Algorithm 1 Word count mapper using the"in-mapper combining"

The mapper builds a histogram of all input documents processed before emitting key-value pairs for unique words observed.   [numbers=left, xleft-margin=7.5mm] class Mapper val counts = new Map() def map(key: Long, value: Text) = for (word <- tokenize(value)) counts(word) += 1   def cleanup() = for ((k, v) <- counts) emit(k, v)