## Oriëntatie

Voor dit onderzoek gaan wij ons verdiepen in “Generative AI”. Hiervoor moeten we natuurlijk wat basisvragen opstellen, die zijn als volgt:

1. Wat is AI?

2. Wat is generative AI?

3. Welke andere soorten AI zijn er?

4. Hoe kunnen we verschillende soorten van generative AI onderscheiden van elkaar?

5. Welk effect heeft AI op ons?

6. Hoe werkt AI?

a.     Evt. in gaan op types

Om deze vragen te beantwoorden willen we gebruik maken van o.a. wetenschappelijke artikelen die we via Google Scholar kunnen vinden. Daarnaast willen we ook gebruik maken van verschillende bronnen van informatie die vrijgegeven worden.

## **­­­**Introductie

Kunstmatige intelligentie (AI) is intelligentie dat geuit wordt door machines, alleen kunnen machines niet voor zichzelf nadenken, deze intelligentie moet dus getraind worden. Een van de applicaties van AI is generative AI. Generative AI is het gebruik van kunstmatige intelligentie om originele media zoals tekst, afbeeldingen, video's en audio te genereren als reactie op input van een gebruiker [8]. Om de media te laten genereren wordt gebruik gemaakt van een Large Language Model (LLM).

Een LLM wordt getraind door middel van machine learning. Dit houdt in dat er een hele hoop data naar de computer wordt gestuurd, hier gebruikt de computer algoritmes om de manier waarop de mens leert te imiteren. Wat hieruit komt is een artificial neural network, oftewel een artificieel brein. [1]

Om generatieve AI beter te kunnen begrijpen is het handig om een basiskennis te hebben over de verschillende manieren waarop AI wordt toegepast. Zo zijn er traditionele AI, conversationele AI, generatieve AI en algemene AI (AGI). Traditionele AI is in principe een AI met tunnelvisie, zo heeft die hele specifieke instructies nodig om een taak te kunnen uitvoeren, zo kan je het bijvoorbeeld leren om te schaken. Conversationele AI is een vorm van AI waar de computer menselijke gesprekken kan simuleren. Voorbeelden hiervan zijn bijvoorbeeld chatbots, ‘text-to-speech’ services of ‘speech-recognition software’. Generatieve AI kan dus met een tekst input verschillende soorten media genereren, denk hier aan ChatGPT of Midjourney. Als laatste hebben we algemene AI (Artificial General Intelligence), dit lijkt heel erg op generative AI, alleen is het dan in alles beter. Er bestaan hier nog geen modellen voor, maar onder andere Google is hard bezig met onderzoek hier naar (Deep Mind). In theorie is AGI net een mens, zo kan het heel snel omschakelen en leren van vorige ervaringen. Een generative AI kan alleen maar doen waar het voor geprogrammeerd is, AGI zou daarentegen op zichzelf kunnen leren, redeneren en meer. Als een van deze AI’s onze banen zou kunnen overnemen, is eigenlijk alleen AGI een kandidaat. [2]

## Generative AI

Generatieve AI refereert dus naar een type AI dat nieuwe content/media kan genereren. Hierin kan onderscheid gemaakt worden in code, tekst, video, audio en nog veel meer. Er zijn veel manieren waarop de output gegenereerd kan worden. Zo heb je bijvoorbeeld General Adversarial Networks (GANs) en Transformer-based Models. Er zijn nog veel meer verschillende types maar aangezien dit de twee meest gebruikte types zijn gaan wij hier verder op in.

### General Adversarial Networks (GANs)

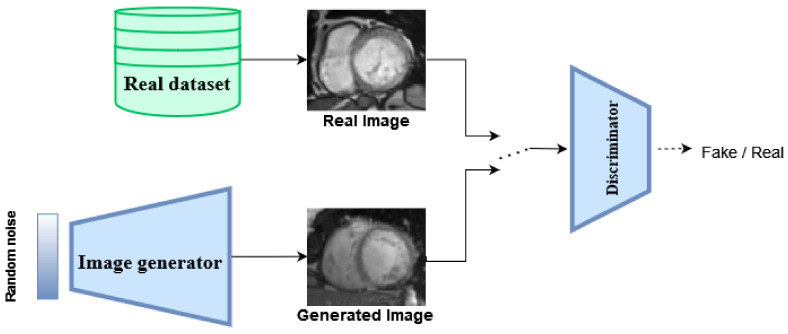
Het concept van GAN werd in 2014 geïntroduceerd door Ian Goodfellow, et al. GANs worden vaak gebruikt bij foto generatie. GANs bestaan uit twee neurale netwerken, een generator model en een discriminator model, die constant in strijd zijn met elkaar. [4]

Het discriminator model wordt getraind om te zien of een foto binnen zijn domein past of niet. Dit domein bestaat, bijvoorbeeld, uit een hele grote van bloemen foto’s. Daarnaast zijn is er ook een grote dataset van voorwerpen/dingen die helemaal geen bloemen zijn aanwezig in het domein. Zo weet de discriminator te onderscheiden tussen bloemen en niet-bloemen.

Nu vragen we aan het generator model, wat tot nu toe nog niks gedaan heeft, om foto’s van bloemen te genereren. Het model genereert dit en stuurt ze door naar de discriminator, deze zegt of het echt een foto is of niet en dat resultaat wordt ook aan het generator model verteld. Nu begint het proces weer opnieuw, tot dat de generator zo’n goede nep-foto heeft gemaakt dat de discriminator niet kan onderscheiden of het een nep-foto is of niet. [5]

Door dat constant afgewezen worden door het discriminator model moet het generator model zich constant aanpassen waardoor het steeds blijft leren. Elke keer als het generator model een nep-foto langs het discriminator model krijgt leert ook het discriminator model ervan aangezien dat wordt toegepast in zijn domein. Hierdoor worden keer op keer de gegenereerde foto’s beter en beter.

Midjourney is een bekende applicatie die gebruik maakt van GANs, hier wordt een tekst-prompt gegeven aan de applicatie en dan door middel van GANs wordt er een output foto gecreëerd. Daarnaast worden GANs ook toegepast in de medische wereld. Zo is er in 2023 een studie gepubliceerd waarin gebruik werd gemaakt van GANs voor het creëren van trainingsdata voor artsen in opleiding.



Een flowchart van GANs architectuur [6]

GANs kunnen ook gebruikt worden voor video-manipulatie. Een video is namelijk niets meer dan heel veel foto’s heel snel achter elkaar.

GANs kunnen gebruikt worden om meer dan alleen afbeeldingen te genereren, maar het is daar niet heel goed in en ook niet heel erg efficiënt. Zo zijn er veel andere types generatieve AI die daar veel meer voor geschikt zijn. Dus op het gebied van baanverlies voor een software engineer, hoeft niemand bang te zijn voor GANs. Het is wel een heel handig hulpmiddel voor software engineers, zo kan het bijvoorbeeld helpen met het maken van een mockup voor een front-end applicatie.

Een andere bekende toepassing van GANs zijn deepfakes. Deepfakes hoeven niet met GANs gemaakt te worden, maar het is wel een van de beste modellen hiervoor. Volgens de website mediawijs.be is een deepfake “een audiovisuele versie van nepnieuws. Een deepfake-filmpje vertrekt van bestaande beelden, die vervolgens bewerkt of gemanipuleerd worden via kunstmatige intelligentie.” [7]. De laatste tijd zijn deepfakes vaak in het nieuws geweest omdat het makkelijk is om nepnieuws te creëren. Door middel van deepfakes kunnen mensen dus allemaal nep informatie verspreiden over specifieke mensen, bedrijven of overheidsinstanties.

Aangezien GANs grotendeels gebruikt worden voor foto/video manipulatie, en deze technologie nog niet goed genoeg is om een gemanipuleerde foto zo goed te laten lijken dat dit verwarring kan brengen, hoeft de Haagse Hogeschool niet al te veel tijd in het bestrijden hiervan te stoppen.

### Recurrent Neural Network (RNN)

Een RNN is een model dat vooral goed is in het verwerken van sequentiële data om voorspellingen te maken. Het kijkt dus naar de data die hiervoor gegeven is om een voorspelling te maken. Dingen zoals audio en tekst zijn in principe sequentiële data. Audio is om te zetten naar sequentiële data door middel van een spectrogram, tekst is van nature al sequentieel, alle letters staan achter elkaar en woorden ook.

RNN’s kunnen dus gebruikt worden om patronen te vinden in een serie van datapunten. Een goed voorbeeld hiervoor zou de aandelenmarkt zijn. Hier zijn heel veel datapunten beschikbaar die een RNN zou kunnen analyseren en aan de hand van die datapunten kan een RNN patronen vinden en dus ook voorspellingen doen. [9]

Een RNN bestaat uit drie onderdelen, de input, de output en de hidden layer. De hidden layer zet de input om naar iets wat de output kan meegeven. Alle data dat uit de hidden layer komt wordt dus naar de output gestuurd, maar ook naar de hidden layer van de volgende node. Dit kunnen we kort uitleggen als volgt: de zin “Wie is daar?” kan worden opgesplitst in vier delen of nodes wie/is/daar/?. Elk onderdeel los zou geen informatie teruggeven, maar zodra we node 1 “Wie” doorgeven aan node 2 “is” krijgen we iets wat steeds meer op een logische zin gaat lijken. Nu gaan we de nieuwe node 2 “Wie is” doorgeven aan node 3 en die dan weer aan node 4. Hier zien we dat de input is opgesplitst in sequentiële delen, de vier nodes, die elke keer meegegeven worden aan het volgende deel. Vandaar de “recurrent” in Recurrent Neural Network. [12]

A diagram of a network

Description automatically generated

Een flowchart voor RNN architectuur [11]

Naast financiële voorspellingen doen, worden RNN’s ook gebruikt bij vertaal applicaties (Google Translate), tekstgeneratie (chatbots), transcriptie applicaties (Siri) en het kan zelfs gebruikt worden om bijvoorbeeld hartfalen te voorspellen [10].

Aangezien RNN’s moeilijk en langzaam zijn om te trainen zijn ze vooral goed in specifiek werk. RNN’s worden vooral gezien als een ‘stepping stone’ voor technologie als transformers of long short term memory. Bekende AI applicaties zoals OpenAI’s ChatGPT en Google’s Gemini maken gebruik van dit transformer model. YouTube en Netflix maken beiden gebruik van long short term memory (LSTM) om video’s/series/films aan hun gebruikers aan te bevelen. Zoals eerder gezegd zijn zowel transformers en LSTM in principe evoluties van RNN.

Momenteel zijn deze evoluties nog niet slim genoeg om ons zorgen te laten maken over baan-kansen die wij kunnen mislopen. Alles wat uit een van deze applicaties komt moet vaker wel dan niet gecontroleerd worden. Vooral als het om een blok code gaat. Aangezien we het nu hebben over het feit dat de evoluties van een RNN nog niet goed genoeg zijn om onze baan-kans te verminderen zal het bij een RNN, op het gebied van programmeren, nooit lukken. Dit kunnen we ook grotendeels zeggen over alle baankansen. Een RNN is in principe gewoon een hulpmiddel, het feit blijft dat dit voorspellingen creëert en deze voorspellingen kunnen niet altijd waar zijn. Aangezien een voorspelling alleen een voorspelling blijft.

Een RNN is een zeer handig model, en zeker ook iets wat gebruikt zou kunnen worden door de Haagse Hogeschool. Zo zou het mogelijk zijn om hoorcolleges te transcriberen voor mensen die slechthorend zijn of een andere beperking hebben. Daarnaast zouden er veel patronen/trends gevonden kunnen worden in het gegeven onderwijs. Denk hier aan vaak slecht scorende vakken die een nieuw jasje krijgen. Of juist als iedereen voor een vak juist hoog haalt, leren waarom dit zo is. Nu moeten wij toegeven dat een evolutie van een RNN hier waarschijnlijk meer geschikt voor zou zijn. Maar een RNN zou ook prima deze taak kunnen voldoen.

### AI op school

AI gaat een steeds groter onderdeel worden van onze samenleving, veel meer dan wat het nu is. Het is daarom juist belangrijk voor iedereen die er in aanraking mee komt er goed les over krijgt. Nu hoeft het natuurlijk niet voor iedereen in zoveel te gaan als wat wij hebben gedaan want daar gaat het merendeel toch niks van begrijpen. Maar het is wel belangrijk, vooral voor de technische opleidingen, dat er überhaupt tijd in wordt gestoken wat AI precies is en wat het allemaal inhoudt.

Het is nu al te zien dat bijna iedereen wel al in contact is gekomen met een vorm van AI, van een chatbot tot aan ChatGPT. Alleen beseffen meeste gebruikers niet hoe de technologie werkt en hebben ze er ook niet echt een behoefte aan om het te leren.

Wij vinden het dus belangrijk dat iedereen op de Haagse Hogeschool een introductie krijgt over de werkingen van AI en dan vooral oppervlakkig over de werkingen en gevaren van AI, maar ook juist de voordelen en hoe je het goed kan gebruiken. Daarnaast vinden we dus ook dat de technische opleidingen van de Haagse Hogeschool dieper op de werking van de verschillende types van AI moet in gaan. Hier zouden zij dan leren over de types, algoritmes en modellen die allemaal toegepast worden om uiteindelijk een product te krijgen als ChatGPT.

### Conclusie

Bijna alles wat we besproken en uitgelegd hebben in dit onderzoek was nieuwe informatie. We zijn hier van een redelijk brede term, kunstmatige intelligentie, heel snel ingezoomd op generatieve AI en daar 2 modellen van. Voor deze modellen, GAN en RNN, hebben we opgezocht en uitgelegd hoe deze werken en ook allemaal voorbeelden van gevonden.

### Reflectie

Pepijn:

Ik vond het een heel leuk en leerzaam proces. Ik vond het sowieso leuk dat we even iets anders kregen dan al dat datawarehouse gebeuren. Ik merkte tijdens het opzoeken van informatie dat ik erg geïnteresseerd was in het onderwerp maar wel alles wat te maken had met wiskunde niet goed begreep of zelfs helemaal niet begreep. Dat maakte het wel wat lastiger aangezien ik graag iets goed onder de knie wil hebben om het te kunnen uitleggen. Verder vond ik het een heel leerzaam proces.

# Bronnen:

[1] <https://www.cloudflare.com/learning/ai/what-is-large-language-model/>

[2] <https://aws.amazon.com/what-is/artificial-general-intelligence/>

[3] <https://www.makeuseof.com/what-is-artificial-general-intelligence-how-it-differs-generative-ai/>

[4] <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>

[5] <https://www.youtube.com/watch?v=TpMIssRdhco>

[6] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10055771/>

[7] <https://www.mediawijs.be/nl/artikels/wat-een-deepfake>

[8] <https://www.scribbr.nl/ai-tools-gebruiken/generative-ai-uitgelegd/>

[9] <https://www.youtube.com/watch?v=LHXXI4-IEns>

[10] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5391725/>

[11] <https://aditi-mittal.medium.com/understanding-rnn-and-lstm-f7cdf6dfc14e>

[12] <https://www.baeldung.com/cs/hidden-layers-neural-network>