



## Wat is ASTRID

ASTRID is een Cognitieve Architectuur. Dit betekent: een softwaresysteem dat in staat is tot redeneren. Het systeem kan werkelijk redeneren met gebruik van (complexe) kennis. Deze kennis bestaat uit een ingewikkeld model van de wereld dat beschrijft hoe alles werkt in deze wereld. ASTRID bouwt en onderhoudt dit wereldmodel helemaal zelf. ASTRID is op dit moment het enige systeem dat, voor zover bekend, hier toe in staat is.

## Wat is "Common sense Knowledge"

Mensen hebben een enorme hoeveelheid informatie over hoe dingen werken en zich gedragen in onze wereld. Wij leren deze informatie gaandeweg als we opgroeien. Van veel van deze kennis zijn we ons niet eens bewust omdat het "vanzelfsprekend" lijkt te zijn. Daarom noemen we het "Common sense Knowledge", of in goed Nederlands "vanzelfsprekende kennis". Om machines de wereld om ons heen te laten begrijpen, hebben zij ook deze grote hoeveelheid vanzelfsprekende kennis nodig.

Verschillende grote bedrijven, en zelfs overheden, hebben aangegeven dat ze het vermogen om machines deze "vanzelfsprekende kennis" te laten leren, als de volgende grote stap zien naar breed toepasbare Artificiële Intelligentie.

## Wat is het verschil tussen ASTRID en GPT-3

ASTRID leert (voornamelijk) van tekst. Ook taalmodellen zoals GPT-3 en Googles LaMDA worden getraind met tekst. Het grote verschil ligt in wat er wordt geleerd. Transformer-modellen, zoals GPT-3 en LaMDA, leren statistisch hoe woorden elkaar opvolgen in een zin. Deze modellen voorspellen vervolgens welke zin als antwoord kan dienen, op basis van deze woord volgordes. Ze kunnen niet beredeneren of het gegeven antwoord klopt, of zelfs een relevant antwoord is. Er zijn dan ook (nog) geen duidelijke toepassingen voor systemen zoals GPT-3.

ASTRID vindt relaties in tekst, en bouwt een wereldmodel met deze informatie. ASTRID leert hoe dingen in onze wereld zich gedragen, hoe ze relateren en interacteren met elkaar, en elkaar (soms) beïnvloeden. Een antwoord van ASTRID is niet gebaseerd op statistische woordparen, maar op begrip over onze wereld. Hiervoor zijn natuurlijk wel directe toepassingen te vinden.

## Wat kan ASTRID doen

ASTRID kan complete kennisdomeinen leren, en dit onvoorstelbaar veel sneller doen dan dat mensen dit kunnen. Met deze kennis kan ASTRID antwoord geven op complexe vraagstukken, of directe besluiten nemen als onderdeel van machines die in onze wereld moeten navigeren, of zelfs daarbuiten.

## Enkele voorbeelden:

### Robotica

Robots kunnen worden ingezet voor diverse belangrijke taken. In de gezondheidszorg, om het groeiende tekort aan personeel op te vangen. In rampenbestrijding, wanneer mensen het gebied niet in kunnen om mensen te redden vanwege straling, vuur, of andere problemen. En in ruimtevaart, waar robots zaken kunnen versnellen en helpen om de mensheid verder te brengen.

Hiervoor moeten robots grotendeels of volledig zelfstandig kunnen werken, want in al deze scenario's zijn er geen mensen beschikbaar of nabij genoeg, om deze machines op afstand te besturen.

### Beslissingssystemen

Beslissingssystemen, bijvoorbeeld in de medische wereld en de advocatuur, zijn nu gebaseerd op statistische informatie. In de medische wereld is dit zeer succesvol gebleken voor het herkennen van bijvoorbeeld melanomen. In de juridische wereld is het gebruik grotendeels gefaald.

Als een beslissing gebaseerd moet worden op een complexe set van feiten, dan werkt de statistische methode niet meer. Het is niet mogelijk om de complexiteit te voorspellen en daarom ook onmogelijk om een statistisch systeem te trainen voor dergelijke situaties. ASTRID leert echter specifiek door het vinden van relaties in complexe informatie. ASTRID kan dergelijke relaties zelfs vinden tussen verschillende kennisgebieden.

### Zelfrijdende auto's

De ontwikkeling van zelfrijdende auto's staat al een tijd stil. Er zijn al lang geen belangrijke doorbraken meer bereikt. Statistische patroonherkenning heeft zijn limiet bereikt. Patroonherkenning is fantastisch voor het herkennen van objecten rond de auto, maar het kan niet anticiperen op de effecten van een constant veranderende situatie. Het kan alleen reageren op basis van de trainingsdata, wat alleen zaken bevat die al eens gebeurd zijn.

ASTRID heeft kennis over hoe dingen zich gedragen in de wereld, en kan het resultaat beredeneren van een combinatie van deze gedragingen. ASTRID kan nieuwe situaties beoordelen, op basis van het gedrag van de onderdelen die in de situatie voor komen. Als ASTRID bijvoorbeeld is getraind met gedragsinformatie over hoe kinderen in het algemeen spelen, dan kan het beredeneren wat een spelend kind kan doen op een voetpad, en dit meenemen in de beoordeling van de situatie.

# ASTRID

## Fact Sheet

