

SALÃO DE  
INICIAÇÃO CIENTÍFICA  
**XXIX SIC**  
  
**UFRGS**  
PROPESQ



múltipla   
**UNIVERSIDADE**  
inovadora  inspiradora

<b>Evento</b>	Salão UFRGS 2017: SIC - XXIX SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2017
<b>Local</b>	Campus do Vale
<b>Título</b>	Identificação de Elementos de Processo a partir de Textos em Linguagem Natural
<b>Autor</b>	RUBENS IDERON DOS SANTOS
<b>Orientador</b>	LUCINÉIA HELOISA THOM

# Identificação de Elementos de Processo a partir de Textos em Linguagem Natural

## Universidade Federal do Rio Grande do Sul

Rubens Ideron dos Santos - Lucinéia Thom (Orientadora)

Gerenciamento de Processos de Negócio (*Business Process Management* - BPM) é uma disciplina que permite aprimorar a forma como o trabalho é realizado nas organizações. Isto significa que as práticas BPM, quando aplicadas corretamente, proporcionam redução de custos, erros e falhas na execução dos processos, aumentando a eficiência da organização. O analista de processos, profissional responsável por aplicar práticas BPM nas organizações, realiza a gestão dos processos de negócio segundo o ciclo de vida BPM. Dentre as tarefas executadas neste ciclo destaca-se a modelagem de processos, pois, além de impactar diretamente nas etapas subsequentes do ciclo, ela é complexa e demorada, ocupando em média 60% do tempo total de um projeto em BPM<sup>1</sup>.

A modelagem de processos consiste em descrever processos tal como são executados na organização em uma notação específica para este uso, a Notação e Modelo de Processos de Negócio (*Business Process Model and Notation* - BPMN). O analista de processo realiza a modelagem de forma manual, coletando informações através de *workshops*, entrevistas com pessoas-chave do processo e análise de documentos<sup>2</sup>. Gerentes de conteúdo consideram que cerca de 85% da informação armazenada nas organizações é desestruturada, destacando-se os documentos textuais em linguagem natural<sup>3</sup> (ex.: *e-mails*, relatórios e contratos). Tais textos possuem características que dificultam a modelagem, como ambiguidades, descontinuidades na descrição do processo e sentenças fora de contexto.

Esta pesquisa propõe uma abordagem para minimizar o esforço do analista na tarefa de modelagem. A abordagem consiste em identificar elementos de processo em textos previamente à modelagem manual utilizando técnicas de processamento de linguagem natural (*Natural Language Processing* - NLP). Para isto, foram deduzidas regras de mapeamento entre construções sintáticas e elementos de processo a partir de textos descritivos de processo. Como esta técnica analisa sentenças individuais, é esperado que tenha desempenho bom para uma grande variedade de textos.

A validação das regras propostas incluiu a implementação de um protótipo desta abordagem através de técnicas de NLP. O protótipo foi alimentado com sentenças que descrevem elementos de processo retiradas de textos variados. As mesmas sentenças foram categorizadas nas mesmas classes por dois analistas segundo suas próprias experiências de modelagem. Foi classificado um total de 417 sentenças, extraídas de 70 textos retirados de 6 fontes distintas. A comparação dos resultados do protótipo com a modelagem manual proveu resultados promissores (em média: 73.61% de precisão, 70.15% de revocação e 71.82% da medida-F).

---

<sup>1</sup> HERBST, J. An inductive approach to the acquisition and adaptation of workflow models. In: Proceedings of the IJCAI'99 Workshop on Intelligent Workflow and Process Management: The New Frontier for AI in Business. [S.l.: s.n.], 1999. p. 52–57.

<sup>2</sup> DUMAS, M. et al. Fundamentals of Business Process Management. [S.l.]: Springer, 2013. I-XXVII 1-399 p. ISBN 978-3-642-33142-8.

<sup>3</sup> BLUMBERG, R.; ATRE., S. The problem with unstructured data. In: . [S.l.]: DM Review, 2003.