

## Solicitação de Prorrogação de Prazo de Defesa

### Relatório de Atividades Desenvolvidas

Aprovado como aluno regular em agosto de 2015, concluí as disciplinas exigidas para obtenção dos créditos e fui qualificado em agosto do ano seguinte. Segue as principais atividades desenvolvidas:

**Revisão bibliográfica:** Foi iniciada uma revisão bibliográfica sobre as principais técnicas de mineração textual como extração automática de tópicos e extração de conhecimento em documentos com vistas ao desenvolvimento do sistema proposto.

**Segmentação e rotulação manual:** Foram colhidas atas das reuniões do departamento de pós-graduação da UFSCar-Sorocaba<sup>1</sup>. Os documentos foram oferecidos à profissionais que participam de reuniões desse departamento. Para isso, desenvolveu-se um *software* onde os voluntários segmentaram e rotularam o texto. O que resultou em um conjunto de referências para avaliação das ferramentas propostas.

**Ferramentas:** Como parte do sistema proposto, desenvolveu-se uma ferramenta que permite a segmentação automática de documentos permitindo a configuração das técnicas e seus parâmetros. A ferramenta oferece uma interface gráfica para visualização do segmentos extraídos bem como as *features* selecionadas na etapa de pré-processamento.

**Sistema proposto:** Iniciou-se o desenvolvimento de um sistema que é capaz de receber um conjunto de atas, segmentá-las em trechos um assunto relativamente independente e fornecer, por meio de técnicas de extração tópicos, as palavras que podem representar o assunto de cada trecho.

**Submissão de Artigo ao ENIAC:** Como parte da revisão bibliográfica e desenvolvimento do sistema proposto, a segmentação automática das atas resultou em um artigo submetido ao ENIAC – Encontro Nacional de Inteligência Artificial e Computacional. No artigo são descritos as principais técnicas empregadas nessa tarefa, além de ajustes no pré-processamento e parâmetros.

**Redação da dissertação:** A dissertação contém os primeiros textos introdutórios referentes à mineração de textos, segmentação textos, extração de tópicos e resultados obtidos com as ferramentas já desenvolvidas.

---

<sup>1</sup>[http://www.ppgccs.net/?page\\_id=1150](http://www.ppgccs.net/?page_id=1150)

## Solicitação de Prorrogação de Prazo de Defesa

### Plano de atividades e cronograma

Para a conclusão dos trabalhos e defesa, são necessárias as seguintes tarefas:

- 1 - Revisão bibliográfica:** A literatura referente a extração de tópicos deve ser estudada com maior profundidade a fim de aperfeiçoar a dissertação e obter referências de técnicas que compõem o estado-da-arte nesse assunto.
- 2 - Implementação do sistema:** O módulo de extração de tópicos necessita de otimizações quanto quanta a eficiência. O módulo de pesquisa necessita de uma interface com o usuário. O sistema após concluído deve ser avaliado comparando-o com a rotulação manual fornecida pelos voluntários.
- 3 - Conclusão da dissertação:** Os resultados obtidos com a versão final do sistema devem ser incorporados ao texto da dissertação bem como uma revisão da literatura.

Tabela 1: Cronograma das atividades

Atividades	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
1	X	X	X	X		
2	X	X	X	X		
3	X	X	X	X	X	X