

# Terrier



Computing  
Science  
GLASGOW



EPSRC

Engineering and Physical Sciences  
Research Council

## Terrier – Terabyte Retriever

The Terrier Information Retrieval (IR) platform is a modular framework for large-scale IR applications. It is based on **Divergence From Randomness** parameter-free probabilistic models for automatic document ranking and query expansion.

Terrier supports the indexing of various types of collections, including TREC and Web collections. It can also be used as a Web, intranet and desktop search engine.

Terrier offers a modular API for querying, allowing the development of applications as diverse as experimenting with standard test collections, or building Web search engines and XML applications.

## Divergence From Randomness

Terrier provides a number of IR models, based on the Divergence From Randomness (DFR) probabilistic framework.

The DFR framework is a generalisation of the 2-Poisson model and generates families of parameter-free IR models. It is based on a simple idea:

*"The more the divergence of a term's frequency in a document from its distribution in the collection, the more the amount of information carried by the term in the document"*



[terrier@dcs.gla.ac.uk](mailto:terrier@dcs.gla.ac.uk)

## Getting Terrier

- Terrier is distributed as open source software, under Mozilla Public License (MPL).
- Terrier is available for download from:

<http://ir.dcs.gla.ac.uk/terrier/>

## Features

- Open Source Software
- Cross-platform, written in **Java**
- Indexing various Web & TREC collections
- High **compression** of indices
- Processing **large-scale collections**, such as the Terabyte track .GOV2 collection
- Ideal for Internet search
- Supports centralised & distributed IR
- Modular architecture
- APIs for querying & indexing
- Provides DFR & various other IR models for easy cross comparison
- Easy addition of new IR models
- Effective automatic **query expansion**
- Proximity search & support for advanced query languages
- Standard evaluation of results

 <http://ir.dcs.gla.ac.uk/terrier/>