**Installation and Setup**

* Download and install Perl (<http://www.activestate.com/activeperl/downloads>)
* Open a command prompt and run the following commands to install the perl modules:
  + PPM install dmake
  + cpan App::cpanminus
  + cpanm Spreadsheet::WriteExcel
  + cpanm Text::Ngrams (may need to be forced –f Text::Ngrams)
  + cpanm FFI::Raw (may need to be forced –f FFI::Raw)
  + cpan IPC::Cmd

**GUI Interface**

1. To open the GUI, double-click on the **ASAP.jar** file from the **src** directory.
2. Select whether you want to run a query on a single file or run an experiment on multiple files.
3. Select the method (SCAP or Burrows)

* SCAP is a language-agnostic method that compares groups of tokens in a file(s)
  + Requires N which is the n-gram length
  + Optional L which is the maximum number of tokens to store
* Burrows fast, non-language-agnostic method
  + Requires N which is the n-gram length
  + Requires a Token file for the language (C++ and Java Token files are provided in the "res" folder)

Running a query

* Running a query checks the test file against the files in the training directory
* Training results are entered into the training output directory, this directory must be empty to run the query properly

Running experiment

* Choose which kind of experiment you want to run: Default split, K-fold, Leave one out
  + Each requires a directory containing the dataset to be tested
  + The default split requires a test directory and a training directory
* Enter the appropriate values and files for the method you want to use

1. After the values and files are selected click the Query/Experiment button to execute.
2. The results are displayed in the output window
3. A spreadsheet is also created and can be opened from the GUI

**Command-line Interface**

1. Open the command prompt.
2. Run the Asap.pl script with the proper command-line arguments (see below).
3. To run a single query, first run in training mode and then run the query.
4. To run an experiment, just run in experiment mode.

**Training:**

perl Asap.pl -train (-scap|-burrows) (-inputdir=<directory>) (-outputdir=<directory>) [-n=<natural number>] [-tokenfile=<file>]

-scap: Use SCAP method of authorship attribution

-burrows: Use Burrows method of authorship attribution

-inputdir=<directory>: Specifies location of training data

-outputdir=<directory>: Specifies where to store output files

-n=<natural number>: Specifies n-gram size (only required for Burrows method)

-tokenfile=<file>: Specifies the file containing the list of feature tokens (only required for Burrows method)

**Query:**

perl Asap.pl -query (-scap|-burrows) (-inputdir=<directory>) (-doc=<file>) [-report=<file>] [-n=<natural number>] [-L=<natural number>]

-scap: Use SCAP method of authorship attribution

-burrows: Use Burrows method of authorship attribution

-inputdir=<directory>: Specifies location of files created during training mode

-doc=<file>: Specifies query document whose author is to be attributed

-report=<file>: Specifies Excel file to store query results (optional)

-n=<natural number>: Specifies n-gram size (only required for SCAP method)

-L=<natural number>: Specifies profile length (only required for SCAP method)

**Experiment:**

perl Asap.pl -experiment (-scap|-burrows) [...]

-scap: Use SCAP method of authorship attribution

-burrows: Use Burrows method of authorship attribution

-testdir=<directory>: Specifies location of testing data (must be used with -trainingdir)

-trainingdir=<directory>: Specifies location of training data (must be used with -testdir)

-inputdir=<directory>: Specifies location of testing and training data (must be used with -k)

-k=<natural number>: Specifies k-fold size (must be used with -inputdir)

-n=<natural number>: Specifies n-gram size

-L=<natural number>: Specifies profile length (only required for SCAP method)

-tokenfile=<file>: Specifies the file containing the list of feature tokens (only required for Burrows method)