



Aplicação de técnicas de PLN para a qualidade de software



Miriam Sayão
Carlos A. Prolo
FACIN-PUCRS

VIVA ESSE MUNDO



Qualidade de Software e Requisitos

- Correção de erros originados na etapa de requisitos pode custar até 100 vezes mais a correção desse mesmo erro ainda na etapa inicial
- Atividades de qualidade devem ser realizadas durante todo o processo de desenvolvimento
- A complexidade das atividades de qualidade aumenta com fatores tais como ambientes distribuídos de desenvolvimento e quantidade de funcionalidades a atender



VIVA ESSE MUNDO



Verificação e Validação em Requisitos

- *verificação:*
 - revisão cuidadosa de um artefato, buscando a localização de erros ou defeitos no mesmo
- *validação:*
 - avaliação do documento de requisitos visando assegurar que os requisitos relacionados no documento de requisitos correspondem ao esperado pelos usuários e clientes





Verificação e Validação em Requisitos

- diferentes visões na validação e verificação do documento de requisitos
- diferentes metas e intencionalidades por parte dos *stakeholders*

Como apoiar essas atividades?



VIVA ESSE MUNDO



Requisitos de Software

- Requisitos de software são freqüentemente expressos em linguagem natural
- Diferentes modelos para registro dos requisitos: sentenças estruturadas de requisitos, casos de uso, texto livre, *user stories*
- Documento de requisitos: estrutura definida pelo engenheiro de software, atendendo a interesses dos desenvolvedores.



VIVA ESSE MUNDO



Requisitos de Software e Visões

- Visões: grupos de requisitos relacionados entre si ou a uma determinada característica do sistema
- Visões permitem trabalhar com conjuntos menores de requisitos relacionados e facilitam atividades de verificação e validação de requisitos (detecção de conflitos, incompletude, inconsistências e imprecisão)



VIVA ESSE MUNDO



Requisitos de Software e Visões

- em ambientes distribuídos de desenvolvimento, visões possibilitam a distribuição do trabalho a participantes geograficamente distantes
- um tema ou índice pode ser associado a uma visão
- sistema de indexação para um documento de requisitos: um conjunto de visões com seus respectivos índices



VIVA ESSE MUNDO



Requisitos de Software e Visões

Desejamos prover aos *stakeholders*:

- habilidade de recuperar informações relacionadas a interesses específicos, como por exemplo *segurança* ou *passagens aéreas*
 - produzir um sistema de indexação
 - extraindo termos para utilização como índices
- OU
- agrupando requisitos e buscando um termo que represente o agrupamento





Requisitos de Software e Visões

- não há um único conjunto de agrupamentos de requisitos, ou um único conjunto de temas que melhor caracterizem o documento
- para diferentes *stakeholders* diferentes organizações podem ser apropriadas
- a técnica para construção das visões deve possibilitar a criação de diferentes indexações para o documento





Requisitos de Software e Visões

- técnicas conhecidas como categorização ou classificação de textos não são aplicáveis – cada documento de requisitos possui um conjunto único de temas para indexação
- no contexto de desenvolvimento de software, termos possuem significado muito diversos do usual: segurança, por exemplo, pode estar associada a bloqueio de acesso a usuários não autorizados



VIVA ESSE MUNDO



Requisitos de Software e Visões

- não podemos utilizar aprendizado supervisionado: pois não é possível treinar um classificador a partir de conjuntos de documentos previamente classificados por especialistas humanos
- algoritmos de clusterização costumam gerar agrupamentos disjuntos, porém necessitamos de agrupamentos não disjuntos



VIVA ESSE MUNDO



Requisitos de Software e Visões

- Premissas para agrupamentos de requisitos:
 - número reduzido de agrupamentos: de 3 a 9
 - agrupamentos de requisitos a partir de um conjunto de temas
 - requisitos podem estar associados a vários agrupamentos
 - temas podem ser extraídos do documento de requisitos ou gerados a partir de um interesse específico de um participante





Experimentos para agrupar requisitos

- documentos de requisitos na área do turismo
- 169 requisitos, estruturados em funcionais e não-funcionais
- 4 requisitos não funcionais



VIVA ESSE MUNDO



Experimentos para agrupar requisitos

- Experimento 1: selecionamos um conjunto de termos do documento de requisitos (frequência absoluta) e para cada termo recuperamos os requisitos que continham o termo
- Experimento 2: selecionamos um conjunto de termos do documento de requisitos com uso de medidas estatísticas como frequência absoluta, χ^2 e LogLikelihood





Experimentos para agrupar requisitos

- Experimento 3: uso de algoritmos de clusterização. Requisitos convertidos para uma representação *bag-of-words*, com uso de stemização e *stoplist* (*k-means*, na implementação *Weka*)
- Experimento 4: uso de uma taxonomia estruturada em dois níveis. Primeiro nível composto de termos selecionados pelas medidas frequência absoluta, χ^2 e LogLikelihood e o segundo nível via colocações identificadas por medidas como score T e informação mútua



Experimentos: alguns resultados

Indexing Term	Precision	Recall	$F_{\beta=1}$
pacote	0.42	0.31	0.36
pagamento	0.88	1.00	0.94
viagem	0.98	1.00	0.63
reserva	0.56	0.63	0.59

Table 2. Evaluation of Experiment 1

Indexing Term	Precision	Recall	$F_{\beta=1}$
pacote	0.55	0.26	0.35
pagamento	0.71	0.69	0.69
viagem	0.80	0.66	0.72
reserva	0.62	0.38	0.47

Table 3. Evaluation of Experiment 4



Avaliação dos resultados.....

- problema de solução difícil
- técnicas de aprendizado supervisionado para classificação não são aplicáveis
- não são aplicáveis algoritmos de clusterização que gerem partições (agrupamentos disjuntos). Também não é simples identificar temas que possam ser associados aos clusters





.... e continuidade do trabalho

- uso de conceitos extraídos de forma semi-automática do documento de requisitos
- Conceito: conjunto de termos relacionados que definem ou representam uma idéia, opinião, pensamento ou objeto
- Exemplo: o conceito **pagamento** pode ter associados os termos *boleto bancário, cartão de crédito, transferência bancária, ordem de pagamento, débito em conta,...*





Outras estratégias

- Identificação de discrepâncias, erros e omissões em requisitos (similaridade entre requisitos, omissões em requisitos não funcionais)
- Geração semi-automática do léxico da aplicação ou mesmo do domínio da organização



VIVA ESSE MUNDO