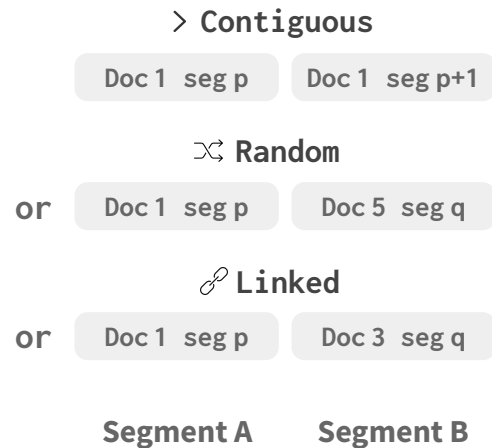
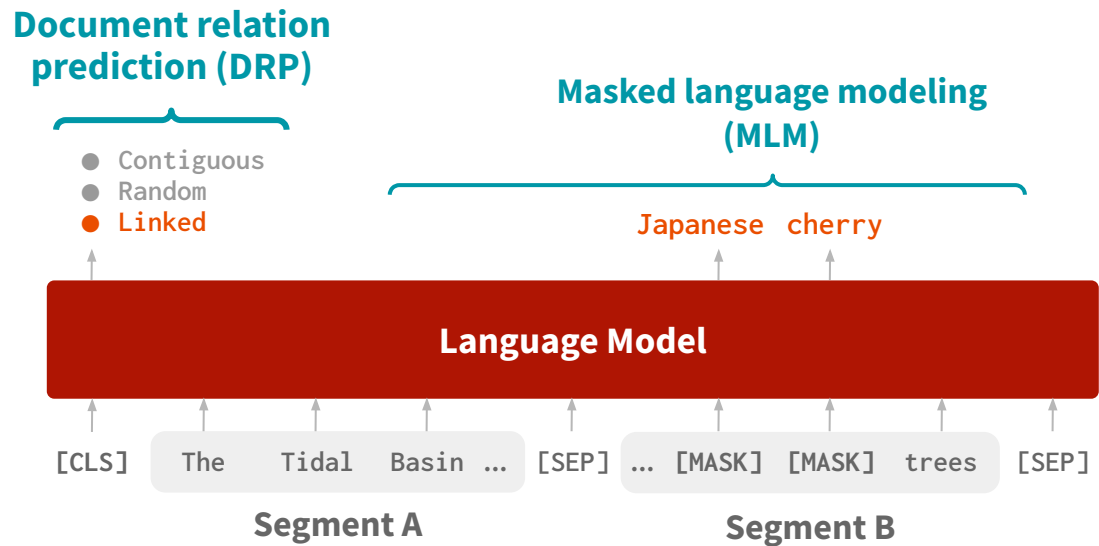


Corpus of linked documents



Create LM inputs



Pretrain the LM