# Verslag BAP 18/03/2024

## Besproken

* De 3 manieren om de BAP aan te pakken:
  + Van de grond op.
  + Finetuning van pretrained modellen.
  + Context window van trained modellen.
* Benchmarks
  + Taal benchmarks:
    - HellaSwag.
    - WinoGrande.
  + Taak benchmark:
    - 20 vragen:
      * 5 komen letterlijk uit de tekst.
      * 5 infereren uit tekst.
      * 5 niks te maken met tekst.
      * 5 het antwoord niet te vinden.
    - Opmerkingen:
      * In literatuur opzoeken en keuze beter onderbouwen  
        2 Papers (python codeer benchmark).
      * Testen voor foute suggesties.
* Modellen en implementatie:
  + Van de grond op:
    - Misschien niet haalbaar in de tijd.
    - Mogelijks hetzelfde als finetunen maar weights op 0 initialiseren.
  + Finetuning:
    - HuggingFace libraries en leaderboards gebruiken om keuze te maken.
    - Autotrain om te finetunen.
  + Context window:
    - Closed source modellen gebruiken.
* Domein
  + Taal:
    - Best Engels, gaat het makkelijkste zijn.
  + Keuze:
    - Misschien Cursustekst inlezen, finale test is dan een examen invullen.

## TODO tegen 25/03

* Keuze opensource en closed source model maken, deze heuristisch kunnen verdedigen aan de hand van benchmark cijfers.
* Open Llama is van nader bestuderen, misschien geeft dit inzichten voor de ground up aanpak.
* De benchmark beter ontwerpen a.d.h.v. literatuur.
* Nadenken over welk domein we willen testen.