

What has been done?

- Gensim
 - memory error on large wiki dataset (full)
 - small dataset working
 - Geometric operations working, but not always relevant due to training set
 - transferred the specifications to later + Questions

short term

- dropping the idea of training on full wikipedia
 - let run on 1 cpu core training for test
- train on larger pretrained datasets from Gensim
- Combine with ANN
- Document in report Gensim life
- Send mail to christophe for Github
- Compare with: FastText, word2Vec, Glove

overall progress

- Gensim is not working as expected on large dataset
- fall back on smaller datasets and officially pretrained

- Wikipedia FR
- Conversational agent framework

Question:

Deliverable next week