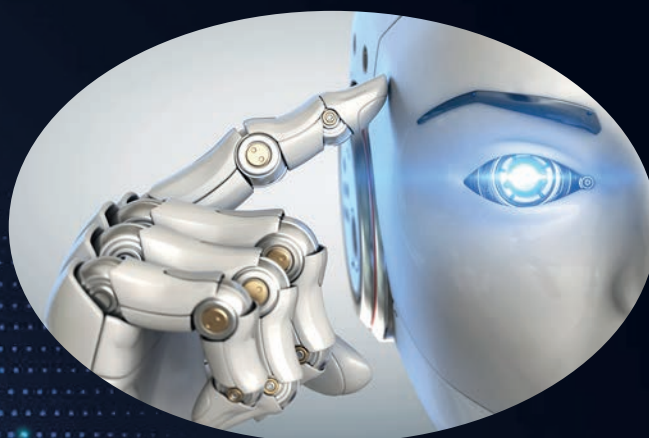




Tweeluik keynotes VOGIN-IP-lezing Van forensisch onderzoek tot taalmachinetrends

VOGIN en IP bieden jaarlijks een gevarieerd programma met lezingen en workshops rond het thema 'zoeken en vinden van informatie'. Dit jaar zijn Zeno Geradts en Antal van den Bosch de keynote-sprekers van de VOGIN-IP-lezing. Zij geven een inkijkje in hun werkpraktijk en een voorproefje van wat zij op 21 oktober gaan vertellen.



Anne van den Dool

Tekstschrijver, auteur
en cultureel journalist

Zeno Geradts:

‘Om goed te worden in het herkennen van een deepfake, moet je er ook zelf eentje kunnen maken’



Wie bewijs zoekt voor de rappe digitale ontwikkeling van onze samenleving, hoeft alleen maar naar het personeelsbestand van het Nederlands Forensisch Instituut (NFI) te kijken. Waar in 1991, toen Zeno Geradts er begon, twee mensen zich bezighielden met digitaal onderzoek, zijn dat er nu ruim tweehonderd. Met name de laatste jaren breidde die afdeling zich gestaag uit. ‘Waar collega’s destijds focusten op faxmachines en het eerste e-mailverkeer, zijn we nu massaal bezig met de analyse van in beslag genomen data’, zegt Geradts. Ook hij werkt aan digitaal forensisch onderzoek, in het bijzonder aan methodes van data-extractie van digitale gegevensdragers tijdens een delict of nadat een delict is gepleegd. Hij is hierin de verbindende schakel tussen het NFI en het Instituut voor Informatica van de Universiteit van Amsterdam (UvA).

‘Van belang is om te bepalen welk type bewijs kan worden verkregen uit de data in relatie tot het delict’, legt Geradts uit. ‘Ons doel is om te komen tot forensisch verantwoorde methodologieën en oplossingen voor

‘Van belang is om te bepalen welk type bewijs kan worden verkregen uit de data in relatie tot het delict’

de verwerking van de grote hoeveelheden data die we voorgeschiedt krijgen. Ook gaat het om het bepalen van de bewijskracht, rekening houdend met verschillende hypothesen en scenario’s. Uiteindelijk moet dit leiden tot grotere efficiëntie en kwaliteitswinst in het forensisch dataonderzoek.’

Steeds complexer

Geradts is sinds 1992 werkzaam als forensisch wetenschapper bij het NFI. Momenteel werkt hij op de afdeling Digitale Technologie en Biometrie, waar hij zich bezighoudt met beeldverwerking, camera-identificatie en biometrie. ‘Wij stellen digitale data veilig en maken die leesbaar of doorzoekbaar. Soms is dat een uitdaging, bijvoorbeeld als een telefoon in het water is gegooid of als data zijn versleuteld.’

Omdat tegenwoordig bijna iedereen in het bezit is van een mobiele telefoon, zijn de hoeveelheden data die worden verzameld gigantisch. Dat biedt mogelijkheden voor het oplossen van forensische vraagstukken, maar maakt de vijver waaruit je moet vissen om het antwoord te vinden tegelijkertijd onmetelijk groot. ‘We kunnen bijvoorbeeld zien waar mobiele telefoons geweest zijn, en dus of ze wel of niet op de plaats van het delict aanwezig waren’, vertelt Geradts. ‘We zoeken zulke zaken uit op verzoek van de politie, of op aanvraag van de officier van justitie of de rechter. Zij moeten ons een opdracht geven: of op iemands computer kinderporno aanwezig is bijvoorbeeld, maar ook welke beelden op beveiligingscamera’s aanwezig zijn of dat er sporen van witwassen op iemands pc terug te vinden zijn. We mogen niet zomaar in allerlei data neuzen.’

‘Op verzoek van politie of justitie kijken we bijvoorbeeld of op iemands computer kinderporno aanwezig is; we mogen niet zomaar in allerlei data neuzen’

Het is vooral belangrijk zo betrouwbaar mogelijk te werk te gaan en de werkwijze goed te kunnen uitleggen aan de rechtbank. ‘Dat wordt een steeds grotere uitdaging’, merkt Geradts op. ‘De forensische methodes worden steeds geavanceerder, evenals de manieren om die te omzeilen. Daarmee wordt ons werk steeds complexer.’

Automatiseren zonder biases

Het NFI kan de vraag nauwelijks aan. ‘We proberen zo veel mogelijk te automatiseren, maar we willen natuurlijk geen fouten maken. Onze methodes mogen geen biases bevatten. Bij de toeslagenaffaire hebben we gezien hoe funest dat kan zijn.’ Bovendien wordt goed afgewogen hoeveel tijd aan een vraagstuk wordt besteed. ‘We zouden wellicht meer zaken kunnen oplossen’, denkt Geradts. ‘We gaan de zwaarste middelen alleen niet inzetten voor een licht vergrijp. De officier van justitie maakt daarin een afweging; wij voeren het alleen uit.’

Voor het forensisch verantwoord zoeken in de grote hoeveelheid data heeft het NFI een systeem ontwikkeld met de naam Hansken waarin de politie kan zoeken, zodat veel vragen die weer terugkomen geautomatiseerd worden.

Deepfakedetectie

De laatste jaren is een nieuw fenomeen aan de oppervlakte komen drijven: de deepfake. Ook steeds meer burgers zien de manipulatiemethode opduiken in de media. Politici die per ongeluk in gesprek gaan met een geanimeerde versie van een collega, presidenten die via deze weg andermans woorden in de mond gelegd krijgen – het komt steeds dichterbij. ‘Bij een deepfake wordt een beeld van een mens samengesteld met behulp van kunstmatige intelligentie’, legt Geradts uit. ‘Bestaande afbeeldingen en video worden over elkaar gelegd, waardoor een pratend hoofd ontstaat dat in het best gelukte geval nauwelijks te onderscheiden valt van

‘Als een video op YouTube belandt, kunnen we maar moeilijk achterhalen wie die precies heeft gemaakt en met welke middelen’



‘Sociale platforms loven zelfs prijzen uit met grote geldbedragen voor de beste deepfake-detectie’

de echte persoon.’

De opkomst van de deepfake brengt voor digitaal forensisch onderzoek nieuwe uitdagingen met zich mee. ‘Mensen claimen al sinds de uitvinding van de fotografie dat beelden niet echt zijn. In de jaren negentig was virtuele kinderporno bijvoorbeeld niet strafbaar, waardoor wij moesten checken of beelden met de computer gegenereerd waren of niet. Ook in de rechtszaal gebruiken mensen steeds vaker het argument dat ze niet degene zijn die op een foto of video wordt getoond. Het is dan aan ons om biometrisch onderzoek te doen: we bekijken iemands lengte en postuur, en doen op basis daarvan een uitspraak. Deepfakes zijn voor ons ook redelijk nieuw: ze maken het een nog grotere uitdaging om te zien of beelden echt zijn of niet.’

Om dat te bepalen worden gelukkig steeds meer technieken ontwikkeld. ‘Toch blijft het ingewikkeld. Als een video op YouTube belandt, kunnen we maar moeilijk achterhalen wie die precies heeft gemaakt en met welke middelen. Met elk nieuw profiel creëer je een nieuwe identiteit.’

Gevecht van sociale platforms

Sociale platforms kunnen dus een grote rol spelen in de detectie van deepfakes. ‘Ook Facebook is hiermee in gevecht’, zegt Geradts. ‘Als je bijvoorbeeld plaatjes uploadt van thispersondoesnotexist.com, een site waarop automatisch portretten van niet-bestaande mensen worden gegenereerd, wordt dat herkend. Voor sociale media is die strijd noodzakelijk: hoe meer deepfakes, des te onbetrouwbaarder zijn de berichten die op hun kanalen te vinden zijn. Beelden manipuleren is tegenwoordig niet meer iets wat alleen grote filmstudio’s of geavanceerde overheden kunnen: iedere burger kan deepfakesoftware installeren en zelf aan de slag.’

Kat en muis

Het is een waar kat-en-muisspel: elke nieuwe methode voor detectie

wordt binnen no-time weer omzeild. ‘Daarom zijn we bij het NFI een samenwerking aangegaan met de UvA. Als bijzonder hoogleraar Forensic Data Science begeleid ik vijf promovendi die zich met deze onderwerpen bezighouden. Daarbij komen de kennis van bijvoorbeeld een master in kunstmatige intelligentie en de praktijk van ons werk samen. Er lopen hier studenten met de nieuwste kennis rond, die later eventueel bij ons aan de slag kunnen. Bovendien hebben wij de capaciteit niet om nieuwe algoritmes voor detectie te kunnen ontwikkelen.’

Om goed te worden in het herkennen van een deepfake, moet je er ook zelf eentje kunnen maken, weet Geradts. ‘Daar zijn deze promovendi heel bedreven in. We kunnen dat materiaal goed gebruiken om onze methodes mee te testen. Verder doen we dat met de grote hoeveelheden data die partijen als Facebook vrijgeven. Zulke platforms loven zelfs prijzen uit met grote geldbedragen voor de beste deepfakedetectie.’

Meer tijd en geld nodig

Waarom herken je een deepfake? ‘Het ingewikkeldst is misschien nog wel het goed synchroniseren van audio en video. Ook het vervalsen van de manier waarop iemand praat, is niet makkelijk; daar moet je als maker goed op oefenen. Verder werkt deepfakesoftware het best als je een beetje op iemand lijkt.’

Wat hoopt Geradts zijn luisteraars op 21 oktober mee te geven? ‘Ik hoop mensen alert te maken op de ontwikkeling dat het maken van deepfakes steeds makkelijker wordt en het detecteren steeds moeilijker. We hebben meer tijd en geld nodig om bij te blijven. Anders geloven we onze ogen binnenkort misschien niet meer en worden we wantrouwend, bijvoorbeeld als we met anderen in een digitale vergadering zitten: hoe weet je zeker dat zij het zijn? Wat daarvan de sociale gevolgen zullen zijn, kan ik niet overzien.’

Antal van den Bosch:

‘Voor AI is het herkennen van sarcasme, specifieke emoties of toxiciteit een belangrijke volgende stap’



Als directeur van het Meertens Instituut van de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) en bijzonder hoogleraar Taal en Kunstmatige Intelligentie aan de UvA is Antal van den Bosch op de hoogte van de laatste AI-trends op talig gebied. Je zou het misschien niet zeggen op basis van de nonchalante toon waarop hij zijn verhaal vertelt, maar ook hij is verbaasd over de grote sprongen die de afgelopen jaren in zijn vakgebied zijn gemaakt. Neem bijvoorbeeld de hoge vlucht die automatische vertaalmechanismen hebben genomen. ‘Systemen kunnen steeds feillozender aanvoelen welke woorden ze moeten gebruiken. Vergelijk dat maar eens met het Google Translate van tien jaar geleden. Die sprongen zijn te danken aan het feit dat steeds sterkere computers, met dank aan de rappe digitalisering, uit steeds grotere databases kunnen putten, terwijl de automatisch lerende systemen steeds verfijnder worden.’ Na een opleiding als computerlinguïst in Tilburg deed Van den Bosch promotieonderzoek bij de

‘Vertaal-
machines
zullen steeds
meer tekst
tot hun
beschikking
hebben om
zichzelf mee
te trainen,
maar er
zit wel een
einde aan de
verbeteringen’



‘In iedere
taal schuilen
bepaalde
culturele
regels en
voorkeuren,
die voor veel
machines nu
nog lastig
mee te
nemen zijn in
de vertaling’

vakgroep Informatica in Maastricht. Vervolgens werd hij hoogleraar Taal- en Spraaktechnologie aan de universiteiten van Tilburg en Nijmegen. In zijn werk staan automatisch lerende systemen centraal. Zo werkte hij met onderzoekers en programmeurs aan systemen voor automatische verwerking van teksten, aan dialoogsystemen, aan automatische vertaalsystemen en aan text mining-systemen met toepassingen in de zorg, in de media en in andere takken van wetenschap.

Klassieke onderzoeker van AI

Dat klinkt alsof hij zich bezighoudt met de laatste hypes. Toch beschouwt hij zichzelf als een ‘klassieke’ onderzoeker van kunstmatige intelligentie, die niet met alle winden meewaait. ‘Toen ik in de jaren tachtig student was, waren dialoogsystemen en neurale netwerken het hipste wat je maar kon bedenken’, herinnert Van den Bosch zich. ‘En dat terwijl de wetenschap zich ook toen al tientallen jaren met die zaken bezighield. De buitenwereld valt interessant genoeg soms technologieën op die al vele jaren bestaan.’

In zijn inmiddels ruim 25-jarige loopbaan zag hij hypes komen en gaan. ‘Interesse komt altijd in golven. In de jaren zestig zagen we de eerste opleving rondom neurale netwerken, gevolgd door een tweede hype in de jaren negentig, en de derde, daar zitten we nu middenin.’ Soms vallen zulke golven te verklaren door de historische context. ‘Kijk bijvoorbeeld naar automatisch vertalen: in de Koude Oorlog ontstond grote behoefte aan vertalingen van het Russisch naar het Engels. Nu zien we dat weer: onze globale oriëntatie laat ons veel va-

ker dan vroeger in aanraking komen met andere talen.’

Hoewel de interesse van de buitenwereld zich in golven beweegt, ontwikkelt de wetenschap zich gestaag. ‘Hypes komen voor op de momenten waarop de buitenwereld zich gaat interesseren in ons werk en de media er vervolgens mee weglopen’, aldus Van den Bosch. ‘Als onderzoekers verbazen die golven van interesse ons soms: wij zien de lange lijnen en alle tussentijdse ontwikkelingen. Soms wordt inderdaad een nieuwe techniek ontwikkeld, maar vaak is het een oude techniek in een nieuw jasje – of getraind op heel veel meer data.’

Doorbraken

Een groot deel van de ontwikkelingen op het gebied van kunstmatige intelligentie is vrij gemakkelijk te voorspellen, zegt Van den Bosch. ‘We weten van bepaalde vraagstukken, zoals automatisch vertalen, dat we die met toenemende hoeveelheden data en reken capaciteit steeds beter kunnen oplossen. De vraag is eerder: hebben we met slimmere leer methoden relatief minder trainingsmateriaal nodig? Dan kunnen de ontwikkelingen nog sneller gaan dan verwacht.’

Wat zullen in dat kader de grote doorbraken zijn op het gebied van automatische vertalingen? ‘Vertaal machines zullen steeds meer tekst tot hun beschikking hebben om zichzelf mee te trainen,’ vertelt Van den Bosch, ‘maar er zit wel een einde aan de verbeteringen. Een grote uitdaging is nog altijd het begrip van de wereld achter de taal. In iedere taal schuilen bepaalde culturele regels en voorkeuren, die voor veel machines nu nog lastig mee te nemen zijn in de vertaling. En toch: we komen al een heel eind met de strategie waarin de machine het hele internet en meer tot haar beschikking heeft om naar de beste match te zoeken met de tekst die ze voor haar neus heeft.’

Op die manier worden machines bijvoorbeeld ook steeds beter in word completion: voorspellen wat iemand

‘Niet iedereen die AI-systemen gebruikt, kan ze ook kritisch bevragen. Daarom is in de toekomst wellicht steeds meer regulering nodig’



‘Het zijn niet de technologische mogelijkheden die ons remmen, maar de terechte ethische vraagstukken’

wil typen op basis van een onaf stuk tekst en eerdere ervaringen. ‘Als je de eerste paar woorden van een tekst typt, blijft er al snel maar een beperkt aantal mogelijkheden over voor wat daarop kan volgen’, legt Van den Bosch uit. ‘Als mens kunnen we meestal wel raden wat het vervolg is – en hoe clichématiger, hoe makkelijker. Als een computer die clichés uit data kan halen, kan die ook gaan raden hoe de tekst verder gaat, zelfs zo goed dat hij na een gestelde vraag ook daadwerkelijk het antwoord geeft.’

Wapenwedloop

Op dit moment is een heuse wapenwedloop gaande tussen bedrijven die steeds beter worden in het automatisch doorgaan op onaffe teksten, vragen en dialogen. ‘Neem bijvoorbeeld GPT-3, afkorting voor Generative Pre-trained Transformer, versie 3, van het bedrijf OpenAI – een van de slimste taalgeneratiesystemen van dit moment. Zulke systemen worden steeds beter in het aanpassen van hun taalgebruik aan de juiste context; als er sprake is van een informele dialoog, vergt dat een andere stijl van vervolg dan bij een formele vraag- en antwoordsituatie. GPT-3 kan bijvoorbeeld helpen een tekst om te zetten van informele naar formelere taal, zodat ook mensen met een lager taalniveau een brief in de passende formele stijl kunnen schrijven. Of juist andersom: een tekst die eigenlijk te moeilijk voor iemand is, zet GPT-3 om in makkelijkere taal.’

Het grootste probleem van zulke slimme computers is misschien wel het ‘garbage in, garbage out’-principe. ‘De computer geeft antwoord op basis van alle teksten waarmee hij is gevoed. Daar kunnen ook racistische of haatzaaiende teksten tussen zitten. Voor kunstmatige intelligentie is dat een belangrijke volgende stap: het herkennen van sarcasme, specifieke emoties of toxiciteit.’

Datahoogtepunt

Er zijn meer mogelijkheden voor

verbetering. ‘Hoe zou het bijvoorbeeld zijn als we taal machines niet alleen de culturele dimensies van een taal kunnen aanleren, maar ook individuele? Stel je voor dat jouw telefoon precies weet welke woorden jij gaat kiezen omdat je telefoon jouw persoonlijke stijl van schrijven kent, waardoor je het aantal aanslagen drastisch kunt verminderen. Zo werkten de allereerste mobiele telefoons eigenlijk ook. Die waren gebaseerd op een woordenboek met veel financiële termen, gewoon omdat de eerste gebruikers van mobiele telefoons op Wall Street rondliepen. *Do* werd *dollar*, *st* werd *stock*. Helemaal voorspelbaar worden we aan de andere kant natuurlijk nooit. Het punt van communicatie is dat we ons met taal juist op een unieke manier willen uitdrukken en elkaar nieuwe dingen willen vertellen.’

Binnen zijn vakgebied probeert Van den Bosch zich altijd goed te blijven realiseren dat de AI-systemen die ontwikkeld worden ook steeds meer worden ingezet door mensen die niet per se technisch onderlegd zijn. ‘Niet iedereen die deze systemen gebruikt, kan ze ook kritisch bevragen. Daarom zijn in de toekomst wellicht steeds meer reguleringen nodig. AI-wapens zijn bijvoorbeeld een heel slecht idee – denk aan een gewapende drone die op basis van een foto op zoek gaat naar haar target. Het zijn niet de technologische mogelijkheden die ons remmen, maar de terechte ethische vraagstukken.’

De VOGIN-IP-lezing 2021 vindt dit jaar weer live plaats, op 21 oktober in de Openbare Bibliotheek Amsterdam (OBA). Meer weten over het gevarieerde programma van sprekers, de lezingen en de workshops? Wil je je aanmelden? Je vindt alle informatie op vugin-ip-lezing.net. <