Step 3 Deliverable:

A description of the ARFF data file including:

1. **An explanation of the correspondence between the 100 keywords and the attribute declaration part of the ARFF file (the lines beginning with @attribute).**

**Answer**:

The @ATTRIBUTE declaration serves to format the 100 keywords such that the program using the ARFF file can see that they are the search terms and that the NUMERIC declaration is the data type that the program returns.

1. **An explanation of the data rows (the portion after @data). For example, pick a tuple and explain what the 0’s and 1’s mean for the document that this tuple represents.**

**Answer**:

The 0’s and 1’s are generated because the attributes are declared with the NUMERIC data type. Since we are searching for hit on the websites from DMOZ, the numbers are representative of the Boolean condition being used when searching for the attributes. A ‘0’ indicates that the search term (keyword) was not found on that page, whereas a ‘1’ shows that it did appear on that site.