## Onderzoeksvragen

* Hoe kunnen AI & ML worden geïmplementeerd binnen een Javaproject?
* Hoe kunnen AI & ML worden onderhouden na implementatie en productie?
  + Hoe zorg je voor minimale downtime van een product wanneer een algoritme her traint moet worden?
* Hoe kunnen AI & ML verantwoordelijk worden gebruikt?
  + Hoe breng je de uitkomsten en gebruik van data over naar de gebruiker?
  + Hoe verantwoord je data gebruik en data transformaties gemaakt door een algorithme?
* Wat zijn de gevolgen van AI-implementaties op de bruikbaarheid van de rest van de applicatie?
  + Moeten andere producten onafhankelijk werken van de AI?
  + Moeten services aangestuurd door AI menselijke input ontvangen?

# Hoe kunnen AI & ML worden onderhouden na implementatie en productie?

Meeste machine learning models volgen een standaard stappenplan voordat deze in een productie omgeving worden gezet. De standaard zes stappen beschreven door datacamp, zijn:

1. Planning
2. Data Preparation
3. Model Engineering
4. Model Evaluation
5. Model Deployment
6. Monitoring and Maintance

(Datacamp, 2022)

### Planning

Onder planning valt het oppakken van domeinkennis en het verbeteren van de probleemstelling. Hierbij kan gedacht worden aan vroege hypothese stelling zetten over wat een netwerk mogelijk zou terug kunnen geven. De toepasbaarheid van het probleem verbeteren, en het zoeken naar databronnen en de hoeveelheid van deze databronnen.

### Data Preparation

Data prep is het vergaren van data, deze labelen, opschonen en onderhouden. Onder onderhouden valt het verwijderen en toevoegen van nieuwe features aan de data, maar ook de opslag van data en hoe deze gebruikt kan worden het model.

### Model Engineering

Model engineering is het onderdeel waarbij keuzes worden gemaakt voor welke vormen van machine learning modellen er gebruikt gaan worden, welke metrics hiervoor gebruikt gaan worden, en het in elkaar zetten van dit model.

### Model Evaluation

Vanuit de vierde stap wordt bepaald of het model klaar is voor productie, ruim 90% van offline getrainde modellen worden niet naar productie gebracht en worden herstart vanuit de data preparation stadium. Naast het verifiëren van de correctheid van een model wordt hier ook gekeken naar regelgeving en ethische risico’s die dit model van het algoritme mogelijk ondergaat.

### Deployment

Vergelijkbaar met andere services, ML deployment kan gedaan worden op lokale services, cloud VMs of op package niveau naar andere programma’s. Vanuit hier kan het worden aangeroepen via API-calls of direct vanuit de browser als dashboard.

In dit stadium van het project worden er ook al vergelijking gemaakt met andere modellen rondom het gebruik van RAM-allocatie, opslag en processing power doormiddel van A/B testen.

### Monitoring & Maintenance

Modellen die uiteindelijk volledig uitgerold zijn worden automatisch onderhouden vergelijkbaar met standaard algoritmes in het logging systeem. Wanneer een model abnormaal gedraagt, errors oplevert of klachten van klanten ontvangt zal het development team deze automatisch ontvangen of deze worden gelogd in het systeem. Vanuit hier worden modellen niet vaak volledige opnieuw het lifecycle in gebracht, ingevallen van software of architectuur updates zal het model opnieuw naar het data preparation stadium worden gebracht.

Voor dit onderzoek zal de focus liggen rondom de laatste twee stappen in de lifecycle van een ML model. Er zal gekeken worden naar hoe verschillende modellen productie ingebracht kunnen worden. Hoe deze op verschillende niveau’s gehost kunnen worden, en hoe deze gemonitord kunnen worden op verschillende niveau’s.

## Model exports

<https://towardsdatascience.com/5-different-ways-to-save-your-machine-learning-model-b7996489d433>

## Continous training

<https://neptune.ai/blog/retraining-model-during-deployment-continuous-training-continuous-testing>