Step One: New User Registration

The first step is new user registration from the CAT home page.

1. Go to the website: http://cat.texifter.com.
2. Select "click here to register for a new account."

The following steps will take you through the **"New User Registration"** for the person who going to be the **primary account** holder. A primary account holder controls who has access to the data management, coding and analysis in the new user's account.

1. Create unique user name and password.
2. Provide required background information (last name, first name, and an e-mail address).
3. Click on the confirmation link you will receive in via email.
4. Acknowledge the privacy and security statement.
5. Your account will be authorized (this may take a few minutes or a few days).
6. You will receive an email that you are ready to log in using your username and password.

> ***Note:*** *As a first time user you will need to create coder sub-accounts. You can create unlimited number of sub-accounts; however this function may be limited in the future. This will be explained further in a later section.*

"Forget your Password?"
If you forget your password and want to reset it, select:

1. "Click Here" on the home page to the right of the username and password field.
2. Enter your username and email address
3. A new password will be generated and sent to you.

If you cannot remember your username or your email address is wrong, contact Dr. Stuart Shulman (stu@texifter.com) for assistance.

To Login

1. Go to the CAT Home Page http://cat.texifter.com.
2. In the Login fields type your USERNAME and PASSWORD.
3. Then select LOGIN.
4. You will be taken to the main page.
5. The username of the person currently logged in will appear in the left hand corner of the main page.

Step Two: Preparing a Raw Dataset for Coding

You can upload and code a dataset from a file with text formatted in one of three ways:

1. a plain text file (.txt)
2. a zip archive of plain text files (.zip)
3. an XML file (.xml).

Data preparation when using a single plain text file:
CAT relies on predefined spans of text to enable the auto-loading of discrete, **"codeable units"** during the coding process. Unless you use one of two special delimiters, the system assumes you want to apply codes to the entire document.

If you upload a single plain-text file, the coding system will present codeable units lying in between each pair of blank lines. The blank link is the delimiter in this case. When uploading a single .txt file as your raw data, prepare it as follow:

<**text to be coded**><hard return>
<hard return>
<**text to be coded**><hard return>
<hard return>
…

Data preparation when using a .zip archive of plain text files:
If you upload a collection of plain text files in a .zip archive, the system assumes that each document is a "codeable unit." You can, however, insert a special delimiter in your raw data (==--endcodeableunit--===) so that you can upload a zip archive and still code at the paragraph rather than whole document level. The delimiter has to be on a line all by itself - that is, you need to have:

<**text to be coded**><hard return>
==--endcodeableunit--===<hard return>
<**text to be coded**><hard return>
==--endcodeableunit--===

Note: *Be sure to save your file as "plain text" (.txt) file.*

If using an xml file (with an .xml extension); the system will verify that it conforms to the correct schema definition and process the file as such. This schema may be found at: http://cat.ucsur.pitt.edu/resources/codeupload.xsd

A sample XML document and tips on using the XML upload functionality can be found at: http://cat.ucsur.pitt.edu/resources/codeupload.xsd

Preparing the Code File (*optional, but recommended*)

If you do not upload a code file, you will have to opportunity to define the codes on the next screen as well as assign the keyboard keys associated with each code.

There are three parts to the code file, which also must be a plain text (.txt) file:

1. the code name (e.g., yes, borderline, no)
2. the definition (e.g., yes, this is new information)
3. the key stroke (e.g., 1, 2, or 3)

The definition is preceded and followed by a "pipe" or vertical line (|). This is found above the Enter key with the backslash. Press Shift and Backslash to make the "pipe".

A code file will look as follows:

No|No, no specific new information provided|1

Uploading New Raw Data File and Code File

Once your dataset and code file (optional) are prepared you may upload them for viewing and coding. Navigate to the Upload New Raw Dataset page. Below the instructions on that page, you will have several options:

Raw Dataset Name: If you do not give it a name it will take the default file name
Raw data file: Click Browse to select the file with your prepared raw data
Code File: Click Browse to select the file with your code list
Data Format Style: Choose Standard or Word-by-Word (most will use Standard)

Two other **extremely important** options on this raw data upload page.

- **Disable Verification** for user-defined and multiple coding for multiple coding projects Check this box to remove a step where coders review their selections and verify the choices are final. Check this box to enable confident coders to work rapidly through a high-volume coding projects.
- **Allow User-Defined Codes** is a feature that enables the creation of new codes during the coding process. Check this box if you want to give coders the freedom to generate new codes on the fly.

Select Upload.