Scripts Documentation

1. Build folder directory: EPU -> (gmrb, rmrb)
2. Download Anaconda
   1. <https://www.anaconda.com/products/distribution>
   2. Install Python, R and Jupter Note Book
3. Find the system PATH
   1. System Paths are the directories that Python will search by default. If the code package is in a folder other than the system path, python will be unable to find it. We need to make sure the packages are saved in the system paths, otherwise, we need to add the package path to the system path
   2. PATH can be found by using: *import sys* and then *print(sys.path)*
   3. Folder can be added to the system path with:
      * *import sys*
      * *sys.path.append("/home/me/mypy")*
4. Transfer *rexepu.py* to the correct PATH
   1. When adding new keywords, two commands are needed for each term to add in the rexepu.py: defining a new term and compile the new term (see comments in the rexepu.py file).
5. Transfer the Renmin Daily and Guangming Daily raw text data
   1. Renmin Daily:
      1. *rmrb1946\_2006\_full.txt, rmrb 2017\_2021.txt*
      2. The raw text of 2007-2016 is downloading.
   2. Guangming Daily:
      1. *gmrb 2008\_2021.txt*
      2. Use the Library of Congress’s Guangming Daily database to search
         * Use *( )* for composition
         * **Use**
6. Transfer the web scraping scripts:
   * 1. For Renmin Daily and Guangming Daily
     2. *scrap gmrb to update.ipynb* and *scrap rmrb to update.ipynb*
     3. Only need to change **month and year to scrape** the new month**.**
7. Update the Mainland China EPU:
   1. *update index.py*
   2. Change the default directory