

Traditional MT “Pipeline”

“Training the Translation Model”

1. Find relevant parallel texts.
2. Align at the level of sentences.
3. Align at the level of words.
4. Extract translation units, with scores (co-oc. stats.).
(Language Model similar, “simple” words co-oc. stats, no alignment.)

“Tuning” (“MERT”) = Actual training in the ML sense

5. Identify TM/LM/other model component weights.

Translation:

6. Decompose input into known units.
7. Search for best combinations of units.