

Ovídio José Francisco

**Aplicação de técnicas de Mineração de Textos
para Organização e Extração de Históricos de
Decisões de Documentos de Reuniões**

Sorocaba, SP

13 de setembro de 2017

Ovídio José Francisco

Aplicação de técnicas de Mineração de Textos para Organização e Extração de Históricos de Decisões de Documentos de Reuniões

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação (PPGCC-So) da Universidade Federal de São Carlos como parte dos requisitos exigidos para a obtenção do título de Mestre em Ciência da Computação. Linha de pesquisa: Aprendizado de Máquina.

Universidade Federal de São Carlos – UFSCar

Centro de Ciências em Gestão e Tecnologia – CCGT

Programa de Pós-Graduação em Ciência da Computação – PPGCC-So

Orientador: Prof.^a Dr. Katti Faceli

Coorientador: Prof. Dr. Rafael Geraldeli Rossi

Sorocaba, SP

13 de setembro de 2017

«Sobrenome», «Nome do aluno»

Aplicação de técnicas de Mineração de Textos para Organização e Extração de
Históricos de Decisões de Documentos de Reuniões/ Ovídio José Francisco. – 2017.

39 f. : 30 cm.

Dissertação (Mestrado) – Universidade Federal de São Carlos – UFSCar

Centro de Ciências em Gestão e Tecnologia – CCGT

Programa de Pós-Graduação em Ciência da Computação – PPGCC-So.

Orientador: Prof.^a Dr. Katti Faceli

Banca examinadora: Prof.^a Dr. Katti Faceli, Prof. Dr. «Membro 1», Prof. Dr.

«Membro 2»

Bibliografia

1. «Palavra-chave1». 2. «Palavra-chave2». 3. «Palavra-chave3». I. Orientador.
II. Universidade Federal de São Carlos. III. Título

**«COLOQUE AQUI A CÓPIA DA
FOLHA DE APROVAÇÃO ASSINADA»**

*Coloque a sua dedicatória aqui;
Coloque a sua dedicatória aqui.*

Agradecimentos

Agradeço,

ao xxxxxxxx por yyyyyyy.

ao xxxxxxxx por yyyyyyy.

*“Coloque aqui a sua epígrafe;
Coloque aqui a sua epígrafe; Coloque aqui a sua epígrafe;
Coloque aqui a sua epígrafe; Coloque aqui a sua epígrafe;
Coloque aqui a sua epígrafe; Coloque aqui a sua epígrafe.
(Fonte, Autor)*

Resumo

Dissertação (do latim, *disertatio*), é uma modalidade de redação ou tipo de composição, escrita em prosa sobre um tema que se devem apresentar e discutir argumentos, provas, exemplos etc. Nos meios universitários equivale a tese diferenciando-se pelo volume de material, a dissertação seria o material que envolvesse poucas páginas até o limite de cem, enquanto a tese rotularia os textos que ultrapassassem esse número; e pelo aspecto qualitativo, a dissertação pressupõe a capacidade de aplicação de um método de análise e interpretação, enquanto a tese implica a originalidade do tema ou da abordagem à luz do qual é exposta e discutida (fonte: wikipedia).

Palavras-chaves: Dissertação. Texto. Mestrado.

Abstract

A thesis or dissertation is a document submitted in support of candidature for an academic degree or professional qualification presenting the author's research and findings. In some contexts, the word "thesis" or a cognate is used for part of a bachelor's or master's course, while "dissertation" is normally applied to a doctorate, while in other contexts, the reverse is true. Dissertations and theses may be considered as grey literature. The required complexity and/or quality of research of a thesis or dissertation can vary by country, university and/or program, therefore, the required minimum study period may vary highly significant in duration. The word dissertation can at times be used to describe a treatise without relation to obtaining an academic degree. The term thesis is also used to refer to the general claim of an essay or similar work (Source: wikipedia).

Key-words: Dissertation. Text. Master.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

ABNT	Associação Brasileira de Normas Técnicas
abnTeX	ABsurdas Normas para TeX

Lista de símbolos

Γ	Letra grega Gama
Λ	Lambda
ζ	Letra grega minúscula zeta
\in	Pertence

Sumário

	Prefácio	25
1	INTRODUÇÃO	27
2	PRIMEIRA CONTRIBUIÇÃO	29
3	PROPOSTA	31
	Conclusão	33
	Referências	35
	APÊNDICE A – TÍTULO	37
	APÊNDICE A – TÍTULO	39

Prefácio

Dissertação (do latim, *disertatio*), é uma modalidade de redação ou tipo de composição, escrita em prosa sobre um tema que se devem apresentar e discutir argumentos, provas, exemplos etc. Nos meios universitários equivale a tese diferenciando-se pelo volume de material, a dissertação seria o material que envolvesse poucas páginas até o limite de cem, enquanto a tese rotularia os textos que ultrapassassem esse número; e pelo aspecto qualitativo, a dissertação pressupõe a capacidade de aplicação de um método de análise e interpretação, enquanto a tese implica a originalidade do tema ou da abordagem à luz do qual é exposta e discutida (fonte: wikipedia).

1 Introdução

Reuniões são tarefas presentes em atividades corporativas, ambientes de gestão e organizações de um modo geral. Seu conteúdo é frequentemente registrado em texto na forma de atas para fins de documentação e consulta posterior.

As atas de reunião possuem características particulares, frequentemente apresentam um texto com poucas quebras de parágrafo e sem marcações de estrutura, como capítulos, seções ou quaisquer indicações sobre o tema do texto. Além disso, o estilo formal contribui para a escrita de textos com poucos detalhes, pois o ambiente dá preferência a textos sucintos.

Devido a fatores como a não estruturação e volume dos textos, a localização de um assunto em uma coleção de atas é uma tarefa custosa, especialmente considerando o seu crescimento em uma instituição. Usualmente, são feitas buscas manuais guiadas pela memória ou com uso de ferramentas computacionais baseadas em localização de palavras-chave. Normalmente esse tipo de busca exige a inserção de termos exatos e apresentam ao usuário um documento com as palavras buscadas em destaque, mantendo o texto longo, o que dificulta por exemplo o ranqueamento por relevância.

Uma vez que a ata registra a sucessão de assuntos discutidos na reunião, há interesse em um sistema que aponte trechos de uma ata que tratam de um assunto específico. Tal sistema tem duas principais tarefas: 1) Descobrir quando há uma mudança de assunto. 2) Descobrir quais são esses assuntos. Este trabalho tem como foco principal, a detecção de mudança de assuntos, que pode ser atendida pela segmentação automática de textos (??????). E a extração automática de tópicos.

A tarefa de segmentação textual consiste em dividir um texto em partes que contenham um significado relativamente independente. Em outras palavras, é identificar as posições nas quais há uma mudança significativa de assuntos.

A segmentação de textos é útil em aplicações que trabalham com textos sem indicações de quebras de assunto, ou seja, não apresentam seções ou capítulos, como transcrições automáticas de áudio, vídeos e grandes documentos que contêm vários assuntos como atas de reunião e notícias.

Pode ser usada para melhorar o acesso a informação solicitada por meio de uma consulta, onde é possível oferecer porções menores de texto mais relevantes ao invés de exibir um documento grande que pode conter informações menos pertinentes. A navegação pelo documento pode ser aprimorada, em especial na utilização por usuários com deficiência visual, os quais utilizam sintetizadores de texto como ferramenta de acessibilidade (??).

Além disso, encontrar pontos onde o texto muda de assunto, pode ser útil como etapa de pré-processamento em aplicações voltadas ao entendimento do texto, principalmente em textos longos.

2 Primeira contribuição

Coloque o texto aqui.

3 Proposta

Proposta

Conclusão

Coloque o texto aqui.

Publicações

Submissões

Referências

APÊNDICE A – Título

Texto ou documento elaborado pelo autor, a fim de complementar sua dissertação/argumentação.

ANEXO A – Título

Texto ou documento não elaborado pelo autor, que serve de fundamentação, comprovação e ilustração.