

# evaluations

## Table of contents

1 Evaluation Events.....	2
--------------------------	---

## 1. Evaluation Events

- [TREC \(2003\)](#)
  - Genomics Track Resource to use for relevance judgments is the Gene References into Function (GeneRIF) field of LocusLink.
  - Task: For gene X, find all MEDLINE references that focus on the basic biology of the gene or its protein products from the designated organism. Basic biology includes isolation, structure, genetics and function of genes/proteins in normal and disease states.
- [TREC \(2004\)](#) 4.5 million MedLine records, two tasks: *ad-hoc retrieval* , *catagorization*
  - Genomics Track: This task mimics a real-world problem faced by the MGI and other model organism databases: how to "triage" documents that are likely to have experimental evidence for GO code assignment. MGI is particularly interested in genes that have been assigned the "weakest" form of evidence: inferred from electronic annotation (IEA). IEA means that the GO code has been assigned because of some possible function inferred in another electronic resource, such as a sequence database. When this happens, and there also happens to be literature discovered about this gene, MGI is keenly interested in determining whether there is true experimental evidence (as opposed to electronic inference) to support the GO code assignment. When a gene has a GO code assigned with evidence level IEA and literature about the gene is known to exist, a two-step process is followed. The first step is a high-level review of each article to determine if it has experimental evidence that would warrant more detailed analysis (also called annotation or curation). The second step is to actually assign the specific GO codes. The task for the track will be to mimic the first step, i.e., determining if experimental evidence exists that would allow GO code assignment. This is, of course, a text categorization task.
- [BioCreative \(2003\)](#)
  - Task1: "entity identification" in text (genes, proteins, small molecules, chemicals, tissues, etc.)
  - Task 2: classification of entities (gene products) into classes based on Gene Ontology (Ashburner et al., 2000). (automatic functional annotation)
    - find the segment of text from the document which supports the annotation of the text with teh GO term
- [KDD Cup \(2002\)](#)
  - Task 1: 32 participants, 862 tagged papers for development, 213 papers for test, determine whether the paper meets the Flybase gene-expression curation criteria, and

*evaluations*

for each gene, indicate whether the full paper has experimental evidence for gene products (RNA and/or protein).