7. REFERENCES

- [1] A. D. Birrell and B. J. Nelson. Implementing remote procedure calls. *ACM Transactions on Computer Systems (TOCS)*, 2(1):39–59, 1984.
- [2] B. Bouchra, B. Fouzia, and B. Chahinez. Data sources integration using viewpoint-based approach. In *Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication, IPAC '15*, pages 25:1–25:6, New York, NY, USA, 2015. ACM.
- [3] G. Coulouris, J. Dollimore, T. Kindberg, and G. Blair. *Distributed Systems: Concepts and Design*. Addison-Wesley Publishing Company, USA, 5th edition, 2011.
- [4] C. Date. Locking and recovery in a shared database system: An application programming tutorial. In *VLDB*, pages 1–15, 1979.
- [5] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers. *Fundamentals of business process management*, volume 1. Springer, 2013.
- [6] X. Guangbin, W. Jianfeng, Z. Gang, H. Yinfei, W. Shaoping, B. Shuo, et al. Fftp: A file-like data exchange method for high-liquid securities information based on extended fast. In *Information and Financial Engineering (ICIFE), 2010 2nd IEEE International Conference on*, pages 603–607. IEEE, 2010.
- [7] G. Hohpe and B. Woolf. *Enterprise integration patterns: Designing, building, and deploying messaging solutions*. Addison-Wesley Professional, 2004.
- [8] W. Kim, N. Ballou, J. F. Garza, and D. Woelk. A distributed object-oriented database system supporting shared and private databases. *ACM Transactions on Information Systems (TOIS)*, 9(1):31–51, 1991.
- [9] M. Lenzerini. Data integration: A theoretical perspective. In *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 233–246. ACM, 2002.
- [10] M. Magnani and D. Montesi. Bpdmn: A conservative extension of bpmn with enhanced data representation capabilities. *arXiv preprint arXiv:0907.1978*, 2009.
- [11] A. Meyer, L. Pufahl, D. Fahland, and M. Weske. *Business Process Management: 11th International Conference, BPM 2013, Beijing, China, August 26-30, 2013. Proceedings*, chapter Modeling and Enacting Complex Data Dependencies in Business Processes, pages 171–186. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [12] OMG. Business Process Model and Notation (BPMN) version 2.0.2. <http://www.omg.org/spec/BPMN/2.0.2/>, january 2014.
- [13] N. Russell, A. H. Ter Hofstede, D. Edmond, and W. M. van der Aalst. Workflow data patterns. Technical report, QUT Technical report, FIT-TR-2004-01, Queensland University of Technology, Brisbane, 2004.
- [14] S. Sadiq, M. Orłowska, W. Sadiq, and C. Foulger. Data flow and validation in workflow modelling. In *Proceedings of the 15th Australasian database conference-Volume 27*, pages 207–214. Australian Computer Society, Inc., 2004.
- [15] Y. Sun, J. Su, B. Wu, and J. Yang. Modeling data for business processes. In *2014 IEEE 30th International Conference on Data Engineering*, pages 1048–1059, March 2014.
- [16] A. Thöni, L. Madlberger, and A. Schatten. Towards a data-integration approach for enterprise sustainability risk information systems. In *Proceedings of the 7th International Conference on Research and Practical Issues of Enterprise Information Systems, Linz*, 2013.
- [17] W. M. P. Van Der Aalst. Getting the data. In *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, pages 95–123. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [18] W3C. Web services activity statement. <https://www.w3.org/2002/ws/Activity>, 2002.
- [19] J. Waldo. Remote procedure calls and java remote method invocation. *IEEE Concurrency*, 6(3):5–7, Jul 1998.
- [20] S. A. White. *BPMN modeling and reference guide: understanding and using BPMN*. Future Strategies Inc., 2008.