program\_flow

Program Flow  
  
vaeet.py  
tweets:  
  
  
parsetweets/csv\_db/tweetsK/tweets.csv  
  
parsetweets> py parse.py tweetsK  
  
db/tweetsK/\*.json  
  
vaeet> py vaeet tweetsK [N=3] # N is max synonym substitutions per sentence  
  
db\_/tweetsK/\*.json  
  
  
  
  
vaeet.py  
emails/phishing:  
(Let <name> = 'emails' or 'phishing')  
NOTE: emails may be created in w3c-email text format (processed by parsemail)  
or created directly in vaeet-json format ready for mail send.  
  
  
parsemail/w3c-emails\_db/<name>K/\*.txt # files are in w3c-email text format  
  
parsemail> py parse.py <name>K  
  
db/<name>K/\*.json  
  
vaeet> py vaeet <name>K [N=3] # N is max synonym substitutions per sentence  
  
db\_/<name>K/\*.json # json-format ready for mailing  
  
  
  
  
vaed.py  
emails:  
  
  
email/emailsK/\*.json  
  
email2corpus> py parse.py emailsK  
  
corpus/emailsK/\*.txt # document-text format expected by vaed.py  
  
src> py vaed.py emailsK  
  
corpusgen/emails/emailsK/\*.txt  
  
corpusgen2emailgen> py parse.py emailsK  
  
emailgen/emailsK/\*.json # json-format ready for mailing  
  
  
documents:  
  
corpus/corpusK/\*.txt # vaed docs format  
  
src> py vaed.py corpusK [p=1.0] # p is the probability of sentence replacement  
  
corpusgen/corpusK/\*.txt