

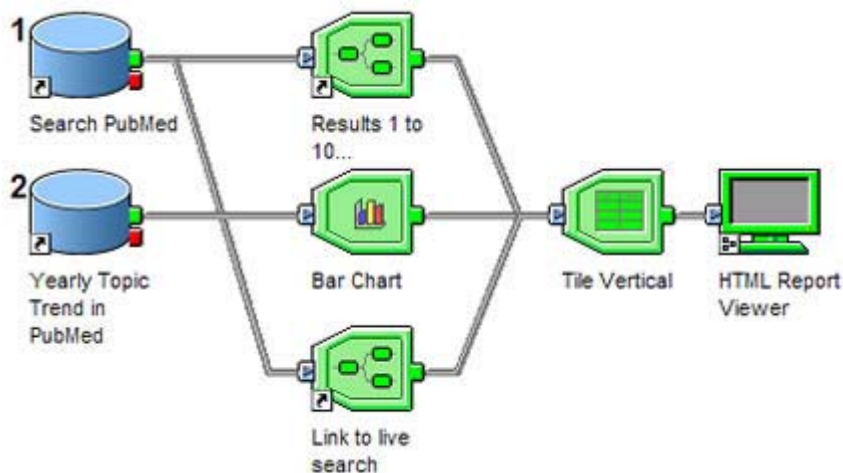


PubMed Trend Analysis using Pipeline Pilot

Keeping up to date with the explosion of published literature is virtually impossible these days. What are the hot topics and trends in science and industry? How are discoveries in one field of research impacting others, particularly yours? It's a catch-22 situation - you need to know what's going on in your field and related fields, but at the same time you need to be doing your own research. How can you most efficiently find the papers you need to read and keep abreast of developing fields of research? One way to help with this problem is to quickly and easily track trends and correlations in publications, for example to PubMed. These analyses can help reveal emerging topics and relationships in science, as well as those that are yesterday's news.

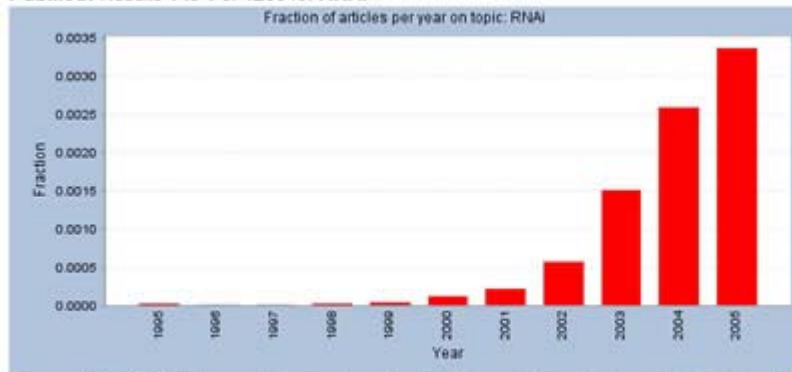
Trend Analysis

Trend and correlation analyses can be performed using the Text Analytics Collection (TAC) for Pipeline Pilot. TAC includes components dedicated for the search, retrieval, analysis and display of documents. To perform a trend analysis to PubMed, and show the results as a bar chart, set up a protocol similar to what is shown below:



The key here is the component labeled "Yearly Topic Trend in PubMed". This component takes a text query (i.e., the topic of interest) and a range of years as input and calculates the number of articles in PubMed matching that query each year. The component also calculates the total number of published articles for a given year, meaning that both raw counts for the topic of interest and the fraction of publications for that topic, compared with all publication, in a given year, can be computed.

The results can be visualized in a number of ways. The graphic below shows the results viewed in a bar chart, which is part of the Reporting Collection for Pipeline Pilot.

PubMed: Results 1 to 1 of 4260 for RNAi

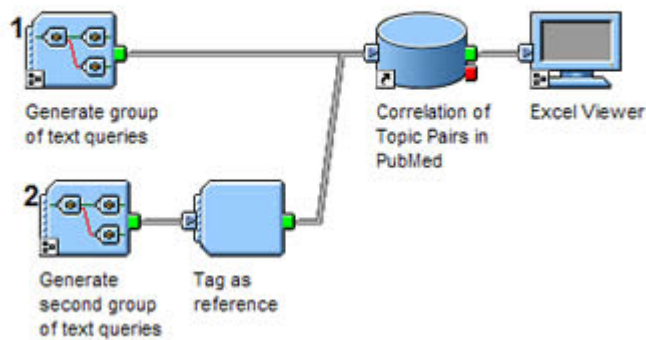
This search at PubMed: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=search&db=pubmed&sort=rank&term=RNAi&from=1>

The graphic shows the tremendous increase in papers published about RNA interference, a technique for controlling and investigating biological processes, over the last five to six years.

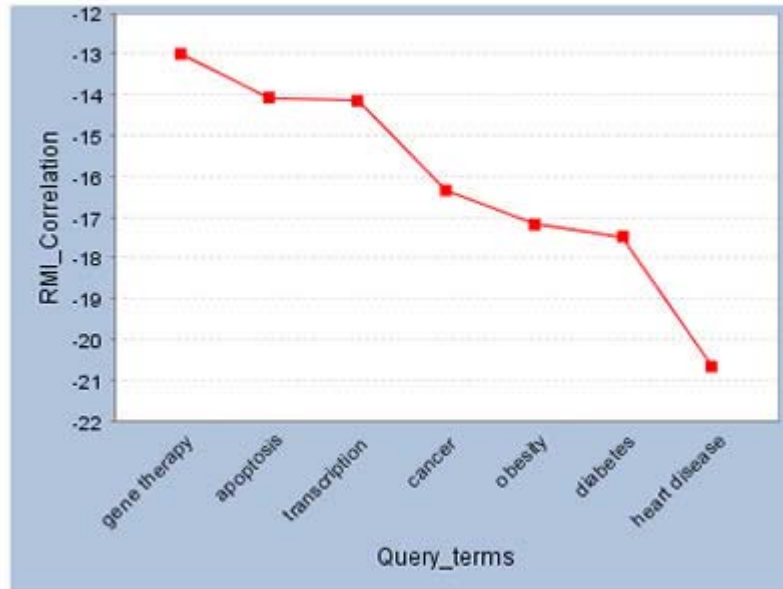
Topic Correlation

A valuable complement to characterizing publication trends for individual topics, is to compute the correlation between topics in the literature. A strong correlation occurs when a pair of topics is often mentioned in the same document, indicating a relationship between the topics. A weak correlation between topic pairs suggests that there is no relationship between them, or at least no recognized relationship.

This protocol demonstrates how terms from streaming data can be examined for relatedness.



Correlation analysis for RNAi



Extending Analyses

The Text Analytics Collection comes prepackaged with components for performing trend and correlation analyses, which can be very valuable. However you may want to extend these analyses further, for example by performing a trend-correlation computation to look for changing correlations over time. With this type of analysis you can ask questions such as "what pairs of topics are becoming hot?" and survey the changing landscape of science and industry. This can point out areas of opportunity or, conversely, areas of tight competition.

One of the great strengths of Pipeline Pilot is the ability it offers you to customize your analyses. Unlike many other applications that limit you to only the prepackaged analyses, Pipeline Pilot makes it straightforward to flexibly combine prepackaged components into analysis pipelines to achieve the desired analysis.

Copyright © 2001-2008 Accelrys Software Inc.

[Careers](#) | [Legal / Terms of Use](#) | [Contact us](#)