

# Aula 10/11 - Notebook - MultiDoc-QA

---

# Escopo

1. Gerar index com `LuceneSearcher`
2. Busca e re-rank CrossEncoder - `cross-encoder/ms-marco-MiniLM-L-6-v2`
3. Geração de resposta com `text-davinci-002`
  - a. Para o ground Truth e buscas
4. F1\_score baixo

# Resultados e tempos

- pyserini.index.lucene

- CPU times: user 986 ms, sys: 124 ms, total: 1.11 s  
Wall time: 1min 46s

- re-rank na faixa de 3 minutos

- Geração de respostas

- 50it [00:24, 2.04it/s]CPU times: user 554 ms, sys: 37.1 ms, total: 591 ms  
Wall time: 24.6 s

# Tempos e Resultados

	F1 score	EM
ground truth	0.308	0.26
Gerados a partir de buscas	0.296	0.26

# Observações

- Colab começou a travar com os notebooks grandes
- O doc2query poderia ajudar a melhorar os resultados! (?)