STIX-D Use Case L1: Import ACE Lexicon

* **Objective:** Seed the stixd\_corpus.lexicon table with lexical entries from the ACE Common Lexicon ([Clex](https://github.com/Attempto/Clex)) or other ACE lexicon files.
* **Actor(s):** Administrator
* **Input:** A URI to the clex\_lexicon.pl file or other ACE lexicon file.
* **Output:** Displays summary information about the import or an error message.
* **Process:**
  1. The administrator provides the URI to a ACE lexicon file. Default URI is: “<https://github.com/Attempto/Clex/blob/master/clex_lexicon.pl>”
  2. The system opens a connection to the stixd\_corpus database on the localhost through the src/mysql\_repository.py module.
  3. For each line in the Clex file, the system performs the following steps:
     + Extracts the characters from the start of the line to the last character before the open parenthesis character (`(` ) and stores the string in the local word\_tag variable.
     + Extracts the characters after the open parenthesis character (`(` ) and before the first comma (`,`). If this resulting string of characters is enclosed in single quotes, the system drops the enclosing single quotes. The system then stores the string in the local word\_form variable.
     + Generates a SHA256 hash for the combination of the word\_tag and word\_form variables and stores it in a local tag\_form\_hash variable.
     + Checks if the tag\_form\_hash value exists in the stixd\_corpus.lexicon table by comparing it with values in the lexicon.tag\_form\_hash field.
       - If the tag\_form\_hash value exists, the system adds a new link between the existing lexicon.lex\_id and the source lexicon.obj\_id in the lexicon.obj\_lex\_jt junction table.
       - If the tag\_form\_hash value does not exist, the system creates a new entry in the lexicon table and imports word\_tag, word\_form, and tag\_form\_hash into the table.
     + Extracts the 2nd argument and imports it into lexicon.logical\_symbol.
     + If the Clex line includes a 3rd argument, extracts it into lexicon.third\_arg field. Otherwise, sets the third\_arg field to NULL in the stixd\_corpus.lexicon
  4. The system displays summary information about the extracted text(s) or an error message.