

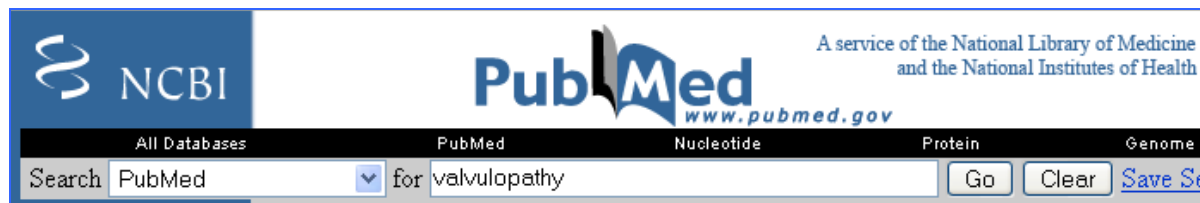


Importing the Results of a PubMed Search

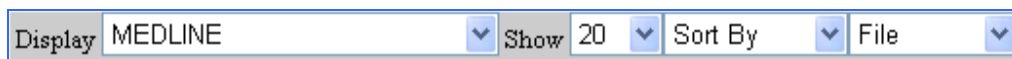
This Quick Start Guide explains how to import the results of a PubMed search into OmniViz and display the data in the OmniViz Galaxy View.

Step 1: Create a MEDLINE® Document Set in OmniViz-Ready Format


1. Open PubMed in your internet browser <http://www.ncbi.nlm.nih.gov/sites/entrez>.
2. Search PubMed for your term(s) of interest (e.g. valvulopathy) and click **GO**.



3. On the **Display** line, change **Summary** to **MEDLINE** and **Send to** to **File**.



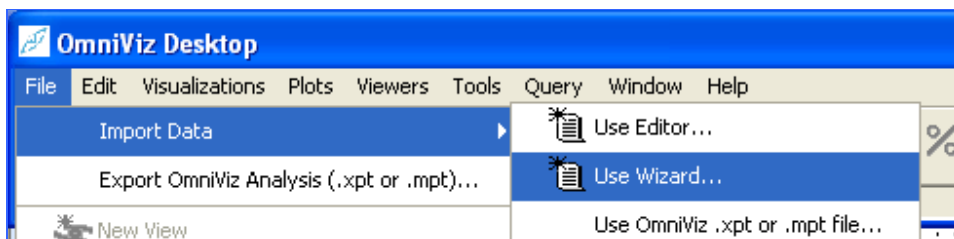
4. A **File Download** dialog box will appear. **Save** the file to a destination, with a filename of your choice. The file is now ready to be imported into OmniViz.

 **NOTE:** Before importing your MEDLINE query file into OmniViz:

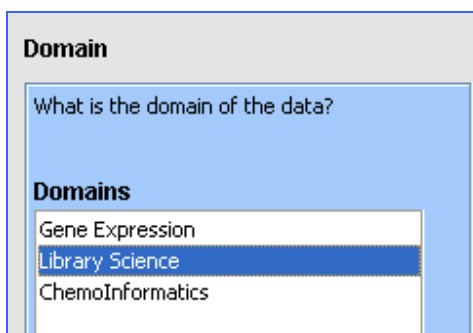
- It is worth knowing the time it takes for this process: on a Sony VAIO laptop with 512MB RAM & Pentium®M1500MHz processor, it takes ~10mins for 5K documents & ~1h for 30K documents.
- It is a good idea to open the MEDLINE query file to check that ALL of the records were downloaded. If not, you will need to repeat the downloading procedure.

Step 2: Import the MEDLINE Document Set into OmniViz

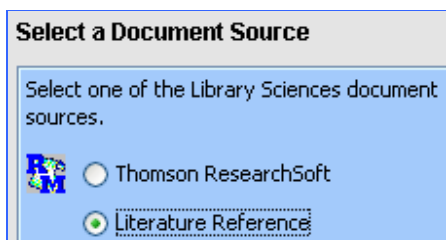
1. Open OmniViz. In the **OmniViz Desktop**, click on **File → Import Data → Use Wizard**. A series of dialogue boxes will follow.



2. For the **Domain**, choose **Library Science** and click **Next**.

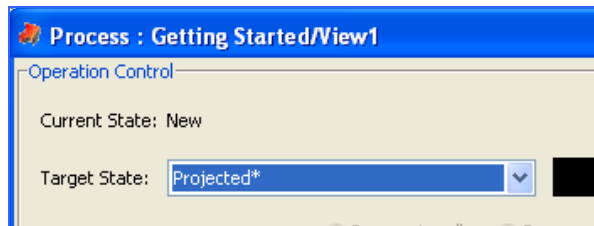


3. For the **Document Source**, select **Literature Reference** and click **Next**.

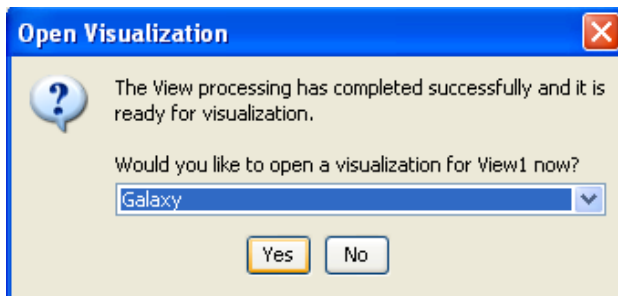


4. Click on **Add** and navigate to the MEDLINE query file that you have just created. Click **Open**, followed by **Next**. OmniViz will then start to import the documents.
5. Give your **Data Set** a name and click **Next**; your files will be loaded by OmniViz.
6. To view the data, click **Yes, create the view now** (the default setting) and click **Next**.

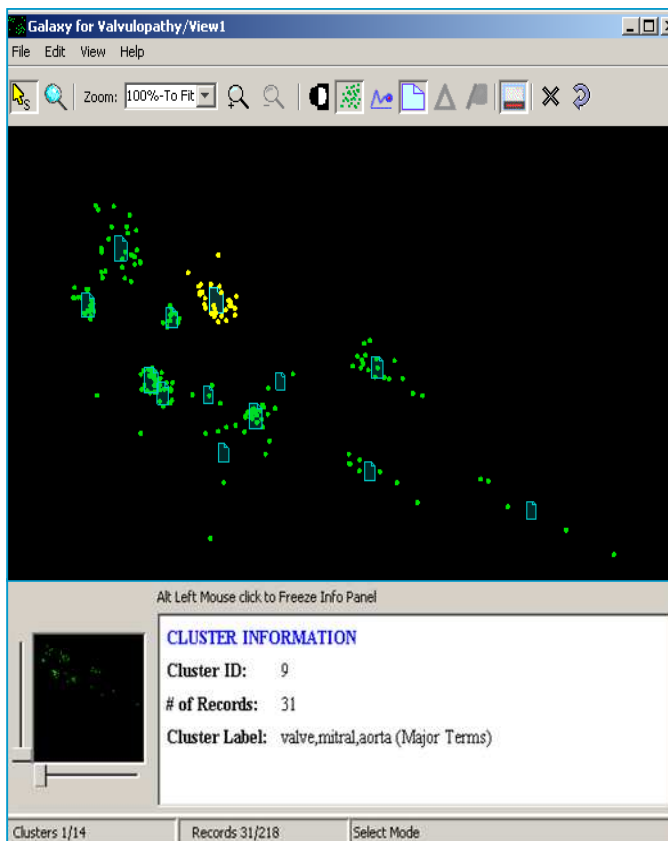
7. Type a name for the view (or leave as the default 'View1') and click **Finish**. OmniViz will then start to process the data.



8. When the processing is complete, you will be able to open a Visualization. Choose Galaxy (the default) and click on **Yes**.



The Galaxy view shows how every document is related to every other document. Documents are shown as points in clusters. Documents and clusters that are near each other will have similar document content, whereas documents and clusters that are far apart will have relatively disparate document content. Clustering is based on mathematical analysis of the word occurrence, distribution and associations to assign a numeric vector to each document.



**NOTE:**

The number of MEDLINE abstracts that are imported into OmniViz is likely to be lower than the number of abstracts retrieved in the PubMed search because, by default, OmniViz is set to ignore records that have titles but no abstract. This setting can be changed. However, it should be noted that including records with titles only will bias the clustering of the records.

By hovering over any cluster centroid (the paper icon), you will see the cluster ID, the number of documents in the cluster and the top 3 discriminating terms of the cluster in the Information Panel. Hovering over a document will display the title, as well as the ID of the document's cluster. Click on a document or cluster centroid to select it and display its contents in tools such as Record Viewer and Gist.

See other Quick Start Guides and Videos for more information about understanding the Galaxy and using Gist to discover themes in your document set.