

Original text:
Verbatim (original)

Cleaning
(only FF data)

Text: *Verbatim*

g2pE: Grapheme to
Phoneme & removing
inter-word punctuation

IPA

CLTF: Phoneme to
sound class mapping

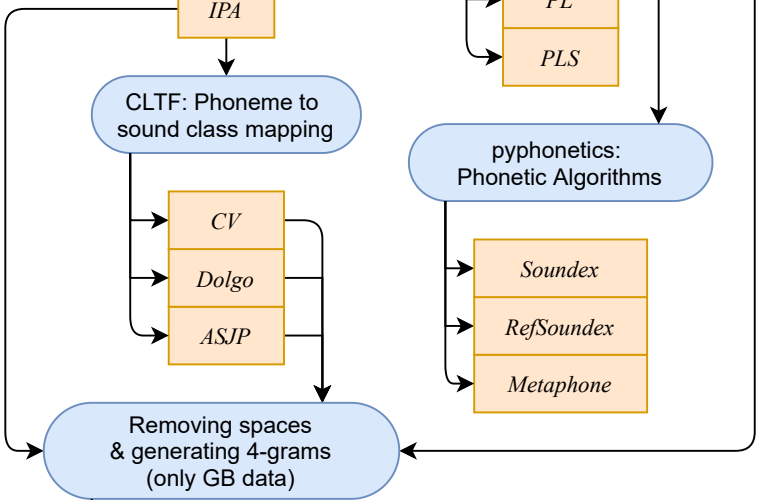
CV
Dolgo
ASJP

spaCy: Normalizing

P
PL
PLS

pyphonetics:
Phonetic Algorithms

Soundex
RefSoundex
Metaphone



IPA 4-grams
Dolgo 4-grams
ASJP 4-grams
CV 4-grams
P 4-grams