Report on Lexicon Development

Department of Information Technology is currently under the process of collecting words for lexicon development. Each word in the lexicon contains the following annotations:

- Word meaning (multiple)
- Part of Speech (POS)

We are also supposed to have the word pronunciation (IPA) but due to lack of know-how in phonetics (ex. IPA), the lexicon currently does not have the pronunciation annotation.

The Part of Speech annotation is currently in Dzongkha since the POS tag set that we submitted is yet to be reviewed by PAN.

The word count for the lexicon is given in the following table:

Sl. No.	Alphabet	Word Count
1	Ka	1576
2	Kha	1249
3	Ga	2228
4	Nga	648
5	Ca	435
6	Cha	598
7	Ja	323
8	Nya	225
9	Pa	1054
10	Pha	884
11	Ha	333
12	Aa	92
13	Sa	1250
	Total	**10895**