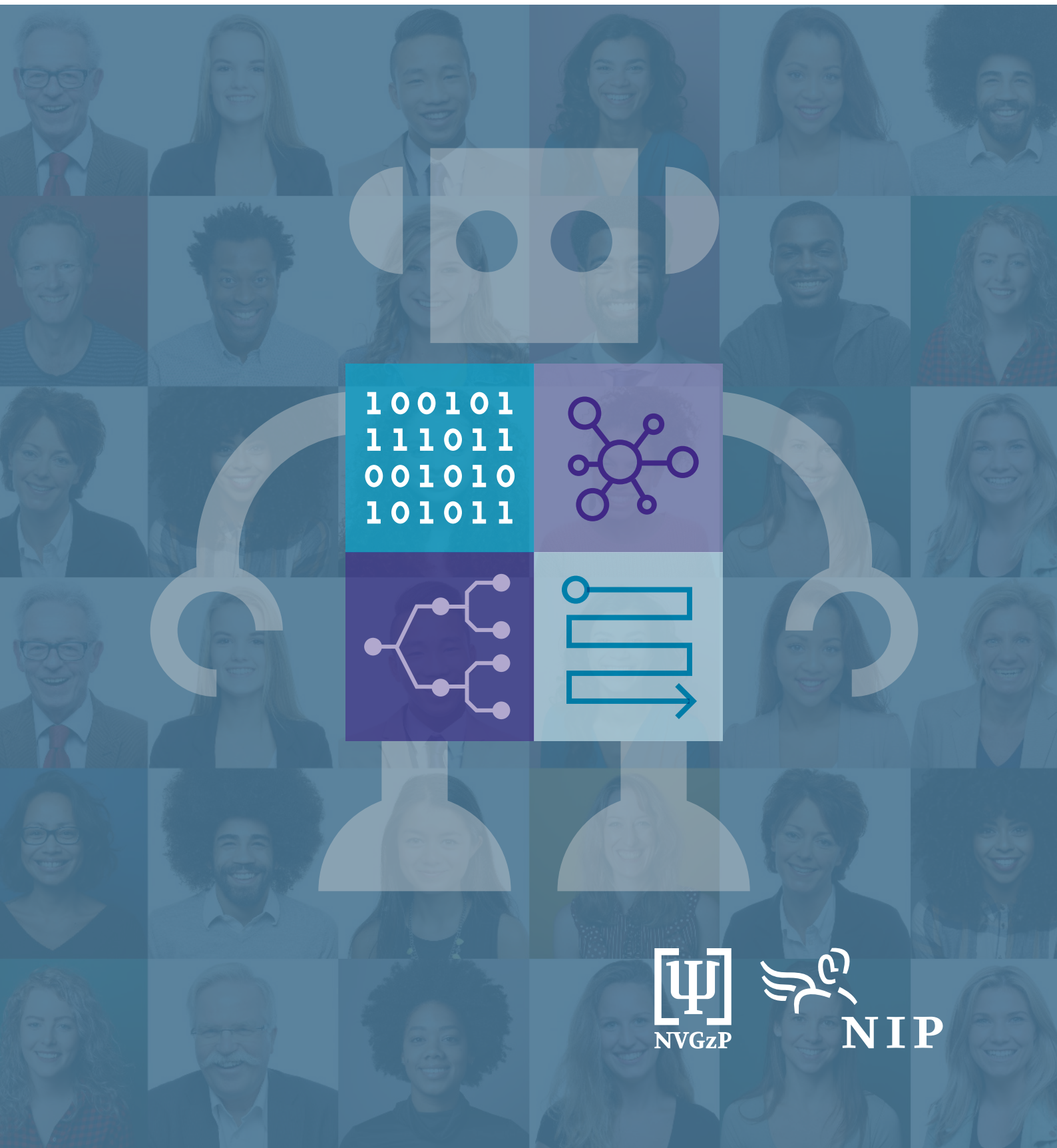


# AI en Personeelsselectie: Het Psychologische Perspectief



# AI en Personeelsselectie: Het Psychologische Perspectief

Artificiële Intelligentie (AI) speelt een steeds grotere rol bij het selecteren van personeel. Zo maken werkgevers al gebruik van AI om LinkedIn-profielen, cv's of digitale interviews automatisch te scoren. De coronacrisis heeft deze trend versterkt, aangezien AI het mogelijk maakt om een groot aantal kandidaten op afstand te screenen.

Psychologische personeelsselectie heeft als doel om relevant werkgedrag voor een functie te voorspellen. Bij het selecteren van kandidaten wordt in kaart gebracht welke kenmerken ze moeten bezitten om de functie goed te kunnen uitoefenen. Dit stelt eisen aan de instrumenten die je gebruikt om de kenmerken in beeld te brengen/te meten. Of je nu AI gebruikt of traditionele psychologische selectie-instrumenten.

AI is een veld binnen de computerwetenschappen waarbij menselijke intelligentie wordt nagebootst. Bij het selecteren van personeel houdt dit in dat AI gebruikt wordt om menselijke oordelen te automatiseren. Doordat AI de oordeelsvorming standaardiseert, belooft het gebruik van AI niet alleen meer efficiëntie maar ook meer objectiviteit.

Toch zijn er ook risico's verbonden aan het gebruik van AI. Een bekend voorbeeld is de AI die Amazon had ingezet om hun selectieprocedure te automatiseren. Al snel bleek dat vrouwen een grotere kans hadden om afgewezen te worden op basis van de AI dan mannen. Er zijn echter nog veel meer voorbeelden waarbij het gebruik van AI tot discriminatie van bepaalde groepen heeft geleid. De Europese Commissie ziet dan ook een hoog risico bij het gebruik van AI voor personeelsselectie, waarvoor strenge voorschriften gaan gelden.

Door deze ontwikkelingen zijn veel werkgevers en HR-adviseurs wantrouwig ten aanzien van AI. Het Nederlands Instituut van Psychologen (NIP) begrijpt dit wantrouwen. Net als andere psychologische tests, zouden AI-toepassingen namelijk aan verschillende eisen moeten voldoen alvorens deze wordt ingezet voor het nemen van belangrijke beslissingen. Binnen het NIP beoordeelt de Commissie Testaangelegenheden Nederland (COTAN) of psychologische tests aan deze eisen voldoen. Tot op heden hebben AI-leveranciers echter nog niet aangetoond dat hun toepassingen aan dezelfde COTAN-eisen voldoen.

In dit artikel zetten we enkele belangrijke eisen van psychologische tests op een rij en leggen we uit waarom deze eisen ook van toepassing zijn bij verantwoord gebruik van AI.

## Eis 1: De AI is gebaseerd op theorie

Binnen de psychologie wordt al meer dan 100 jaar onderzoek gedaan naar eigenschappen van mensen. Zo zijn er verschillende modellen ontwikkeld die laten zien dat een beperkt aantal eigenschappen (bijv. bepaalde persoonlijkheidstrekken, intelligentie) gerelateerd is aan hoe mensen presteren op het werk. Ook weten we dat voor verschillende banen verschillende eigenschappen van belang zijn. Een goede selectieprocedure begint daarom met een zogeheten functieanalyse waarin deze eigenschappen op een rij worden gezet.

Bij de ontwikkeling van AI wordt bestaande psychologische kennis echter vaak genegeerd. De AI-ontwikkelaar is gericht op het vinden van een optimaal verband tussen de aangeleverde gegevens van kandidaten die verzameld zijn in een selectieprocedure (bijv. informatie op cv's) en aangeleverde scores op een uitkomstmaat. De uitkomstmaat is hetgeen men wil voorspellen op basis van de selectieprocedure (bijv. beoordelingen van werkprestatie). Bij de ontwikkeling van AI wordt meestal niet gekeken of er met deze gegevens relevante eigenschappen worden gemeten.

Het gevaar van dit soort AI is dat men de achterliggende reden niet weet waarom de AI werkprestatie voorspelt en dus ook niet voor welke banen de AI gebruikt kan worden. Ook is de kans groot dat de AI is gebaseerd op gegevens die niet uitlegbaar zijn (eis 2) of die tot discriminatie kunnen leiden (eis 6). Zo nam de AI van Amazon mee waar kandidaten gestudeerd hadden; een universiteit waar voornamelijk vrouwen studeren leverde een lagere score op.

AI-toepassingen die wel gebaseerd zijn op psychologische kennis beogen specifieke eigenschappen te meten. Net als bij andere psychologische tests kan de ontwikkelaar in dat geval aantonen dat de AI-toepassing ook echt deze eigenschappen meet en geen irrelevante gegevens. Dit kan bijvoorbeeld door de AI-toepassing samen met andere tests af te nemen die dezelfde eigenschappen meten. Indien de scores samenhangen, dan levert dit bewijs op dat de AI-toepassing ook echt de boogde eigenschappen meet.

Bij het ontwikkelen van AI is het dus belangrijk om te bouwen op de psychologische kennis over menselijke eigenschappen. Dit maakt het mogelijk om aan te tonen wat er met de AI gemeten wordt. Daarnaast voorkomt men dat de AI gegevens meeneemt die irrelevant zijn voor de baan.

## Eis 2: De AI is uitlegbaar

Als werkgever of HR-adviseur moet je uit kunnen leggen hoe je selectiebeslissing tot stand is gekomen. AI is echter vaak ondoorzichtig en moeilijk te begrijpen vanwege de grote complexiteit van het algoritme. Daarnaast zien veel AI-ontwikkelaars deze informatie als bedrijfsgeheim.

Als werkgever ben je verantwoordelijk voor de AI die je gebruikt. AI-ontwikkelaars dienen dus inzicht te geven in hoe de AI-toepassing is ontwikkeld. Enkele belangrijke vragen om te stellen zijn: Onder welke groep kandidaten zijn de gegevens verzameld die voor de ontwikkeling van het algoritme zijn gebruikt? Welke uitkomsten is de AI-toepassing getraind te voorspellen? Welke gegevens van kandidaten zijn er opgenomen in het AI-model? In welke mate beïnvloeden deze gegevens de uiteindelijke score?

Deze informatie is niet alleen noodzakelijk om de AI-toepassing uit te kunnen leggen aan kandidaten, maar ook om de kwaliteit van de AI-toepassing te kunnen controleren. Hierbij is het ook belangrijk om de letten op de uitkomstmaat die gebruikt is om de AI te ontwikkelen. Indien de AI getraind is om menselijke oordelen te voorspellen, is de kans groot dat de AI de bias in deze oordelen heeft gekopieerd. De ideale uitkomstmaat is dus zo objectief mogelijk.

### **Eis 3: De AI levert consistente scores op**

Voor elke psychologische test geldt dat scores niet te veel afhankelijk moeten zijn van toevallige factoren. Het zou dus niet moeten uitmaken op welke dag, door welke HR-adviseur, of op welke computer de test wordt afgenomen. Met andere woorden, de test dient consistente scores op te leveren. Hetzelfde geldt voor AI-toepassingen.

Dit betekent dat de AI-ontwikkelaar moet kunnen aantonen dat indien dezelfde kandidaten op twee verschillende tijdstippen worden gescreend door de AI-toepassing, hun scores grotendeels overeenkomen. Voor de meeste AI-toepassingen is echter onbekend in hoeverre deze consistente scores opleveren. Dit bewijs is echter noodzakelijk voordat de AI-toepassing gebruikt kan worden voor personeelsselectie. Selectiebeslissingen dienen namelijk niet afhankelijk te zijn van toevallige factoren.

### **Eis 4: De AI voorspelt relevant gedrag**

Indien een psychologische test gebruikt wordt voor selectiebeslissingen, dan zou de testontwikkelaar aan moeten tonen dat de test scores samenhangen met relevant werkgedrag (bijv. werkprestatie). Met andere woorden, een kandidaat met een relatief hoge score op de test zou ook meer relevant werkgedrag moeten laten zien.

Sommige AI-toepassingen zijn specifiek ontwikkeld om relevant werkgedrag te voorspellen (zie eis 1). Ondanks dat vaak de theoretische basis ontbreekt voor deze AI-toepassingen, kan de AI-ontwikkelaar relatief makkelijk aantonen dat de scores samenhangen met dit gedrag.

In veel gevallen is de AI-toepassing echter ontwikkeld op basis van aangeleverde uitkomstmaten die werkgedrag benaderen (bijv. selectiebeslissingen van recruiters of lijnmanagers) of die psychologische eigenschappen weergeven (bijv. persoonlijkheidsinschattingen van getrainde beoordelaars). Het zal duidelijk zijn dat deze uitkomstmaten niet in de buurt komen van echt werkgedrag. In zulke gevallen dient de AI-ontwikkelaar dan ook aan te tonen dat de AI-scores ook echt gerelateerd zijn aan het gedrag van werknemers in de baan.

Dit is waarschijnlijk de meest uitdagende eis, die geldt voor alle psychologische tests en ook voor AI. Om aan te tonen dat een AI-toepassing relevant gedrag voorspelt, dient de AI-ontwikkelaar twee keer gegevens te verzamelen onder dezelfde groep kandidaten: de AI-scores op het moment dat ze solliciteren op een baan en een meting van relevant gedrag als ze de baan al enige tijd bekleden.

## Eis 5: De AI is gebaseerd op een grote en representatieve steekproef

Om de AI te kunnen gebruiken voor het nemen van belangrijke beslissingen, is het van belang dat de aangeboden AI is ontwikkeld op basis van een voldoende grote en representatieve groep kandidaten. De AI-leverancier dient deze informatie te verschaffen, zodat je kunt nagaan of de AI wel geschikt is voor de selectieprocedure binnen de eigen organisatie.

Een kleine groep kandidaten – voor AI is minder dan 800 al gauw klein – kan het probleem opleveren dat de AI gebaseerd is op toevalligheden in de steekproef. In dat geval zal de AI geen goede voorspellingen van werkgedrag opleveren. Een minder representatieve groep kandidaten kan ook problemen opleveren. In dat geval zal de AI geen goede voorspellingen opleveren voor kandidaten die niet of nauwelijks voorkwamen in de steekproef. Indien de AI bijvoorbeeld is ontwikkeld op basis van een steekproef van voornamelijk mannen of hoogopgeleide kandidaten, is het onwaarschijnlijk dat deze ook werkgedrag kan voorspellen voor vrouwen of laagopgeleide kandidaten. In het ideale geval is de AI-toepassing ontwikkeld op basis van een steekproef van kandidaten die overeenkomt met de beoogde doelgroep.

## Eis 6: De AI discrimineert niet

Voor psychologische tests geldt dat persoonlijke kenmerken (bijv. leeftijd, geslacht, etniciteit, sociaal-economische klasse) die niet van belang zijn voor de te meten psychologische eigenschap of voor het werk geen invloed mogen hebben op de scores. Op het moment dat deze kenmerken wel invloed hebben, is de kans groot dat bepaalde groepen benadeeld worden bij het gebruik van de test.

Dit is een van de redenen waarom een AI-toepassing uitlegbaar moet zijn (eis 2). Op het moment dat de ontwikkelaar inzicht geeft in de gegevens waarop de AI-toepassing gebaseerd is, kan gecontroleerd worden of de AI-toepassing persoonlijke kenmerken meeneemt. Indien nodig, kan daar vervolgens voor worden gecorrigeerd. In de meeste gevallen betreffen het echter indirecte metingen van deze kenmerken, zoals de postcode op een cv (indirecte meting van sociaal-economische klasse) of gezichtskenmerken in een virtueel interview (indirecte meting van leeftijd, geslacht en etniciteit).

Dit risico is kleiner voor AI-toepassingen waarmee specifieke psychologische eigenschappen (bijv. persoonlijkheidseigenschappen) worden gemeten. Het risico kan echter nooit uitgesloten worden. Ook voor deze AI dient de ontwikkelaar dan ook aan te tonen dat bepaalde groepen niet benadeeld worden.

## Gouden driehoek

Om een AI-toepassing te ontwikkelen die voldoet aan deze eisen, is het raadzaam om computerwetenschappers samen te laten werken met A&O-psychologen en juristen. A&O-psychologen kunnen keuzes wetenschappelijk onderbouwen, functie-analyses uitvoeren, relevante eigenschappen identificeren en advies geven met betrekking tot de dataverzameling. Aangezien er bij de ontwikkeling van AI persoonlijke gegevens worden verzameld, dienen ook de AVG en de Europese wetgeving voor het gebruik van AI in acht genomen te worden. Daarom dient men al in een vroeg stadium advies in te winnen bij juristen.

## Check je leverancier goed

Alvorens AI in te zetten voor het selecteren van personeel, adviseren wij om bij AI-leveranciers na te gaan in welke mate hun AI aan deze zes eisen voldoet. Indien de AI-leverancier dit bewijs niet kan leveren, dan zou de AI-toepassing (nog) niet gebruikt moeten worden voor selectiebeslissingen. Als werkgever of HR-adviseur kun je wel aanbieden om de AI-toepassing mee te laten lopen in de selectieprocedure om zo de leverancier te helpen met het verzamelen van het benodigde bewijs. Zo kunnen we gezamenlijk toewerken naar het verantwoord gebruik van AI-toepassingen in personeelsselectie.

### De eisen op een rij:

**Eis 1** De AI is gebaseerd op theorie

**Eis 2** De AI is uitlegbaar

**Eis 3** De AI levert consistente scores op

**Eis 4** De AI voorspelt relevant gedrag

**Eis 5** De AI is gebaseerd op een grote en representatieve steekproef

**Eis 6** De AI discrimineert niet



## Meer weten?

- Aan de slag met AI bij werving en selectie? Een **Arbeid & Organisationspsycholoog NIP** kan u helpen. U vindt er een via de **website van het NIP**.
- Of lees eerst meer over de **Registerpsycholoog NIP in het werkveld Arbeid- en Organisationspsychologie** en de vakbekwaamheid waar u op kunt rekenen als u hen inschakelt.

Deze paper is in 2022 opgesteld door Janneke Oostrom, arbeid- en organisationspsycholoog en Associate Professor of HRM and Organizational Behavior bij de VU in Amsterdam, in opdracht van de werkgroep Artificiële Intelligentie van het Nederlands Instituut van Psychologen. NIP-leden Jeroen Meliëzer, Marise Born, Jozien Silvester en Cora Reijerse danken we eveneens voor hun bijdrage aan deze paper. Ook danken wij de stakeholders en andere experts die we in aanloop naar deze paper hebben gesproken.

Het NIP is de grootste vereniging van psychologen in Nederland. We zetten ons in voor de psycholoog als professional en voor de psychologie als vak. Dit doen we door de standaard van de professionals hoog te houden en het vak stevig op de kaart te zetten. We zijn een serieuze gesprekspartner in Den Haag en werken samen met diverse andere belangenverenigingen. Bij vragen kun je contact opnemen met het NIP via [info@psynip.nl](mailto:info@psynip.nl)

Arthur van Schendelstraat 650  
3511MJ Utrecht  
(030) 820 15 00  
[www.psynip.nl](http://www.psynip.nl)

