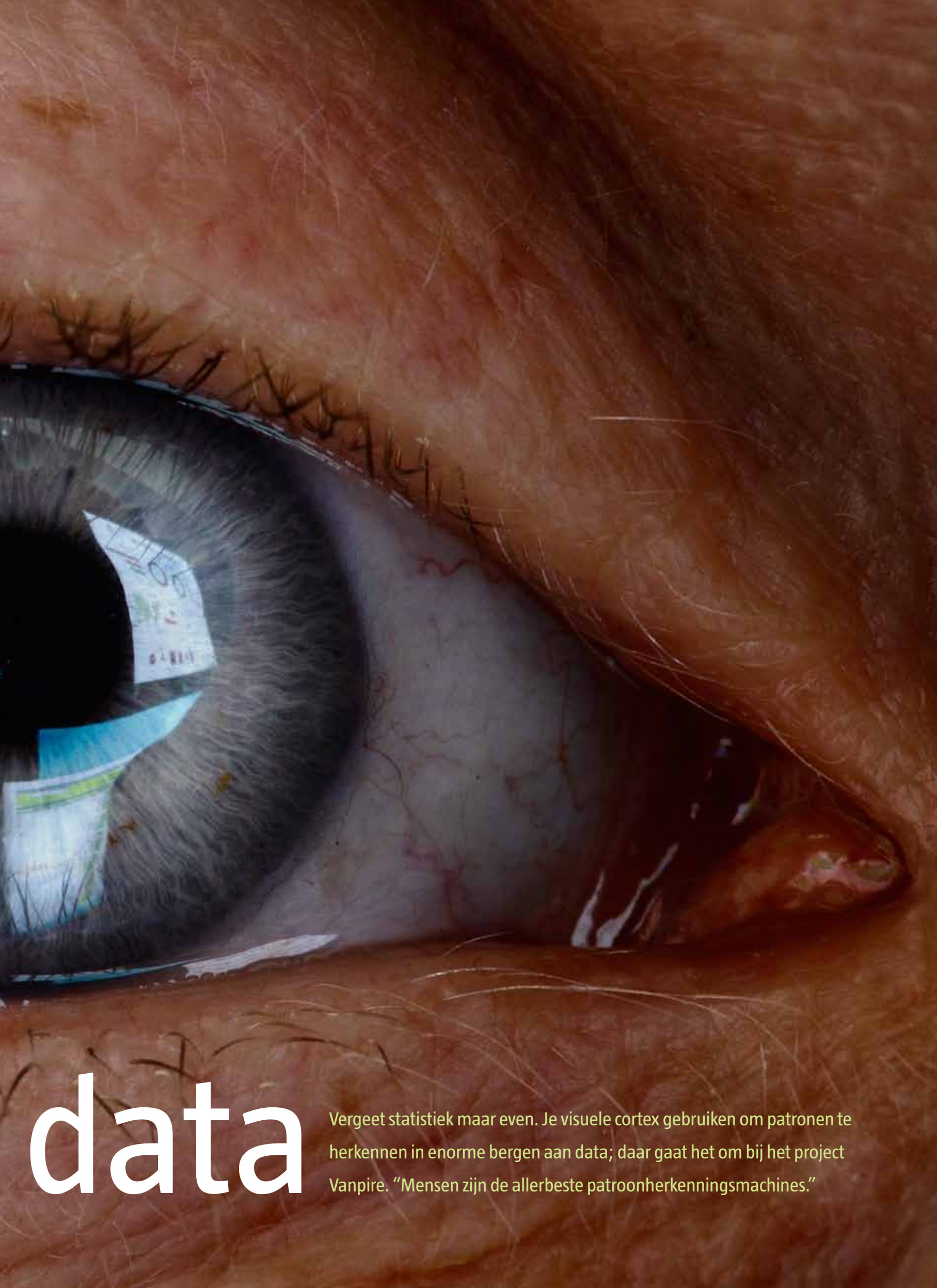




Doorzichtige



data

Vergeet statistiek maar even. Je visuele cortex gebruiken om patronen te herkennen in enorme bergen aan data; daar gaat het om bij het project Vanpire. "Mensen zijn de allerbeste patroonherkenningsmachines."

MRI-scans, hersenvolumes, scores van cognitieve tests, concentraties antistoffen in het bloed; het waren zeer uiteenlopende soorten data die de Delftse onderzoeker ir.

Martijn Steenwijk als input gebruikte voor zijn computerprogramma Papilio.

En het waren véél data. Zowat alle gegevens afkomstig uit een grote Leidse studie naar 156 patiënten die leden aan een variant van de auto-immuunziekte lupus, nam Steenwijk onder de loep. Uit die enorme berg data moest iets interessants te destilleren zijn, redeneerde hij. En inderdaad, na lang spelen met variabelen, combineren van gegevens, creëren van clusters, inzoomen en weer uitzoomen, werden bepaalde verbanden duidelijk. Zo bleek er een link te zijn tussen geheugenverlies bij lupuspatiënten en de aanwezigheid van de antistof lupus anticoagulant.

Het onderzoek van Steenwijk, die in 2010 cum laude in Delft afstudeerde, vormt het voorwerk van het project Vanpire. Vanpire staat voor visual analysis in population imaging research. De bedoeling van dit project is medici in staat te stellen op interactieve en intuïtieve manier grote hoeveelheden data te bekijken die op het eerste gezicht moeilijk te combineren lijken.

Dr. Charl Botha is penvoerder van de sectie computer graphics & visualisation (EWI). Hij wil het

computerprogramma Papilio de komende vier jaar verder uitbouwen, samen met collega's uit Leiden en Rotterdam. Zo kunnen nog meer gegevens - waaronder ook genetische data - toegevoegd worden.

HERKENNEN

Eind vorig jaar kreeg Botha driekwart miljoen euro, voornamelijk van onderzoeksfinancier STW. Hij kreeg dit geld samen met dr. Julien Milles van het Laboratorium voor Klinische en Experimentele Beeldverwerking (onderdeel van het Leids Universitair Medisch Centrum LUMC) en prof.dr.ir. Boudewijn Lelieveldt, hoogleraar biomedical imaging in Delft (EWI) en hoogleraar bij de afdeling radiologie van het LUMC.

De achterliggende gedachte bij het project Vanpire is dat mensen heel goed patronen kunnen herkennen. "Computerprogramma's kunnen makkelijk lineaire verbanden registreren, maar als de verbanden ingewikkelder worden en niet meer lineair zijn, zijn ze daar minder goed toe in staat", weet Steenwijk uit ervaring. "Je visuele cortex daarentegen is de beste patroonherkenningsmachine die er is. Door menselijke visuele perceptie handig te gebruiken, kun je dergelijke ingewikkelde patronen wél ontdekken."

BREINONDERZOEK

Bij deze manier van werken, ook wel *data driven research* genoemd, hebben de onderzoekers meestal geen hypothese over een ziekte

'Je kunt duizenden variabelen overzichtelijk representeren in één oogopslag'

die ze willen verifiëren. Daarvoor is er vaak nog te weinig bekend over de ziekte.

"Statistiek loslaten op data is natuurlijk prima", zegt Botha. "Maar dit impliceert dat je vooraf een hypothese hebt die je gaat verifiëren. Wij willen wat anders doen. Wij willen op zoek naar nieuwe hypothesen door naar de data te kijken."

Hoeveelheid grijze hersenmassa, witte hersenmassa, bloedwaarden, leeftijd, gewicht en genetische afwijkingen. Botha maakt zwaaiende bewegingen voor zijn beeldscherm en doet alsof hij via zijn computerinterface allerlei parameters naar hartenlust verandert middels imaginaire schuifjes en knopjes.

De parameters waar Botha mee wil stoeien, zullen onder meer afkomstig zijn uit de 'Rotterdam Scan Study'. In deze studie onderzoekt het Erasmus MC sinds 1990 via het maken van CT- en MR-scans van de hersenen van mensen van 45 jaar en ouder wat er gebeurt met ons brein als we ouder worden. Het onderzoek moet onder andere inzicht verschaffen in het ontstaan van Alzheimer en dementie. Inmiddels zijn er scans gemaakt van ruim negenduizend inwoners

van Rotterdam Ommoord. Deze wijk is destijds gekozen omdat hij volgens de onderzoekers model kon staan voor de gemiddelde Nederlandse bevolking. Behalve hersenscans bestaat de studie uit cognitieve tests, bloeddrukmetingen en enquêtes.

APPELS EN PEREN

Het maken van een mooi gebruikersvriendelijk computerprogramma is nog maar het halve verhaal. Een groot deel van het werk gaat zitten in het prepareren van gegevens.

"Wat we eigenlijk gaan doen is appels met peren vergelijken", zegt prof. Boudewijn Lelieveldt. "Alle variabelen moeten geschaald worden naar waarden met een vergelijkbaar bereik. Dat is de *preprocessing* stap. Vervolgens moeten er wiskundige formules ontwikkeld worden waarmee deze 'appels en peren' vergeleken kunnen worden. En het geheel moet afgewerkt worden met een mooie gebruiksvriendelijke interface om de kracht van de visuele cortex te kunnen gebruiken."

Lelieveldt beschouwt de Rotterdam Scan Study slechts als een voorbeeld van data waar Delftenaren mee kunnen gaan werken. "Wereldwijd zijn er talloze

Twintigduizend tinten grijs

Vanpire maakt deel uit van een veel groter project, genaamd Imagene (acroniem voor population imaging genetics). Dat overkoepelende project is gericht op het bestuderen van genetische data in combinatie met beelden, zoals hersenscans, verkregen bij populatiestudies. Een tiental universiteiten en onderzoeksinstituten en een vergelijkbaar aantal bedrijven doen mee aan diverse onderzoeklijnen binnen Imagene.

Een van die andere onderzoeksprojecten is 'Genes in Space', dat wordt geleid door prof.dr.ir. Boudewijn Lelieveldt en prof.dr.ir. Marcel Reinders van de sectie pattern recognition and bioinformatics (faculteit Elektrotechniek, Wiskunde en Informatica). Hebben de onderzoekers bij het project

Vanpire geen of weinig voorkennis en kijken ze 'bottom up' of de data hen iets kunnen vertellen; bij Genes in space is het net andersom. Bij Genes in space zoeken ze naar verbanden uitgaande van kennis over ruimtelijke genexpressie.

De onderzoekers gaan gebruikmaken van een onlangs gepubliceerde driedimensionale atlas van de menselijke hersenen. Deze zogenaamde 'Allen Brain Atlas' (een database die is samengesteld door het Allen Institute for Brain Science in de Verenigde Staten) geeft door de hersenen, op elk plekje (elke voxel), aan welke genen actief zijn. De activiteit van genen wordt aangegeven met grijswaarden en doordat de activiteit van alle genen gemeten wordt, heeft elke voxel dus zo'n twintig-

duizend grijswaarden.

"Het is een *mind-boggling* hoeveelheid data", zegt Lelieveldt. "Die ruimtelijke component van genexpressie op deze schaal bestond nog niet in de genetica."

De database stelt de onderzoekers in staat om zeer gericht te zoeken naar genen die een rol spelen bij ziektes. "Van bepaalde genen weten we al dat ze een rol spelen bij een specifieke ziekte. In de atlas kunnen we nagaan waar deze genen in de hersenen actief zijn. En we kunnen dan andere genen opzoeken die in exact datzelfde gebied actief zijn en die mogelijk dus ook met de ziekte te maken hebben."

