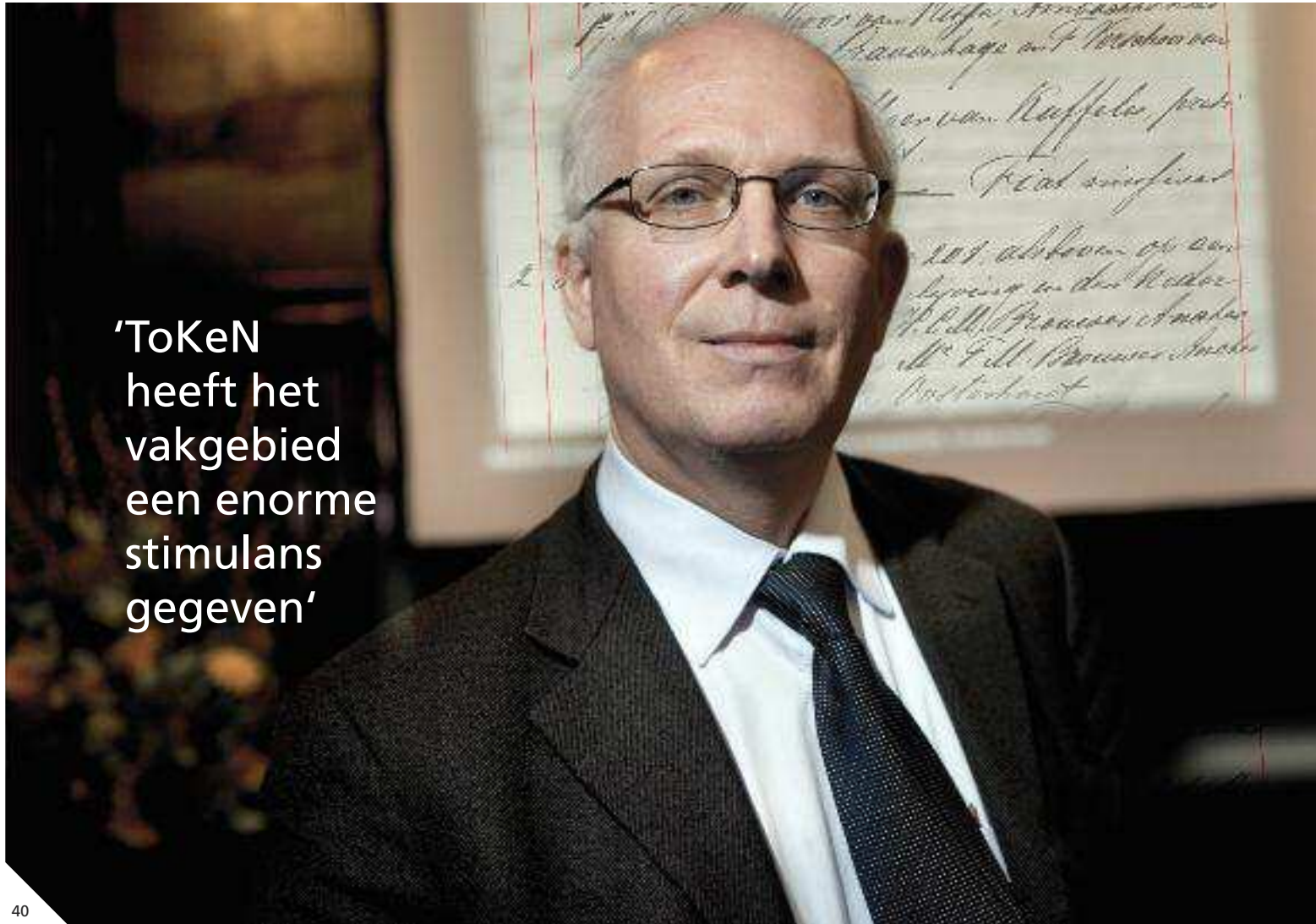


'ToKeN
heeft het
vakgebied
een enorme
stimulans
gegeven'



Forensische schrijveridentificatie

Prof. dr. Lambert Schomaker

Informatici van de Rijksuniversiteit Groningen en de Radboud Universiteit zoeken naar manieren om automatische schrijveridentificatie te verbeteren. In het project TRIGRAPH deden zij dat samen met het Nederlands Forensisch Instituut. Projectleider Lambert Schomaker vertelt hoe dat in zijn werk ging.

‘Door de enorme verbetering van beeld- en patroonherkenning kun je de computer inzetten voor taken die je eerst met de hand deed. Bijvoorbeeld om handschriften te onderzoeken’, zegt Lambert Schomaker van de afdeling Kunstmatige Intelligentie van de RUG. Hij werkt al jaren samen met de afdeling Schriftexpertise van het Nederlands Forensisch Instituut (NFI), en kent de vraagstukken van dichtbij: ‘Het gaat om handgeschreven documenten, variërend van dreigbrieven tot de handtekening onder een hotelrekening. Daarvan wil je de schrijver kunnen identificeren, of kunnen verifiëren dat het geschrevene afkomstig is van iemand van wie je al een schriftmonster bezit.’

Bijzonder

Er bestaan drie methoden voor identificatie en verificatie van handschriften, vandaar de projectnaam TRIGRAPH, legt Schomaker uit. 'Je kunt de computer naar de statistiek van kleine onderdelen in het schriftbeeld laten kijken, zoals de textuur van de inkt en de richting en kromming van de lijnen. De tweede manier is om losse letters of letterfragmenten door de computer te laten vergelijken met een tabel die lettervormen en varianten bevat. Bij de derde methode verricht een expert van het NFI handmatig metingen in het schriftbeeld. Binnen het onderzoek van TRIGRAPH zijn de eerste twee manieren verbeterd. Daarnaast hebben we onderzocht wat je met de afzonderlijke methoden kunt bereiken, en of je tot een beter systeem komt door ze te combineren.' Eerder onderzoek van de RUG leidde in 2004 tot een beeldgebaseerde methode die een doorbraak betekende. Het in Groningen uitgevoerde deel van TRIGRAPH bodde daarop voort en leidde tot het inzicht dat de helling en kromming van een handschrift niet alleen samenhangen met de schrijfhand, maar ook met de manier waarop iemand zijn pen vasthoudt. Schomaker:

'Voor mij persoonlijk een bijzonder resultaat, omdat ik vroeger onderzoek heb gedaan naar het motorisch schrijfproces. Hier kwamen mijn twee onderzoeksterreinen, beeldanalyse en motoriek, onverwacht bij elkaar.'

Uitknippen

Niet alleen de beeldgebaseerde methode is verbeterd, ook bij het vullen van de tabel met lettervormen is een grote sprong voorwaarts gemaakt. Dit deel van het project speelde zich af in Nijmegen. 'Voorheen werden letters met de hand uitgeknipt; nu scannen we de handschriften in, waarna de computer de afzonderlijke letters selecteert en deze vergelijkt met prototypes in de tabel. Daar is bij elke letter ook te vinden wat de exacte schrijfolgorde is', zegt Schomaker. 'Een schrijver wordt gekenmerkt door het specifieke gebruik van lettervormen. Voor het vergelijken van twee schrijvers gebruiken we statistische methoden en combineren we informatie over textuur en lettervorm.' De ontwikkelde methoden blijken zo krachtig, dat geschiedkundigen ze willen gebruiken bij hun onderzoek naar oude handschriften. 'Onze software helpt hen nu om de schrijvers van

oorkondes in tijd en plaats te volgen. Ook proberen historici te bepalen hoeveel verschillende personen de Dode Zeerollen hebben geschreven.' Voor zowel paleografisch als forensisch onderzoek bleek beeldverbetering belangrijk. 'We hebben de software voor beide beroepsgroepen verbeterd met de optie om vlekken en kreukels te verwijderen.'

In een vervolproject binnen het aandachtsgebied politie en justitie hopen de onderzoekers de bruikbaarheid voor forensisch onderzoek verder te vergroten: 'In rechtszaken speelt de verklaarbaarheid van resultaten een grote rol: waarom lijken twee schriftmonsters op elkaar? Het resultaat van de software laat op dat vlak nog te wensen over, de expert moet die uitleg zelf opstellen. Ook het inbedden in de werkprocessen van een forensisch laboratorium is nog een puzzel. Zo'n vervoltraject past goed in het nieuwe NWO-programma Forensic Science.'

Paradox

Gaandeweg is steeds meer helder geworden wat de relatieve voor- en nadelen van de drie benaderingen

zijn, vertelt Schomaker. 'We weten nu dat onze statistische methoden bruikbaar zijn, maar alleen als je een tekst van tenminste honderd letters hebt. Anders kun je beter een expert inzetten. De lettergebaseerde methode met een tabel blijft arbeidsintensief, omdat de gebruiker op het scherm de letters van het schriftmonster moet herschikken en controleren.'

Hoe kijkt Schomaker terug op het onderzoek? 'Het ToKeN-programma heeft het vakgebied patroonherkenning een enorme stimulans gegeven, omdat we zonder ToKeN dit type onderzoek niet hadden kunnen doen. Het is erg specifiek, en de resultaten zullen niet snel in het bedrijfsleven inzetbaar zijn. Voor zo'n project krijg je niet makkelijk een aanvraag gehonoreerd. Ondanks dat blijken de resultaten toch breed bruikbaar, wat het project ook voor de wetenschap geslaagd maakt. In de samenwerking met politie en justitie doet zich overigens een merkwaardig verschijnsel voor: de hoge werkdruk in dat werkveld vraagt om innovatieve technologie, terwijl men tegelijkertijd vanwege de hoge werkdruk niet te veel wil veranderen. Voor informatici is het een uitdaging om die paradox te doorbreken.'