

Zesduizend uittreksels doorspitten? Dat kan sneller, dacht deze hoogleraar

➔ **'Systemic reviews' selecteren relevante studies**

➔ Samenvattingen lezen kost nu vaak maanden

➔ Met een slim algoritme kan dat proces worden versneld

Jan Fred van Wijnen
Amsterdam

Matthijs Oud maakt richtlijnen voor psychiaters en therapeuten. Daarmee kunnen ze de beste therapie voor hun patiënt kiezen. Met een groep collega's uit Engeland en Nederland vlooidde hij 55 wetenschappelijke studies door over de behandeling van een 'bipolaire stoornis'.

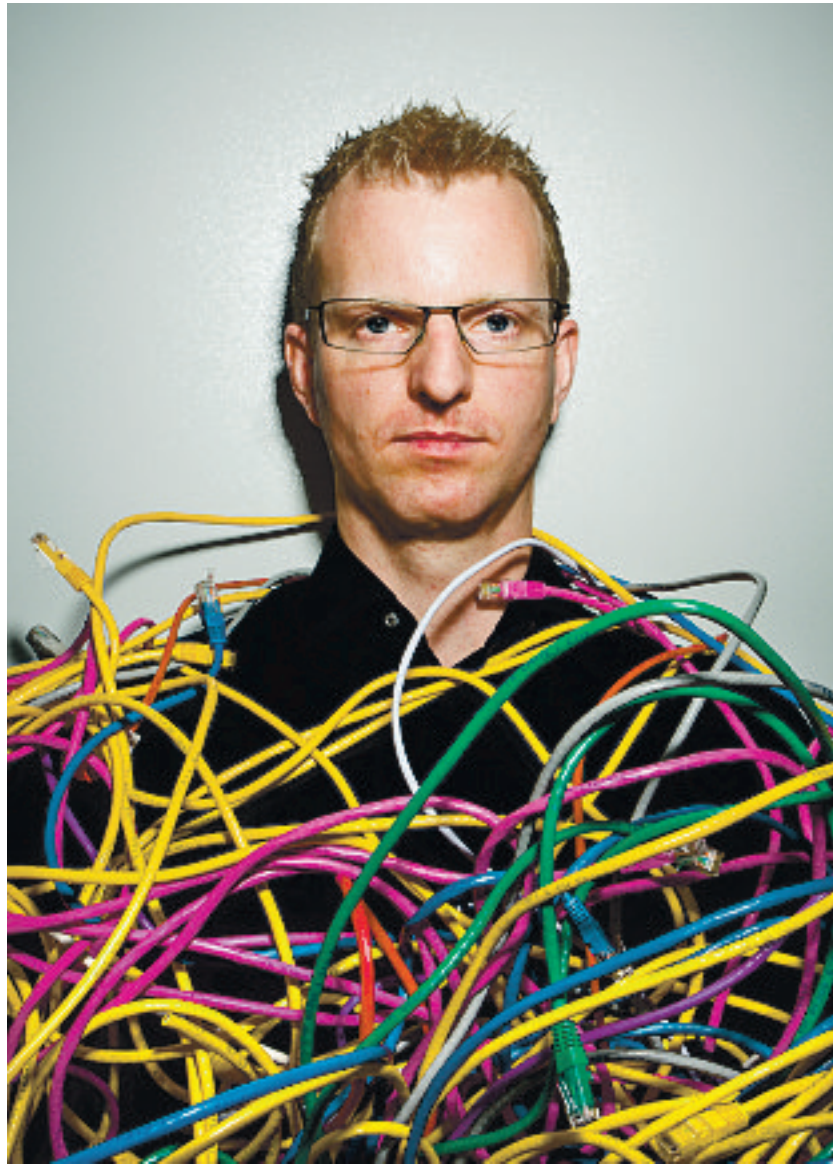
Mensen met zo'n aandoening hebben vaak grote moeite om sociaal te functioneren: de ene keer zijn ze energiek en overmoedig, de andere keer depressief en teruggetrokken. Hun conclusie: bij sommige behandelingen keert de manie of de depressie minder vaak terug, maar eigenlijk kun je niet zeggen dat de ene beter werkt dan de andere.

Teleurstellend? Niet voor een wetenschapper. 'Dit biedt therapeuten een richtlijn voor effectieve behandelingen', zegt Oud. Hij werkt bij het Trimbos Instituut, waar onderzoek wordt gedaan naar geestelijke gezondheid en verslaving.

Het was wel een enorme berg werk. De onderzoekers zochten in negen databestanden naar wetenschappelijke artikelen over bipolaire stoornissen. Ze vonden er 13.645. Volgens de regels van het onderzoek moest elke samenvatting door twee wetenschappers worden beoordeeld. Het duurde enkele maanden om de stapel terug te brengen tot 55 relevante studies.

Wat Oud en zijn collega's deden, toen ze die vele duizenden samenvattingen lezen, wordt in de wetenschap een 'systematisch overzichtsartikel' genoemd. In de gezondheidszorg zijn ze onmisbaar om behandelingen te moderniseren, maar ze worden in alle wetenschappen geschreven.

De 'systematic review' nam de laatste



jaren een hoge vlucht. Aan de Universiteit van Utrecht, bijvoorbeeld, is het aantal sinds 2008 verviervoudigd tot 283. Uit een onderzoek naar reviews wereldwijd blijkt dat voor één studie soms meer dan 90.000 samenvattingen beoordeeld moeten worden.

Rens van de Schoot, zojuist benoemd als hoogleraar statistiek voor 'kleine databestanden' aan de Universiteit van

Rens van de Schoot, hoogleraar statistiek, ontwikkelt een algoritme dat net zo goed artikelen beoordeelt als mensen kunnen. FOTO: PETER BOER

13.645

Om 55 relevante studies over een ziekte te vinden moesten wetenschappers 13.645 samenvattingen lezen.

Utrecht, had zelf ook een frustrerende ervaring met schrijven van een overzichtsartikel. Hij zocht relevante studies over het verloop van post-traumatische stress. Om er 34 te vinden, moest hij 6185 samenvattingen lezen.

Dat kan beter, vond hij. Hij vroeg een subsidie van €25.000 bij het 'innovatiefonds voor IT in onderzoeksprojecten' van zijn universiteit. Daarmee onderzocht hij of het werk geautomatiseerd kan worden. Er bestaan al verschillende programma's en websites voor, maar de meeste reduceren het aantal artikelen maar met zo'n 30%. Samen met andere onderzoekers en programmeurs probeerde hij een slim algoritme te ontwikkelen, dat net zo betrouwbaar leest als een mens.

De eerste experimenten lijken succesvol. Het algoritme oefende op reviews die al eerder werden gedaan door mensen. Het resultaat: tot 80% van alle samenvattingen werd eruit gefilterd als 'niet relevant'. De machine negeerde steeds één artikel, dat door de mens wél als relevant werd beschouwd.

Voor Matthijs Oud is zo'n misser nog 'eng'. Hij zegt: 'Ik word onzeker als ik van tevoren weet dat er één artikel wordt gemist.' Maar Van de Schoot vindt dat zijn algoritme niet onderdoet voor het handwerk. 'Ook een mens kan één relevant artikel over het hoofd zien', zegt hij. De beste oplossing lijkt hem reviews te laten doen door één mens en een algoritme, in plaats van twee mensen. 'Dat scheelt al de helft van alle leestijd.'

De angst voor missers begrijpt hij wel. In de VS overleed ooit een jonge vrouw aan een middel tegen astma. Over de giftige werking was een publicatie verschenen, maar juist dit artikel was niet meegenomen in een review.