

Ferramenta para anotação semântica de processos de negócio de uma redação jornalística

Marcelo Fonseca¹, Edison Ishikawa¹, Benedito M. Neto², Edgar Costa³, Marcio Victorino⁴

¹Dept. de Ciência da Computação, Universidade de Brasília(UnB) – DF – Brasil

²Faculdade de Comunicação, Universidade de Brasília(UnB) – DF – Brasil

³Dept. de Engenharia de Produção, Universidade de Brasília(UnB) – DF – Brasil

⁴Faculdade de Ciência da Informação, Universidade de Brasília(UnB) – DF – Brasil

Abstract. *This work implements a tool that assists the modeling of new business processes in a flexible newsroom, in which process could be created or modified on the fly. Thus we created a light domain ontology for this newsroom by using the methodology 101 and Protégé modeling tool. The BPMN standard was also used to implement the information system that would automatically suggest the most appropriate role to perform a given task in a news production process of a newsroom. In order To achieve this goal, a tool was developed to aid in the semantic annotation of the application domain of the processes.*

Resumo. *Este artigo propõe uma ferramenta que auxilia a modelagem de processos de negócio em uma redação jornalística flexível, onde os processos podem ser criados ou reconfigurados on-the-fly. Criamos assim uma ontologia de domínio leve de uma redação jornalística, usando a metodologia 101 e o editor Protégé. Em seguida usamos o padrão BPMN para implementar uma aplicação que sugere automaticamente o papel mais indicado para executar uma determinada tarefa do processo de produção de notícia da redação. Para alcançar este objetivo foi desenvolvida uma ferramenta para auxiliar na anotação semântica do domínio de aplicação dos processos.*

1. Introdução

Considerando a dinâmica do ambiente jornalístico, este artigo apresenta um artefato para permitir a edição de processos de negócios semanticamente descritos para apoiar o jornalista na criação e modificação de processos de negócios, agregando valor a essa atividade. Para tanto, definiram-se os papéis dos autores e o fluxo de trabalho da redação, usando uma ontologia leve de domínio para anotar semanticamente as atividades dos processos com a utilização da linguagem gráfica BPMN 2.0, que também tem seu diagrama nomeado de *Business Process Diagram* (BPD), apresentada na Figura 1. Nesta figura apresenta-se um exemplo de parte do processo de produção da notícia do Jornal Campus Multimídia na UnB, contexto para o qual direcionamos a solução aqui proposta.

2. Web Semântica e Ontologias

A Web Semântica oferece a capacidade de se criarem repositórios de dados na Web, construírem vocabulários e escreverem regras para interoperarem com esses dados. A conexão

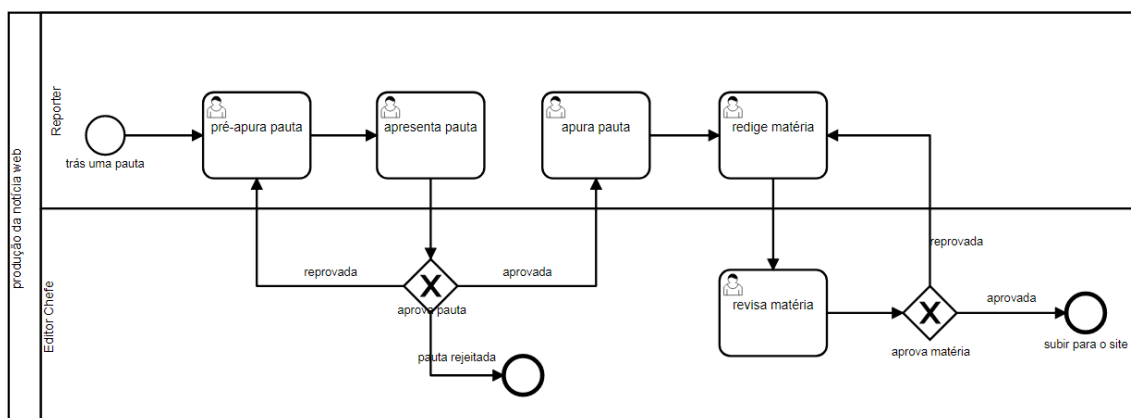


Figura 1. BPD de parte da produção da notícia no Campus Online/UnB

de dados é possível com tecnologias como RDF, SPARQL e OWL onde tanto pessoas como computadores sejam capazes de entender esses dados [Dairon 2012].

Para o conceito de ontologia nesta pesquisa, é utilizado o conceito de Gruber [Gruber 2004] e dos autores em [da Silva et al. 2014, Breitman 2005] que definem ontologia de domínio e ontologia leve.

3. Anotação semântica de processos

Para que uma máquina entenda o que um processo de negócio faz é preciso antes anotá-lo semanticamente fazendo o uso de ontologias. Existem diferentes formas de realizar a anotação semântica de processos mas com destaque para as baseadas em Ontologias e Petri Nets. Portanto neste estudo seguimos à abordagem com o uso de ontologia, semelhante aos trabalhos [Weber 2009, Bor et al. 2007, Kluza et al. 2015, Tobias Weller 2016]. Dessa forma, no artigo [Kluza et al. 2015] é proposto enriquecer *frameworks* de código aberto de execução de processo com semântica. O autor compara as tecnologias disponíveis de *workflow engine* **Camunda**, **jBPM** e **Activiti**. Todas utilizam a linguagem BPMN e são apresentadas como boas candidatas para adicionar anotação semântica aos processos.

4. Tecnologias utilizadas

O desenvolvimento do artefato utiliza a ferramenta de modelagem web bpmn-js utilizada nos trabalhos [Tobias Weller 2016, Kluza et al. 2015] e o framework **Ruby on Rails**(RoR). A linguagem de programação Ruby fornece suporte à bibliotecas semânticas o que viabiliza o desenvolvimento do artefato no framework RoR. Para integrar o RoR ao bpmn-js, utiliza-se neste estudo a biblioteca webpacker¹ que gerencia pacotes javascript no RoR e a biblioteca Nikogiri para a leitura de arquivo em formato XML produzidas pelo bpmn-js. O desenvolvimento da workflow engine foi realizada na aplicação com o uso de um banco de dados relacional.

5. Arquitetura e Implementação

O artefato possui as funcionalidades básicas de uma *workflow engine* e apresenta uma interface de comunicação no padrão Representational State Transfer (REST). Com isto,

¹<https://github.com/rails/webpacker>

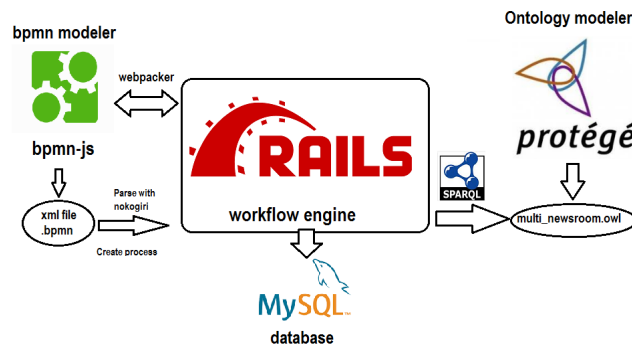


Figura 2. Arquitetura do artefato

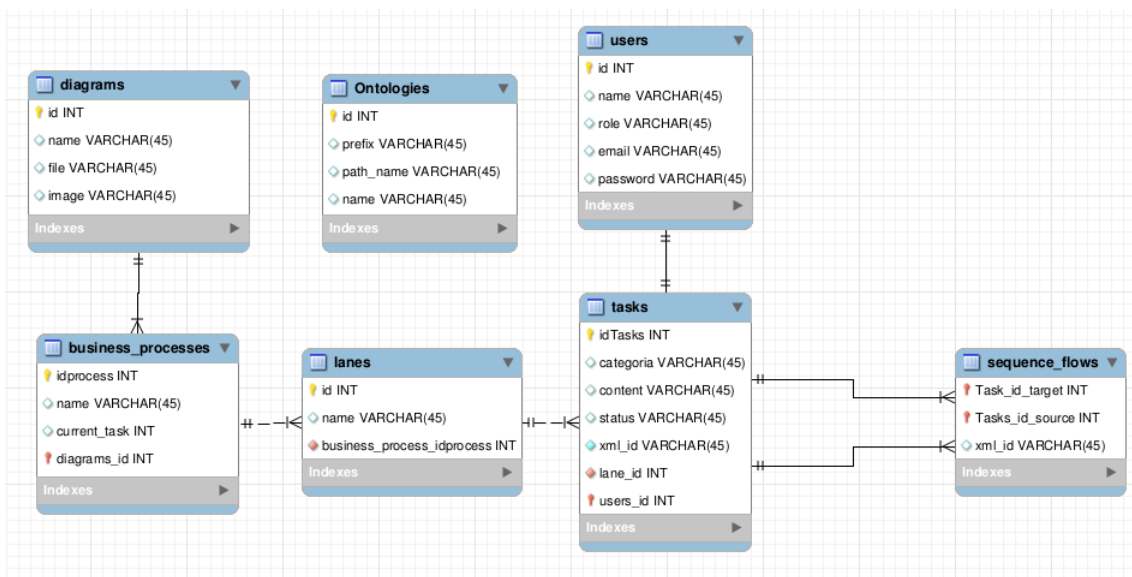


Figura 3. Modelagem no mysql workbench.

o artefato provê os dados de processos via *web service*. A Figura 2 representa a ontologia modelada no *Protégé*(*multi_newsroom.owl*) e as ferramentas de desenvolvimento utilizadas.

Para a modelagem dos dados levamos em consideração o *workflow engine*, a armazenagem do BPD e a ontologia OWL. O *workflow engine* exige uma modelagem das entidades de forma a poder simular de maneira genérica qualquer processo. Podemos nos beneficiar do modelo relacional para simular a transição pelos seus estados de um processo e instanciar usuários para as tarefas. Desta forma, modelamos no MySQL Workbench o banco de dados na Figura 3, descritas a seguir.

diagrams representa a entidade de um BPD. Um diagrama é composto de seu arquivo XML e sua imagem em *.svg*, como a Figura 1. A entidade do diagrama armazena os dois arquivos e tem uma relação 1 pra N com processos, desta forma vários processos podem atuar sobre um diagrama.

business_processes representa a entidade relacional de um processo genérico. Um processo tem um identificador para acessar a tarefa em execução, *current_task*, para fácil acesso em qual tarefa do processo se encontra. O processo é composto de varias

raias(lanes), o que permite mais de um papel atuando no processo.

lanes representa uma raia na linguagem BPMN. Uma raia é composta do nome da raia e um identificador para o seu processo. Uma raia é composta de várias tarefas através de *lane_id*.

tasks representa o nodo da linguagem gráfica BPMN. Uma *task* contém uma enumeração, *category*, para identificar qual o tipo do nodo. *Content* é referente ao conteúdo escrito na tarefa (ex: Apurar pauta), *status* classifica o estado da tarefa como feita, em andamento ou não inicializada. O *users_id* identifica para quem aquela tarefa é designada.

sequence_flows representa as setas de direcionamento da linguagem gráfica BPMN. Assim, identificamos a tarefa anterior e a tarefa seguinte. Uma tarefa tem uma relação 1 pra N com um direcionamento, o que permite direcionamento para N tarefas seguintes e anteriores presentes no diagrama.

users representa os usuários da organização. Um usuário é constituído de nome, email e *role*. *Role* identifica qual o papel do usuário na aplicação(ex:reporter). A relação 1 para N com tarefas permite um usuário a ser instanciado a mais de uma tarefa em diferentes processos.

ontologies representa o armazenamento dos arquivos OWL. Para este trabalho apenas a ontologia da redação jornalística desenvolvida foi utilizada. O *path_name* representa o local de armazenamento do arquivo, seguidos de seu nome e prefixo para acesso.

5.1. Modelagem de processos bpmn

A interface bpmn-js permite modelar e salvar diagramas de acordo com o padrão BPMN. Para complementar, foi desenvolvida a criação automática de um processo de acordo com algum BPD modelado pela interface de modelagem. Para tal foi desenvolvido um *parser* para receber o arquivo XML de um diagrama e com a leitura da estrutura, é criado um processo no MySQL dinamicamente.

Para tornar a aplicação efetivamente uma *workflow engine*, é implementado as seguintes funcionalidades: Um processo deve ser inteiramente exibido e onde ele se encontra, ter início e fim, uma tarefa pode instanciar um usuário para sua execução e um usuário instanciado visualiza as suas tarefas;

5.2. Modelagem da ontologia de domínio

Após a determinação dos papéis dentro da redação jornalística Campus Multimídia, elaboramos a modelagem da ontologia no *Protégé* de acordo com os papéis e tarefas no processo de produção da notícia na Figura 1. A criação da ontologia se dá pela: hierarquia de classes, hierarquia de objetos(predicados), atributos da classe e indivíduos das classes.

A hierarquia de classes da ontologia se encontra de acordo com a Figura 4. As classes que não se associam são modeladas disjuntas (ex:pessoa e material). A Figura 5 representa a modelagem dos objetos de acordo com as tarefas presentes no processo.

Para atender a busca semântica de forma mais generalizada, modelamos os indivíduos **reporter**, instância da classe Reporter e **editor**, instância da classe Editor-Chefe.

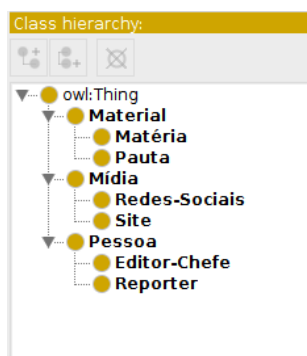


Figura 4. Modelagem das classes no protégé

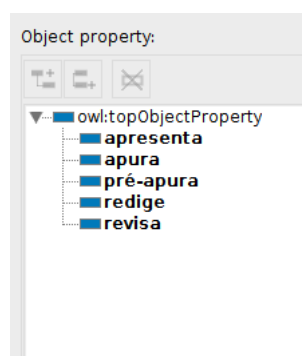


Figura 5. Modelagem das propriedades de objeto no protégé

56	http://www.semanticweb.org/2017/multi-newsroom#reporter	http://www.semanticweb.org/2017/multi-newsroom#apresenta	http://www.semanticweb.org/2017/multi-newsroom#pauta
57	http://www.semanticweb.org/2017/multi-newsroom#reporter	http://www.semanticweb.org/2017/multi-newsroom#apura	http://www.semanticweb.org/2017/multi-newsroom#pauta
58	http://www.semanticweb.org/2017/multi-newsroom#reporter	http://www.semanticweb.org/2017/multi-newsroom#pre-apura	http://www.semanticweb.org/2017/multi-newsroom#pauta
59	http://www.semanticweb.org/2017/multi-newsroom#reporter	http://www.semanticweb.org/2017/multi-newsroom#redige	http://www.semanticweb.org/2017/multi-newsroom#matéria

Figura 6. As triplas RDF no W3C RDF Validation Service

O mesmo se segue para matéria e pauta. Com essa modelagem construímos triplas rdf mais genéricas para atender nosso objetivo. O arquivo OWL gerado é incorporado ao artefato.

Para cada tarefa presente no processo em análise, fizemos uma busca com o predicado e objeto da tarefa. O processo segue o padrão mostrado na Figura 7 para a organização das triplas RDF para encontrar o sujeito da tripla.

5.3. Busca SPARQL

A busca SPARQL retorna o sujeito mais indicado de a tarefa. Ela é realizada em cada tarefa de um processo para em seguida fazer a sugestão automática de papéis para cada uma delas. No validador RDF do W3C², podemos verificar as triplas RDF formadas com os indivíduos criados na Figura 6. No artefato isso funciona através de um formulário é recebido o nome do BPD e submetida para um que módulo faz a leitura do BPD e retorna uma lista das tarefas existentes no diagrama. Em seguida, para cada elemento da lista é feita a busca SPARQL que retorna as sugestões dos papéis exibidas em tela como mostra a Figura 8.

6. Conclusões

O desenvolvimento do artefato demonstra a capacidade de usar a ontologia para realizar a anotação semântica dos processos de negócios. Os resultados estão disponíveis³ e podem ser incrementados em trabalhos futuros. Os objetivos de implementação foram alcançados, no entanto o artefato desenvolvido pode não apresentar uma interface trivial para a utilização por um profissional na área de jornalismo, mesmo tendo a aplicação a capacidade de abstrair grande parte das tarefas.

²<https://www.w3.org/RDF/Validator/>

³<https://github.com/marcelobbfonseca/workflow-api>

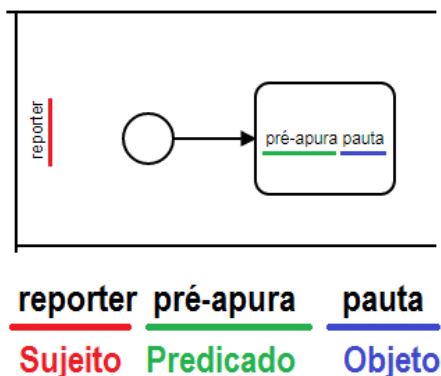


Figura 7. Tripla RDF no processo



Figura 8. Captura de tela da aplicação: Sugestões de papéis da consulta SPARQL.

Como trabalhos futuros pretendem-se implementar um tutor online que baseado no conhecimento das anotações semânticas dos processos de negócios permitam auxiliar na configuração e usabilidade do artefato, diminuindo assim as dificuldades para quem não tiver familiaridade com a solução. O funcionamento, teste e demonstração do artefato, no escopo deste trabalho, demonstra a sua viabilidade. A avaliação do mesmo pode ser conduzida em um ambiente experimental ou em um contexto real. Para isso haverá a necessidade de interação entre o pesquisador e a organização na qual o artefato está sendo instanciado.

Referências

- Bor, M., Dorr, F., and Weber, I. (2007). User-friendly semantic annotation in business process modeling. *Lecture Notes in Computer Science, vol 4832*.
- Breitman, K. K. (2005). *WEB SEMÂNTICA A INTERNET DO FUTURO*. LTC.
- da Silva, A. R. L., Borges, M. A., Fernandes, M. C. P., Sartori, V., and Spanhol, F. J. (2014). Ontologia como representação do conhecimento: aplicação no curso de formação continuada em tecnologias educacionais na web. *RENOTE - Revista Novas Tecnologias na Educação ISSN 1679-1916*.
- Dairon, J. (2012). Introdução a web semântica. *DEVMEDIA*.
- Gruber, T. (2004). What is an ontology? www.ksl.stanford.edu/ksl/what-is-an-ontology.html. acessado em 22/11/2017.
- Kluza, K., Kaczor, K., Nalepa, G. J., and Slazynski, M. (2015). Opportunities for business process semantization in open-source process execution environments. *2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 1307–1314.
- Tobias Weller, M. M. (2016). Adaptive semantic process modeling tool. *SEMANTiCS (Posters, Demos, SuCCESS)*.
- Weber, M. B. B. M. P. (2009). Auto-completion for executable business process models. *Lecture Notes in Business Information Processing, vol 17*. BPM 2008. Springer, Berlin, Heidelberg.