Recherche:

* tokenization: <https://www.youtube.com/watch?v=D8j1c4NJRfo>
* batch: <https://www.youtube.com/watch?v=T0VSn-XkjGM>, <https://www.youtube.com/watch?v=cjBn4ATxZoQ>
* decoder: https://www.youtube.com/watch?v=Qo\_0PGusJfU
* pytorch: <https://www.youtube.com/watch?v=ORMx45xqWkA>, <https://www.youtube.com/watch?v=r2JKV4_zlmM>
* transformers: <https://www.youtube.com/watch?v=SZorAJ4I-sA>

Versuche, ob existierende KIs unser Problem lösen:

* ChatGPT, BingChat und Co: ohne Probleme (noch weiter ausführen)
* GPT4All:  
  - Mistral: klappt gut, auch auf DE (noch weiter ausführen)  
  - Falcon: Klappt kaum, lädt lange und stürzt ab, hängt sich oft auf (noch weiter ausführen)

Modell-Versuche:

* Orca 2:  
  - lädt erst lange (20Min) und ist dann abgebrochen