SCIENCESPRINGDAY

FCT FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

Departamento de Informática

Search and Data Mining

CITI / Multimodal Systems

CHI CENTER FOR INFORMATICS AND INFORMATION TECHNOLOGIES

ZON MULTIMEDIA

RTP

sapo.pt Quidgest UT Austin | Portugal

João Magalhães

Prof. Auxiliar

Ongoing funded projects:

ImTV
media data

CS4SE
medical data

QSearch
generic search

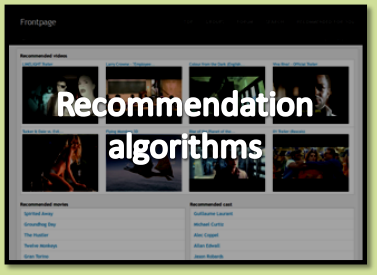
MEOrecommender
media data

Objectives

Current challenges in search and mining applications are deeply related to the heterogeneity of data and the size of data repositories.

We research new **machine learning and data storage approaches** capable of efficiently processing, accessing and locating large data repositories.

Our current topics of research are: image-concept mining and ranking; large-scale indexing and search with probabilistic methods; opinion mining for recommender systems; text mining for concept-based querying and searching.



Methodology

We are currently researching innovative methods to analyse and index large-scale data.

Machine learning and Information retrieval methods aim at bridging user queries and indexed information effectively and efficiently.

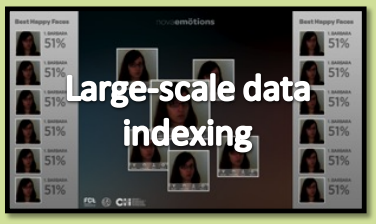
All methods are researched on a medium-size GPU cluster delivering over 22Tflops in single-precision and 5 Tflops in double-precision.

Expected Results

A principled and generic approach for processing heterogeneous data (e.g., medical records, media metadata, Web data) will be implemented as a free software framework.

Our goal is to explore human text (e.g., medical notes, user comments) and visual data (e.g., medical images, user videos).

Effective technology-transfer to our industrial partners.



Funding:

FCT Fundação para a Ciência e a Tecnologia

labs.sapo.pt

Quidgest

QREN QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL