

## COMUNICAÇÃO MULTILÍNGUE APOIADA POR TRADUÇÃO AUTOMÁTICA: ESTUDOS NA ÁREA DE DESENVOLVIMENTO DE SOFTWARE

Renata Vieira - Facin/PUCRS - Processamento de Língua Natural  
Bianca Pasqualini - PPGLet/UFRGS - Teorias Linguísticas do Léxico  
Aline Evers - PPGLet/UFRGS - Teorias Linguísticas do Léxico  
Lucas Welter Hilgert - Facin/PUCRS - Processamento de Língua Natural  
Maria José B. Finatto - PPGLet/UFRGS - Teorias Linguísticas do Léxico  
Aline Villavicencio - PPGC/UFRGS - Processamento de Língua Natural  
Fabiano Bruno Gonçalves - PPGLet/UFRGS - Teorias Linguísticas do Léxico  
Leonardo Zilio - PPGLet/UFRGS e PPGC/UFRGS - Teorias Linguísticas do Léxico

Estudos recentes apontam o Brasil como um potencial candidato a realizar projetos internacionais com grandes organizações de software. Um dos fatores que tem levado a isso é a sobreposição de fuso horário com os Estados Unidos e a Europa, o que permite a prática de desenvolvimento distribuído de software (DDS). Esse tipo de prática tem se conformado em um diferencial competitivo para as organizações. No entanto, a execução de tarefas colaborativas envolvendo participantes de diferentes países traz consigo uma série de dificuldades, destacando-se as dificuldades linguísticas. Este projeto tem como objetivos principais o estudo e a aplicação de métodos, técnicas e ferramentas da área de Processamento de Língua Natural (PLN) para a facilitação da inserção do Brasil no mercado global de desenvolvimento de software. Dentre as principais possíveis colaborações da área de PLN, destacam-se as máquinas de tradução automática, permitindo que usuários utilizem seu idioma nativo para a comunicação durante a execução de tarefas colaborativas. Os serviços de tradução de máquina estão longe da perfeição em relação à qualidade das traduções produzidas. Portanto, uma das propostas do presente projeto é o estudo dos principais problemas apresentados na comunicação apoiada por esses serviços e alternativas de melhorias no processo. Para tal, serão construídos corpú e vocabulários multilíngues especializados da área de Engenharia de Software e analisados registros de comunicação entre equipes com diferentes idiomas nativos. Nesta comunicação, relatamos a avaliação de serviços de tradução simultânea em equipes distribuídas usando como suporte vocabulários controlados. Tais vocabulários facilitam a identificação e o mapeamento de termos, sua organização hierárquica, desambiguação, além do alinhamento bilíngue. A partir da avaliação das traduções simultâneas, identificamos que problemas de tradução não recaem apenas sobre aspectos terminológicos da comunicação. Assim, fizemos um levantamento de estruturas discursivas conhecidas como *lexical bundles*, a fim de testar a utilidade da automatização de blocos discursivos recorrentes que facilitem o diálogo entre as equipes. Para gerar tais recursos, procedemos ao processamento de corpú da área em questão, usando técnicas de alinhamento lexical e sentencial. No futuro, pretende-se avaliar os recursos produzidos pelo projeto em novos experimentos com estudantes e profissionais da área de Engenharia de Software. Inicialmente serão considerados Brasil e Itália, com possibilidade de inclusão de outros países.

### Referências bibliográficas

- BIBER, D.; CONRAD, S., *Register, genre and style*. Cambridge: Cambridge, 2009.
- CALEFATO, F.; LANUBILE, F.; PRIKLADNICKI R. "A controlled experiment on the effects of machine translation in multilingual requirements meetings". In *Global Software Engineering (ICGSE)*, 2011 6th IEEE International Conference on, pp. 94 -102, ago. 2011.
- DAMIAN, D.; ZOWGHI, D., Requirements Engineering Challenges in Multi-Site Software Development Organizations, *Requirements Engineering Journal*, 8-3, 2003, pp. 149-160.
- LOPES, L., VIEIRA, R., FINATTO, M., MARTINS, D. Extracting compound terms from domain corpora. *Journal of the Brazilian Computer Society (Impresso)*. pp.1 - 13, 2010.
- ODGEN, W. A Task-Based Evaluation Method for Embedded Machine Translation in Instant Messaging Systems, in *Advanced Decision Architectures For The Warfighter: Foundations and Technology* (P. Mcdermott And L. Allender eds.), pp. 341-357, ago. 2009.