



II JORNADAS SASIG, Évora, 2-4 Novembro 2009

LGTE: Sistema aberto de Recuperação de Informação Textual, Geográfica e Temporal.

Jorge Machado⁽¹⁾, José Borbinha⁽²⁾ e Bruno Martins⁽³⁾, Gilberto Pedrosa⁽⁴⁾

⁽¹⁾ IST / INESC-ID; Rua Alves Redol 9, 1000-029 Lisboa; jorge.r.machado@ist.utl.pt

⁽²⁾ IST / INESC-ID; Rua Alves Redol 9, 1000-029 Lisboa; jlb@ist.utl.pt

⁽³⁾ IST / INESC-ID; Rua Alves Redol 9, 1000-029 Lisboa; bruno.g.martins@ist.utl.pt

⁽⁴⁾ IST / INESC-ID; Rua Alves Redol 9, 1000-029 Lisboa; gilberto.pedrosa@ist.utl.pt

RESUMO

Este artigo apresenta o LGTE¹ (Lucene Geo-Temporal Extensions), um sistema de Recuperação de Informação (RI) textual, geográfica e temporal que estende o sistema aberto Lucene², um motor para indexação de texto escrito em Java. O LGTE é o motor por trás do serviço de pesquisa³ do DIGMAP⁴. O LGTE permite indexar coleções de documentos em XML e inclui uma serie de utilitários para funcionalidades comuns dos motores de busca que podem ser facilmente verificadas numa DEMO⁵ que vem com o pacote. O LGTE inclui ainda um componente para criar experiências de avaliação de RI que usa o formato CLEF/TREC e que disponibiliza diferentes utilitários como *stemmers* linguísticos e de n-gramas, modelos de expansão de query, modelos de ranking geográfico [1], modelos de ranking textual tais como o Okapi BM25, o modelo de linguagem do sistema Lucene-LM⁶, o Vector Space Model do Lucene, e os modelos *divergence from randomness* do sistema Terrier⁷. Este artigo apresenta ainda uma serie de interfaces [2] demonstrativas das capacidades Geográfico-Temporais do LGTE que foram construídas a pensar na WEB 2.0 usando uma serie de *web-services* do projecto DIGMAP juntamente com motor do LGTE.

Palavras chave: *software livre, recuperação de informação, lucene, sistemas de informação geográficos*

REFERENCIAS

¹ <http://code.google.com/p/digmap/wiki/LuceneGeoTemporal>

² <http://lucene.apache.org>

³ <http://digmap2.ist.utl.pt:8080/mitra/>

⁴ <http://www.digmap.eu>

⁵ <http://digmap2.ist.utl.pt:8080/lgte/indexDigmap.jsp>

⁶ <http://ilps.science.uva.nl/resources/lm-lucene>

⁷ <http://ir.dcs.gla.ac.uk/terrier/>



II JORNADAS SASIG, Évora, 2-4 Novembro 2009

- [1] Jorge Machado, Bruno Martins e José Borbinha (2009), "*LGTE: Lucene Extensions for Geo-Temporal Information Retrieval*". GIW, ECIR, Toulouse.
- [2] Jorge Machado, Gilberto Pedrosa, Nuno Freire, Bruno Martins, Hugo Manguinhas, (2009), "User interface for a geo-temporal search service using DIGMAP components". ECDL, Corfu, Grecia.