



I/O VIVAT

JAARGANG 28
NUMMER 1

Oracle vs. Google
Het auteursrecht op een API

Thales Cryogenics

Geen tv-weerbeelden zonder
miniatuurkoelers

Het onzichtbare web

Ga daar waar geen zoekmachine geweest is

Modeling biology with ANIMO

Using computer models to understand the
communication system of living cells

Peer-to-Peer Infor- mation Retrieval

Stappen richting een gedistribueerde
zoekmachine

En verder...

Amerikaanse ISP's gaan copyright-waarschuwingen versturen

Hitachi onderzoekt permanente opslagmethode

Nieuw onderzoek naar cybercrime met psychologische invalshoek

Van de voorzitter

Van het ENIAC-bestuur

Turing Award-winnaar in beeld

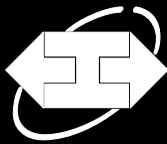
Van Rom



Inter-Actief

Advertentie
YEP IT Adver-
tentie.pdf

//Colofon



Jaargang 28, nummer 1,
December 2012
ISSN: 1389-0468

I/O Vivat is het populair-wetenschappelijke tijdschrift van I.C.T.S.V. Inter-Actief, de studievereniging voor Technische Informatica, Bedrijfsinformatie-technologie en Telematica van de Universiteit Twente. I/O Vivat verschijnt vier maal per jaar en heeft een oplage van 1800 exemplaren.

// Hoofdredactie
Herman Slatman

// Redactie
Ralph Broenink, Rick van Galen,
Ronald Meijer, Niek Tax

// Vormgeving
Niels Witte

// Gastschrijvers
Pim Jager, Rom Langerak,
Johan Noltes, Theo Rijks,
Stefano Schivo, Almer S. Tigelaar

Voor vragen, suggesties en tips is I/O Vivat bereikbaar via e-mail op vivat@inter-actief.net, twitter op @iovivat, telefonisch op 053-489 3756 of per post: Studievereniging Inter-Actief Postbus 217 7500AE Enschede

De studievereniging wil de adverteerende bedrijven bedanken voor de samenwerking.

// Drukwerk
Drukkerij van den Bosch & Fikkert

© 2012 I.C.T.S.V. Inter-Actief



I/O VIVAT

//Redactioneel

Het nieuwe collegejaar is weer spetterend van start gegaan met een uitbundige Kick-In en intussen zit je al weer een paar weken achter je bureau of in de studiebanken. Ook de redactie is na de vakantie weer hard aan de slag gegaan om een nieuwe editie uit te brengen, wat dit nummer wat extra inspanning kostte. De helft van de redactie zat namelijk op het moment van schrijven in Zuid-Korea, en dan kom je er achter dat de Vivat niet zonder haar helden kan. Ook in de toekomst kunnen we nog helden gebruiken, dus als je zin hebt om met ons mee te werken aan dit mooie blad, spreek dan vooral eens een commissielid aan.

Zoals gezegd hebben de redacteurs ook dit nummer weer geprobeerd een gevarieerde inhoud te geven aan het blad. Zo kunnen we onder andere lezen over de historie van het zoeken op het web, de problemen die aan het licht komen bij het huidige zoeken en recent onderzoek om deze problemen op te lossen. Daarnaast gaan we ook in op de heftige discussie en rechtszaak die is ontstaan tussen Oracle en Google betreffende de Java-implementatie op het Android-platform en de consequenties die dit kan hebben voor het programmeren zoals we dat kennen.

Veel leesplezier toegewenst!

Herman Slatman
Hoofdredacteur

//Inhoud 28.1



Nieuws



Oracle vs. Google



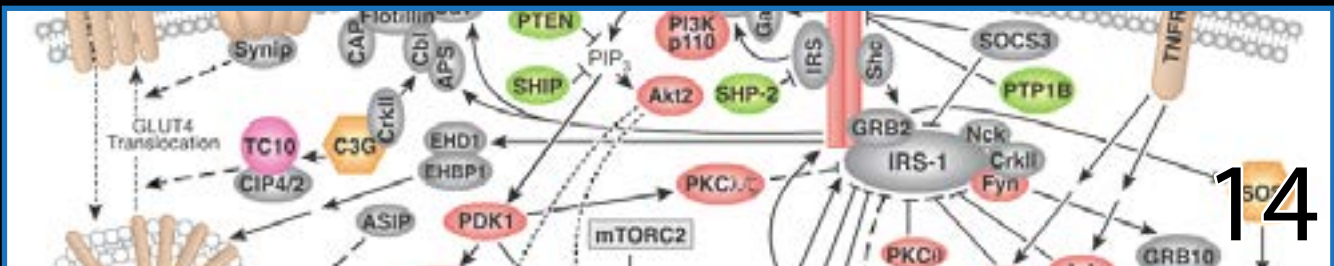
Thales Cryogenics: Geen tv-weerbeelden zonder miniatuurkoelers



Van de voorzitter



Van het ENIAC-bestuur



Modeling biology with ANIMO

Technolution



Peer-to-Peer Information Retrieval



Het onzichtbare web



Van Rom



Marvin Minsky



Volgende keer in I/O Vivat 28.2

THALES



Nieuw onderzoek naar cybercrime met psychologische invalshoek

Het onafhankelijke National Science Foundation heeft een gift verleend aan informatici van de University of California, the International Computer Science Institute en George Mason University, waarmee zij gedurende vijf jaar onderzoek zullen doen naar de psychologische aspecten van cybercrime.

‘Computers zijn slechts het medium waarover het menselijke conflict, cybersecurity, plaatsvindt’ aldus Stefan Savage, één van de hoofdonderzoekers. ‘Het verdedigen tegen cyberdreigingen vraagt niet enkel om technische kennis, maar er is ook adequate kennis van de

menselijke natuur nodig’. De onderzoekers hebben zich onder andere ten doel gesteld inzicht te krijgen in de manier waarop cybercriminelen geld verdienen en hoe zij gebruik maken van sociale netwerken als aanvalsvector en hoe zij deze sociale netwerken zelf gebruiken om contact te houden met andere cybercriminelen.

Uit eerder onderzoek bleek spam wellicht beter aangepakt kan worden op de economische aspecten dan op de technische aspecten. Zo handelden slechts drie banken 95% van de betalingen af voor producten die in spam geadverteerd werden. De geldsom die

nu beschikbaar komt zal de onderzoekers in staat stellen hun onderzoek uit te breiden, en hierbij gebruik te maken van een zeer interdisciplinaire aanpak, tussen computertechnici en –engineers en sociale wetenschappers, om zo tot de beste oplossingen tegen cybercrime te komen.

bron: http://ucsdnews.ucsd.edu/pressrelease/10_million_nsf_grant_to_help_computer_scientists_understand_the_world_of_cy

Ohio State University zet project ‘Science DMZ’ op poten

De Ohio State University heeft een subsidie van de staat Ohio gekregen om een project dat Science DMZ genoemd wordt, op te zetten. Het project is erop gericht om onderzoekers verspreid over de hele wereld via het internet te laten samenwerken zonder dat men zich zorgen hoeft te maken over de veiligheid van de communicatie.

De gedemilitariseerde zone zal bestaan als een sub-netwerk op het huidige internet, zodat gebruik gemaakt wordt van de huidige infrastructuur. “We hebben het internet nodig om effectief samen

te kunnen werken met collega’s overal op de wereld, maar het gebruik ervan vormt een risico. Het internet heeft de snelheid waarmee wetenschappelijke ontdekkingen gedaan worden enorm versneld, maar firewalls en andere noodzakelijke veiligheidsmaatregelen hinderen het delen van onderzoeksresultaten”, aldus Caroline Whitacre, vice-president voor onderzoek aan de Ohio State University en hoofdonderzoekster voor het project.

De huidige oplossingen, zoals de hogesnelheidsverbindingen die het Internet2 consortium aanbiedt, voldoen tegen-

woordig niet meer aan de eisen. Te vaak bestaan er op lokaal niveau bottlenecks in de vorm van firewalls. Door het verminderen of verwijderen van deze bottlenecks hoopt Whitacre de mate waarin onderzoeksresultaten gedeeld kunnen worden te vergroten.

De onderzoekers zullen onder andere op zoek gaan naar hardware, software, methoden en protocollen die gebruikt kunnen worden om onderzoeksdata op een veilige manier te delen.

Bron: <http://researchnews.osu.edu/archive/ScienceDMZ.htm>

Oracle vs. Google

Het auteursrecht op een API



Door: Ralph Broenink
Redacteur I/O Vivat

In april van dit jaar begon Oracle een rechtszaak tegen Google over hun Android-platform. Zij waren van mening dat de Java-implementatie van Google wel erg leek op de API's die in de Java VM zitten. De discussie heeft jaren geduurd voordat het zover als een rechtszaak kwam. Maar heeft Oracle eigenlijk wel gelijk?

De zaak van Oracle

Sinds 2010 is Oracle de eigenaar van Java; tot die tijd was Sun de eigenaar. 'Java' kan verwijzen naar drie dingen: de programmeertaal, de API's en software gemaakt met Java. In principe is de programmeertaal vrij om te gebruiken, maar de API's – alle ingebouwde objecten en klassen – zijn intellectueel eigendom van Oracle.

Voordat Android werd gelanceerd, was Java al de nummer één programmeertaal voor mobiele apparaten. Veel programma's op 'dumb' phones gebruikten al Java en het zou de overstap voor ontwikkelaars naar het nieuwe platform eenvoudiger maken. De Open Handset Alliance, de eigenaar van Android, besloot daarom om Java te gebruiken voor hun platform. "We investigated what technical alternatives exist to Java for Android. They all suck," concludeerde Tim Lindholm, ontwikkelaar voor Google, in 2010.

De Java API's die Google hiervoor open source wilde maken, zodat program-

meurs eenvoudiger over konden stappen, zouden echter eigendom zijn van Oracle. Google wilde echter geen licentie afnemen: "I don't see how we can work together and not have it revert to arguments of control," aldus Andy Ru-

tuur, vergelijkbare code en vergelijkbare documentatie.

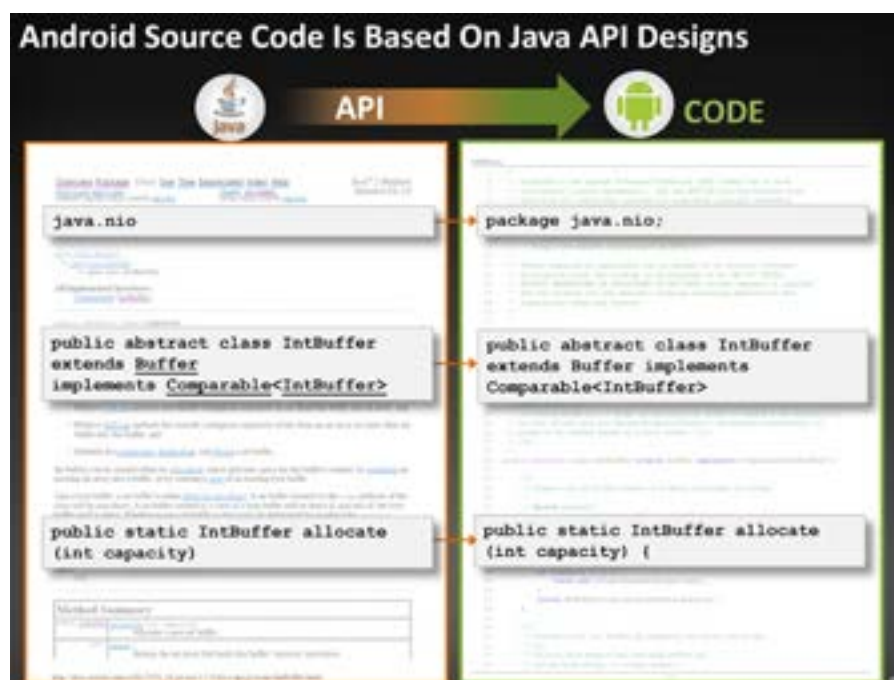
De verdediging van Google

Uiteraard is Google het niet eens met

"Nobody owns the Java programming language."

bin, één van de bedenkers van Android. Zonder licentie is Google toch aan de slag gegaan en heeft het Android uitgebracht. Met vergelijkbare API-struc-

ture. De CEO van Oracle zei in 2011 zelf dat "nobody owns the Java programming language." Hoewel er wel onderscheid gemaakt kan worden tussen



Figuur 1: Java API code overeenkomsten met de Android broncode.

de API's en de taal zelf, is Java niet functioneel zónder die API's, waardoor het dus raar zou zijn als die API's níet bij de taal horen. Bovendien heeft Google de implementatie niet letterlijk overgenomen, maar (voor compatibiliteit) alleen de structuur en naamgeving.

Anderzijds lijkt Oracle vooral veel baat te hebben bij het meeliften op het succes van Android. Er leken bij Oracle plannen te zijn geweest om een Java-gebaseerde iPhone-concurrent te bouwen, maar het 'Project Java Phone' was mislukt. Volgens Google was de rechtszaak

Oracle creatief genoeg om een auteursrecht te rechtvaardigen. Google vindt de symbolen en de grammatica zelf niet onder het auteursrecht vallen, maar een stuk tekst – of programma in Java-terminen – kan er wel onder vallen.

Het einde van programmeren zoals we het kennen?

Na een aantal lange dagen vond de jury dat Google inbreuk maakte op het auteursrecht van Oracle. Zo vonden zij dat 37 Java API's inbreuk maakten op het auteursrecht van Oracle, mits deze

ook in het gevaar komen. Dat kan nooit de bedoeling zijn.

Gelukkig is het zover niet gekomen. Voor de negen regels code die Google overgenomen zou hebben en nog eens acht Java-bestanden die gedecompileerd zouden zijn, maakt Oracle kans op ongeveer \$300.000 aan schadevergoeding. Aan de andere kant mag Oracle wel ruim een miljoen dollar aan advocatenkosten aan Google betalen. Oracle heeft dus verloren, maar de rest van de wereld kan opgelucht ademhalen. Totdat het hoger beroep dient natuurlijk.

"We investigated what alternatives exist to Java for Android. They all suck."

de laatste strohalm voor Oracle om iets op de smartphonemarkt te beginnen, en ging het niet direct om een inbreuk op het auteursrecht.

Klingon

Uiteindelijk kwam het in de rechtszaak neer op de vraag of er auteursrecht kan rusten op een programmeertaal. Google vindt dat een specifieke set instructies eventueel wel auteursrecht kan omvatten, maar dat de taal zélf slechts een concept of een middel is. Onder Amerikaans auteursrecht is dat laatste niet te beschermen. Oracle nam natuurlijk een ander standpunt in en zegt dat de manier van het opschrijven van de taal, waaronder de API's die gekopieerd zouden zijn, op zich wel origineel genoeg kan zijn om onder het auteursrecht te vallen.

Een analogie kan worden getrokken naar Klingon. De symbolen die het Klingon-alfabet maken zijn volgens

überhaupt onder auteursrecht zouden kunnen vallen. Maar tot een definitief besluit kon de jury hierover niet komen; over het in Amerika bekende 'fair use' werd de jury het niet eens.

De rechter maakte aan deze besluite-loosheid een einde toen hij besloot dat Oracle geen auteursrecht kon hebben op de API's zelf. Als het auteursrecht namelijk toegewezen zou worden, dan zou Oracle het exclusieve recht krijgen op alle mogelijke (deel)implementaties van de API, ook al heeft het maar het auteursrecht op één specifieke implementatie.

In het extreme getrokken, iedereen die een programma schrijft zou anderen ervan kunnen weerhouden vergelijkbare programma's of delen daarvan te schrijven. Zo zouden browsermakers misschien geld schuldig zijn aan Netscape voor delen van Javascript, zou Linux misschien inbreuk maken op de API's van Unix en ports van libraries zouden

```
private static void rangeCheck(int arrayLen,
                               int fromIndex, int toIndex) {
    if (fromIndex > toIndex)
        throw new IllegalArgumentException("fromIndex ("
            + fromIndex + ") > toIndex (" + toIndex + ")");
    if (fromIndex < 0)
        throw new ArrayIndexOutOfBoundsException(fromIndex);
    if (toIndex > arrayLen)
        throw new ArrayIndexOutOfBoundsException(toIndex);
}
```

Listing 1. De gewraakte rangeCheck-methode.

rangeCheck

De rechtszaak kwam uiteindelijk neer op een discussie over het ontwerp van de API, een achttal gecompileerde Java-bestanden en negen regels rangeCheck-code. Wat betreft API-ontwerp kreeg Oracle geen gelijk; maar de rangeCheck in listing 1 zou wel gekopieerd zijn. Heel spannend is deze methode niet en waarschijnlijk zou een ieder deze functie op eenzelfde manier kunnen schrijven.

Bronnen

TheVerge–TheOraclevs.Googletrial
<http://www.theverge.com/2012/4/22/2967626/oracle-vs-google-trial-java-android>

Dr. Dobb's – Oracle and the End of Programming As We Know It
http://www.drdoobbs.com/jvm/oracle-and-the-end-of-programming-as-we/232901227?itc=edit_stub

Thales Cryogenics

Geen tv-weerbeelden zonder miniatuurkoelers



Door: Theo Rijks
Thales Cryogenics

Voor optische instrumenten die vanuit de ruimte de aarde bespieden, is koeling tot 150 à 200 graden onder nul een must. Alleen zo verkrijgt u een goede signaal-ruisverhouding. Thales Cryogenics ontwerpt en bouwt de benodigde compressors voor miniatuurkoelsystemen voor nieuwe satellieten, zoals weer- en defensiesatellieten. In 2015 gaan de eerste vernieuwde versies de lucht in. De eisen aan de betrouwbaarheid van de compressors zijn extreem hoog. 'Van elke gebruikte bout moet zowel het staal als het productieproces gecertificeerd zijn. Om early failure op te sporen gaat geen enkele veer van de zuigerophanging de ruimte in zonder miljoenen testbewegingen te maken.'

Weersvoorspellingen betrouwbaarder en nauwkeuriger maken: dat is de ambitie voor de derde generatie Meteosatweersatellieten. De eerste van de zes geplande zogenoemde MTG-satellieten is waarschijnlijk in 2017 klaar voor lancering. Aan boord: hightech infraroodcamera's die vochtigheid in de atmosfeer detecteren lang voordat waterdruppels wolken vormen; en infraroodcamera's die vanuit de ruimte details zien met een dwarsdoorsnede van slechts 500 meter.

Dit lukt alleen door het ruisniveau zo laag mogelijk te houden met behulp van geavanceerde koelsystemen voor de optische instrumenten. 'Bij kamertem-

peratuur zien de hightech optische systemen in deze satellieten onvoldoende', vertelt dr. ir. Theo Rijks, programmamanager ruimtevaartprojecten en lid van het managementteam van Thales Cryogenics, vestiging Eindhoven.

Rijks gaat met zijn team de compressors voor de koelsystemen van de MTG-satellieten ontwerpen en bouwen en werkt nu al aan compressors voor een

vermoeiingsonderzoek testen we zelfs 50 miljoen keer bij een extra grote uitwijking van de veer.'

Rijks, die zowel als technisch natuurkundige afstudeerde als promoveerde aan de TU Eindhoven, vertelt verder: 'Qua ontwerp zijn de grote G-krachten en trillingen die de compressors tijdens de lancering moeten doorstaan bovendien een extra complicerende factor. In

"Thales Cryogenics combineert de slagkracht van een mkb-bedrijf met de voordelen van een groot concern"

nieuwe Franse defensiesatelliet. De compressors voor beide satellieten – afmetingen ongeveer 21 bij 11 bij 12 centimeter - moeten op velerlei vlak aan extreem hoge eisen voldoen. Een kapot onderdeel vervangen is immers geen optie als een satelliet eenmaal in een baan om de aarde draait. Ze mogen dus gedurende de levensduur van een satelliet van pakweg vijftien à twintig jaar niet kapot gaan. 'Mensen willen immers wel elke avond op het journaal mooie weerbeelden en accurate weersvoorspellingen zien. Daarom gaan bijvoorbeeld de veren van de zuigerophanging pas de ruimte in nadat ze eerst 10 miljoen testbewegingen hebben gemaakt om early failures op te sporen. En voor

dat krachtenveld mag de compressor natuurlijk niet beschadigen. Daarom hebben we in onze testomgeving schokbanken staan die krachten van meer dan 1000 G kunnen veroorzaken. Uiteraard doen we ook duurtesten. De langst draaiende compressor is in ons testlab al meer dan tien jaar aan het werk.'

Druk golf

Voor de satellieten maakt Thales compressors voor zogenoemde pulse tube cryocoolers, waarbij de compressor een drukgolf produceert die bij de 'koude vinger' van het koelsysteem voor koeling zorgt. 'Twee lineaire motoren in de compressor bewegen twee zuigers -

opgehangen in bladveren – die samen een sinusvormige drukgolf maken door tegen elkaar in te bewegen, licht Rijks toe. Bij de koude vinger van dit type koelsysteem zijn geen bewegende delen nodig. En dat is een groot voordeel voor de betrouwbaarheid en levensduur van de koelers. Extreem lage temperaturen bereikt de koude vinger van een pulse tube koeler doordat het heliumgas een compressie- en expansiecyclus doorloopt, vergelijkbaar met een Stirling-koelcyclus.

Voor een lange levensduur van de compressor mag de zuiger de omhullende cilinder niet raken. Echter voor een goede compressie, nodig voor extreem lage temperaturen tot 200 graden onder nul, moet de zuiger wel zo nauw mogelijk in diezelfde cilinder passen. Met een precisiedraaibank en een geheim uitlijnprincipe lukt het Thales om de ruimte tussen zuiger en cilinder kleiner dan 20 micrometer, pakweg 5 keer dunner dan krantenpapier, te maken. Ook het maken van de stalen bladveren vereist enorme precisie. 'Deze worden gemaakt via een etsproces om de benodigde toleranties te halen. De veren zijn zo ontworpen dat ze alleen beweging van de zuiger toestaan langs de as van de cilinder. In de richting van de cilinderwanden moet de veer juist een hele grote stijfheid hebben om te voorkomen dat de zuiger de wand raakt.'

Strakke tijdsplanning

De compressor mag ook geen trillingen doorgeven die de optische instrumenten verstoren. Rijks en zijn team kiezen er daarom voor om voor de satellietprojecten aanzienlijk meer motoren te bouwen dan er nodig zijn. 'We zijn op zoek naar de belangrijkste parameters waarmee je kunt bepalen welke twee motoren het best bij elkaar passen en dus voor de allerminste trillingen zorgen.'

De tijdsplanning voor beide satellietprojecten is heel strak. 'Drie jaar voor de lancering van de eerste MTG-satelliet moeten wij onze compressors al leveren, waarna onze opdrachtgever ze inbouwt in het gehele koelsysteem en verder test.' Deadlines bewaken is daarom een belangrijke taak voor Rijks als programmamanager. Afwijkende onderdelen die terug naar leveranciers moeten is een van de tijdvreeters. 'We eisen van onze toeleveranciers redelijk extreme

nauwkeurigheden. Daarom proberen we in een vroeg stadium met hen mee te denken, eventuele afwijkingen goed door te spreken en de meetmethodes op elkaar af te stemmen.'

Van elke gebruikte bout moeten zowel het staal als het productieproces gecertificeerd zijn. 'Dat betekent soms dat een fabrikant een speciale batch moet maken van bouten met gecertificeerd staal met bijbehorende langere levertijd. Ook daar moet je qua planning goed rekening mee houden.'

Mini-cleanroom

Voor satellieten moeten compressors bovendien veel schoner opgeleverd worden dan voor andere toepassingen. 'Voor het maken van de compressors gebruiken we technieken als lassen met een laserbundel, niet echt een schone techniek. Je gebruikt ook lijmen waarvan de dampen voor contaminatie kunnen zorgen. Zulk soort vervuiling gaan we verwijderen via een combinatie van uitstoken en spoelen met stikstofgas, waarna we de compressor verpakken in een stolp met superschoon stikstofgas, een soort mini-cleanroom.' Voordien is het nog een kunst om aan te tonen dat je de maximaal toegestane vervuiling door koolwaterstoffen niet overschrijdt. 'Hoe test je dat? Daar heb je bijna CSI-achtige methodes voor nodig door bijvoorbeeld een stukje oppervlakte af te poetsen met een steriel wattenstaafje en dat naar een gespecialiseerd lab te sturen. Ik denk nu mee met het ingehuurde lab in hoeverre onze wensen voor deze testen haalbaar én betaalbaar zijn.'

Rijks koos bewust voor deze uitdagende baan. 'We zijn een gespecialiseerde club binnen Thales die voldoende autonomie krijgt. Daardoor combineren we voor mij de slagkracht van een mkb-bedrijf met de voordelen van een groot concern. We hebben een groot moederbedrijf achter ons staan en kunnen tegelijkertijd snel van richting veranderen om op nieuwe ontwikkelingen in te spelen, zoals nu op de vraag vanuit de ruimtevaart.' Aan Rijks de taak een speciale productieomgeving voor ruimtevaartapparatuur van de grond te tillen. 'We willen deze tak van sport graag uitbreiden.'

Over

Thales Nederland is de Nederlandse dochter van het internationale Thales-concern. De omzet in 2010 bedroeg 500 miljoen euro. Het bedrijf is actief in professionele elektronica voor defensie- en veiligheidstoepassingen zoals radar en communicatiesystemen en openbaar vervoerssystemen zoals het OV-chipkaartprogramma. Thales is negende op de nationale lijst van R&D-bestedingen en de grootste hightechwerkgeveroostelijk van de IJssel.

Vestigingen: Hengelo, Huizen, Delft, Eindhoven, Houten en Enschede

Werknemers: 2.000 in Nederland, 68.000 wereldwijd in 50 landen

Actief: wereldwijd

Info: www.thalesgroup.com/nl,
www.jobs.thalesgroup.com

Van de voorzitter

Nieuw jaar, nieuwe helden



Door: Pim Jager
Voorzitter Inter-Actief

Nu de “gewone” student weer hard aan het studeren is, zijn wij ons als bestuur hard aan het aanpassen aan ons nieuwe leven. Het leven als bestuurder verschilt immers nogal van het leven als student. De agenda's zitten ineens propvol met activiteiten, borrels, afspraken en vergaderingen, het balkon en de kamer moeten netjes gehouden worden, vragen moeten beantwoord worden en daarnaast moet de vereniging ook nog gestuurd worden. Kortom, voor ons als bestuur is er nog genoeg te leren en om aan te wennen.

Voor onze nieuwe lichting is dat niet anders; zij mogen de evolutie van scholier naar student gaan maken. Simpel is dat niet, verschillen genoeg tussen het Lutjebroek Lyceum en de Universiteit Twente. De vaste stapavond wordt verplaatst naar woensdag – en soms op donderdag, en maandag, misschien dinsdag, af en toe ook nog wel op zaterdag, gewoon omdat het kan-, de voorbereiding voor hoor- en werkcolleges, eraan wennen dat voorbereiding te noemen in plaats van huiswerk, naar college gaan met docenten in plaats van lessen met leraren. Kortom, de nieuwe student heeft nogal wat te leren.

Gelukkig hoeven onze nieuwe helden dat niet alleen te doen. Nee, er is een vereniging die ze daarbij kan helpen, die ze daarij graag wil helpen. Inter-Actief doet er ook dit jaar weer alles aan om het onze nieuwe leden zoveel mogelijk bij te staan bij alle veranderingen. We helpen ze met studeren, via de tentamenbank, hulp bij klachten en natuurlijk gewoon door te helpen als ze vragen hebben over de zoveelste ogenschijnlijk onmogelijke DiWi-opgave. Daarnaast

helpen we ook met het maken van nieuwe vrienden. Op de vele activiteiten, al dan niet georganiseerd speciaal voor eerstejaars, maken ze kennis met jaargenoten en ouderejaars. Het is deze verticale band die het betrokken zijn bij Inter-Actief extra interessant maakt. Wie kan je immers beter vertellen over het nut van de huiswerkopgaven, welk keuzevak interessant is of welke werkcollegedocent je vooral moet ontwijken, dan een ouderejaars die dit alles al een keer meegemaakt heeft.

Het is dan ook mooi om te zien dat zoveel van de eerstejaars hun weg naar Inter-Actief al gevonden hebben. De nieuwe-oogstborrel liep weer ver na sluitingstijd door, en bij het aXi bowlen was het, ondanks een aantal agendaconflicten, toch gezellig druk. Goed om te zien ook dat velen weer de bank weten te vinden voor een kopje koffie en een goed verhaal.

Met de onderwijsweek en alle andere mooie activiteiten die daarop gaan volgen in het vooruitzicht kunnen we dan ook niet anders dan tot de conclusie komen dat er voor de eerstejaars genoeg ruimte zal zijn om te leren, niet alleen binnen het onderwijs, maar vooral ook over het leven als student, en de fantastische sociale kant daarvan.

Pim Jager opende voor de eerste maal zijn ogen op 29 augustus 1990 in het altijd bruisende Utrecht. Na een glansrijke carrière op basischool de Zonheuvel in het immer gezellige Driebergen (incidenteel ook de plaats waar Pim tot zijn studie in Twente woonachtig is geweest) was het tijd voor de volgende logische stap en begon Pim aan zijn VWO-opleiding aan het Revius Lyceum in het pittoreske Doorn. Na zes jaar was duidelijk dat het VWO-onderwijshemweiniguitdagingmeer kon bieden en het tijd was voor een nieuwe uitdaging.

DezenieuweuitdagingvondPimaan de Universiteit Twente, na een wat moeilijk begin, waarin de verkeerde keuzes zijn gemaakt (de nasleep van deze keuzes zijn nog terug te vinden in zijn primair lidmaatschap bij E.T.S.V. Scintilla) zag hij na twee weken studeren alsnog het licht en schreef hij zich in bij I.C.T.S.V. Inter-Actief. Na onder andere te hebben gezeten in de ECie, Kick-IT, FlitCie, Cinema, TostCie en EWL-trip is het tijd voor de volgende stap. Hij mag zich sinds 4 mei dan ook kandidaat-voorzitter/intern noemen.

Van het ENIAC-bestuur

Beperkte keuze of grote vrijheid?



Door: Johan Noltes
Voorzitter ENIAC

Op het moment dat ik deze column schrijf zijn de stemmen voor verkiezingen van de Tweede Kamer nog maar net geteld, en beginnen formateurs Henk Kamp en Wouter Bos aan hun opdracht om een kabinet te vormen. Op het moment dat jij deze column leest is er vermoedelijk een kabinet gevormd en weten we tot welk resultaat onze keuze in het stembokje daadwerkelijk heeft geleid.

We maken regelmatig keuzes, maar vaak weten we niet welke gevolgen deze keuzes hebben. Dat blijkt ook bij de Tweede Kamerverkiezingen, geen enkele opiniepeiling had voorspeld dat VVD en PvdA samen een meerderheid zouden hebben. Die onzekerheid maakt het maken van keuzes misschien wel extra lastig, waardoor we vaak gaan voor de veilige en bekende opties. Extra opvallend was het daarom ook dat alleen de Twentse Pieter Omtzigt dankzij voorkeursstemmen in de Tweede Kamer komt. De overige stemmen waren voornamelijk uitgebracht op de nummer 1 van de verschillende partijen.

Ook in het onderwijs zien we dat studenten het lastig vinden om keuzes te maken. In mijn tijd als lid van de Opleidingscommissie (OLC) Informatica hadden studenten in hun derde studiejaar nog de ruimte om 10 studiepunten te besteden aan twee vakken binnen óf buiten Informatica; een volledig vrije keuze. Je zou verwachten dat deze vrijheid een makkelijke keuze voor de student zou zijn, je mag immers iets doen wat je leuk vindt. Het omgekeerde blijkt echter het geval: juist op het moment dat de student deze brede keuze moet maken gaan de studierementen om-

laag, en loopt de student vertraging op.

Als maatregel hebben we als Opleidingscommissie geadviseerd om deze vrije keuze van 10 studiepunten drastisch te beperken: studenten kunnen nu nog maar een keuze maken van twee vakken uit een lijstje van 10 vakken. Opvallend genoeg lijkt dit goed te werken, en maken studenten nu veel sneller en makkelijker deze keuze. Doel behaald dus.

Tegenwoordig ben ik geen lid meer van de Opleidingscommissie, maar vraag ik me nog wel af hoe studenten die keuze uit dat lijstje van 10 vakken nu maken. Het antwoord zag ik onlangs op de Facebookgroep van Inter-Actief voorbij komen. De deadline voor het inschrijven voor vakken was dichtbij, en een student vroeg andere studenten om hun ervaringen met de keuzevakken. Binnen vier uur waren er twintig reacties met adviezen! Social media hebben dus een steeds grotere invloed op de keuzevakken. Dat hadden we een paar jaar geleden niet verwacht toen we de keuzeruimte beperkten in de OLC.

De invloed van social media op ons leven geldt natuurlijk niet alleen voor keuzevakken. Als bestuur van ENIAC proberen we daarom ook zoveel mogelijk gebruik te maken van social media om het doel van onze vereniging te bereiken. We zijn bezig met het inrichten van onze Facebook-groep, Twitterstream en LinkedIn-groep. Het goede nieuws is dat dit goed blijkt te werken, dankzij het gebruik van deze media was onze lustrumactiviteit een groot succes, met maar liefst 45 deelnemende alumni op het strand van IJmuiden! We nodigen je daarom ook van harte uit om lid te worden van de ENIAC community,

zodat je in contact kunt blijven met je oud-studiegenoten en andere interessante alumni.

Ik ben daarom ook enorm benieuwd wat de toekomst ons brengt: stemmen we tijdens de volgende Tweede Kamerverkiezingen nog steeds met een rood potlood, of doen we dat via een Facebook-poll?

Johan Noltes, voorzitter

Johan Noltes is voorzitter van ENIAC: de ENSchedese Informatica Alumni Club. ENIAC is de alumnivereniging voor oud-studenten Informatica, bedrijfsinformatietechnologieën Telematica aan de Universiteit Twente.

Voor slechts € 5,- per jaar kan je al lid worden van deze club. Je krijgt dan in ieder geval de Vivats die jaarlijks verschijnen (meestal zo'n 4 stuks, maar niet helemaal per kwartaal) en uitnodigingen voor de activiteiten die we organiseren (meestal per mail). Daar mag je dan vervolgens (veelal gratis!) aan deelnemen. En al doe je maar eens in de paar jaar ergens aan mee, die € 5,- kan toch bijna iedere informatica-alumnus wel missen? Zo houd je toch nog wat binding met je wetenschappelijke roots en af en toe contact met vrienden uit je studietijd.

Johan Noltes
voorzitter@eniac.utwente.nl

Modeling biology with ANIMO

Using computer models to understand the communication system of living cells

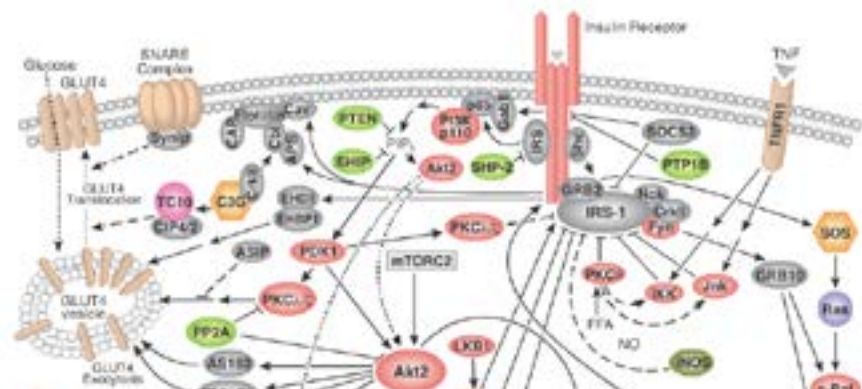


Door: Stefano Schivo
Formal Methods & Tools

When we fall sick, we take a medicine that (hopefully) makes us improve and feel better soon. But how does this medicine actually work *under the hood*? Why do we need exactly that amount of exactly those substances to get better? And what do computers have to do with this?

In the microscopic, our cells receive and process signals of many types: each type of molecule, such as the active principle of a medicine, is interpreted as a different signal. When a particular molecule comes in contact with a cell surface receptor (a sort of “detector” specific for that molecule), the receptor becomes *activated*, and proceeds to activate other molecules on the inside of the cell. These in turn activate other molecules, in a cascade that allows the signal to travel inside the cell to the nucleus, where the DNA is kept. When a signal reaches the DNA, it is interpreted as an order to produce a particular type of protein: this is considered as the response to the signal. Every cell can produce and use a large variety of proteins, some of which are also involved in signaling tasks, thus the correct “understanding” of a signal is of the utmost importance for a cell to work properly.

Moreover, often a signal does not simply travel along a “relay race” to the nucleus: interferences from other signals or molecules drive and change the response of the cell, making the task of



Signaling pathway of insuline (Cell Signaling Technology, Inc. www.cellsignal.com)

understanding how these mechanisms work very hard. As if this was not enough, performing the kind of biological experiments needed to obtain deeper insight into a particular signaling network requires considerable amounts of time (and money). It is thus very important to have a clear vision of how a signaling network is supposed to look like and behave, in order to organize existing knowledge and plan the laboratory experiments. And this is exactly what our software tool ANIMO (Analysis of Networks with Interactive Modeling) does: thanks to an intuitive user interface, ANIMO allows to *play* with a signaling network, seeing how it would evolve over time when certain signals are given. Moreover, as ANIMO is based on the formal language of Timed Automata, we can exploit the predictive power of a model by running *in silico* experiments, speeding up the process of understanding signaling networks. The cost, both in terms of time and money, of an *in silico* experiment is usually much lower than the corresponding *wet lab* experiment, and can be an asset in helping the development of medicines.

Models defined in the ANIMO user interface are automatically translated into Timed Automata systems, which are run “under the hood” thanks to the tool UPPAAL. The results are then shown to the user in the form of graphs plotting the activity intensity for each reactant in the modeled network. A slider under each graph lets the user scroll through the time points in a simulation series: the nodes in the network model are colored according to the activity of the corresponding reactants at the selected instant. Experimental data series can be superposed to simulation results, allowing for a quick comparison between model prediction and experimental facts, thus accelerating model development.

We plan to apply ANIMO in a research on tissue regeneration, aimed at making life easier for people affected by osteoarthritis: computer science can be applied even to *real science!* ;-)

For more information, see <http://fmt.cs.utwente.nl/tools/animo>

Peer-to-Peer Information Retrieval

Stappen richting een gedistribueerde zoekmachine



Door: Almer S. Tigelaar
Databases

Was vroeger het zoeken naar informatie nog veelal beperkt tot het af en toe gebruiken van stoffige grijze computers met groene schermen in openbare bibliotheken, tegenwoordig zoeken we iedere dag en gaat het allang niet meer alleen om boeken. Het World Wide Web is een enorm informatie reservoir waar we de weg niet meer in kunnen vinden zonder adequate hulpmiddelen. We zijn hiervoor hoofdzakelijk afhankelijk van enkele grote zoekmachines, denk aan Google of Bing. Deze ontsluiten het web door zoveel mogelijk documenten te downloaden, wat ook wel crawling word genoemd, en te indexeren. Hoewel ze hier soms heel veel computers voor gebruiken vallen deze allemaal onder een autonome partij, daarom worden ze ook wel gecentraliseerde zoekmachines genoemd. De enorme index van dergelijke zoekmachines kunnen mensen raadplegen door zoekvragen in te typen.

Nuttig zijn deze gecentraliseerde zoekmachines zeker, maar ze hebben ook zo hun technische beperkingen en er zijn ethische bezwaren tegen aan te dragen. Zo kunnen ze bijvoorbeeld zoekresultaten censureren, producten van eigen makelij 'boosten' in de resultaten en indirect bijvoorbeeld ook de aandelenmarkt sterk beïnvloeden. Vanuit ethisch oogpunt is de vraag of het verstandig is om dat soort macht bij een enkele commerciële partij te leggen, ook in verband

met de privacy van gebruikers, immers: macht corrupteert. Vanuit technisch oogpunt indexeren de zoekmachines maar een deel van het World Wide Web. Een ander deel, veelal verborgen achter webformulieren of bestaande uit dynamische interactieve webpagina's, is niet via hen vindbaar. Bovendien is het maar de vraag hoe haalbaar het in de

wel veel meer flexibiliteit qua oplossingen.

Het belangrijkste probleem bij een decentrale peer-to-peer zoekmachine is het bepalen welke zoekvraag naar welke peer moet gaan. Hierbij is het belangrijk dat we van elke peer een indruk hebben wat deze qua informatie beschikbaar

“Hoe haalbaar blijft het om alles te crawlen?”

toekomst blijft om alles te blijven crawlen, gegeven hoe snel het web groeit en de informatie er op veranderd.

Een alternatief voor het centrale model is een peer-to-peer aanpak. Hierbij worden de computers van mensen thuis gebruikt om een grote gedistribueerde zoekmachine op te zetten voor het web. Er wordt geprobeerd het crawlen te vermijden en zoveel mogelijk de daadwerkelijke bronnen zelf hun informatie te laten indexeren: autonome websites met een eigen lokale zoekfunctie die hun opgeslagen informatie beschikbaar stellen op de manier waarop zij dat het liefst willen. Daarnaast helpen alle computers in het netwerk, de peers, de laadbelasting te verdelen door zoekresultaten vast te houden voor hergebruik: caching. Overigens is dit model geen heilige graal, het lost niet in een keer alle problemen van de centrale aanpak op, en introduceert ook weer nieuwe technische problemen. Echter, het biedt

heeft. Deze indruk, een bronbeschrijving, slaan we op bij de zogenaamde tracker. Dit is een peer, of een groep van peers, in het netwerk die het geven van adviezen voor zoekvragen als taak heeft. Een advies is in de vorm van een lijstje van peers met vermoedelijk relevante resultaten voor een specifieke zoekvraag gesteld door een peer in het netwerk. Dit lijstje kan zowel peers bevatten die zelf zoekresultaten genereren, evenals peers die zoekresultaten voor een gelijkende zoekvraag hebben vastgehouden in hun cache.

De beschreven architectuur biedt een basis voor de ontwikkeling van een grootschalige peer-to-peer zoekmachine voor het web. Voor dat dit daadwerkelijk gerealiseerd kan worden moeten er nog wel een aantal deelproblemen worden opgelost. Echter, een alternatief voor de grote gecentraliseerde zoekmachines gloort aan de horizon.

Het onzichtbare web

Ga daar waar geen zoekmachine geweest is



Door: Herman Slatman
Redacteur I/O Vivat

Aiedereen kent de vele mogelijkheden van het internet. Een mooi voorbeeld is dat er gigantisch veel informatie beschikbaar is op het world wide web die we kunnen gebruiken voor allerlei zaken. We maken gebruik van zoekmachines als Google en Bing, die het internet crawlen en indexeren, om snel aan de informatie te komen die we op dat moment nodig hebben. Veel query's leveren bruikbare informatie op, maar er zijn er ook veel die slechte, of zelfs geen resultaten opleveren. Soms begrijpt de zoekmachine een query verkeerd, en krijg je resultaten uit de verkeerde context voorgeschoteld. Andere keren zijn de gegevens die je zoekt simpelweg niet beschikbaar op het internet, en zal je andere bronnen moeten raadplegen.

Dit artikel gaat over de informatie die wél beschikbaar is via het internet, maar die niet gevonden kan worden door gebruik te maken van een general purpose zoekmachine of web directory. De informatie bevindt zich als het ware op een onzichtbaar deel van het internet. Dat deel wordt ook wel 'the Invisible Web' of 'the Dark Web' genoemd.

In dit artikel zet ik uiteen hoe het onzichtbare web heeft kunnen ontstaan, hoe informatie op het onzichtbare web gevonden kan worden en beschrijf ik het onderzoek en de ontwikkelingen die tegenwoordig plaatsvinden om het onzichtbare web beter begaanbaar te

maken.

De historie van zoeken op het web

In de jaren '70 van de vorige eeuw werd de grondslag gelegd voor het internet waar we tegenwoordig veelvuldig gebruik van maken. Het duurde wel enkele jaren voordat er op relatief grote schaal gebruik van werd gemaakt; in die tijd bestonden de gebruikers voornamelijk uit overheids-, onderzoeks- en opleidingsinstellingen. Met de komst van het TCP/IP protocol konden meer mensen gebruik gaan maken van het internet.

Het opzoeken en opvragen van een bestand ging in de vroege jaren '90 nog erg lastig. Voor het opvragen van een bestand diende men eerst verbinding met een server te maken en daarna moest men de exacte locatie het op te vragen bestand op de server weten om het daadwerkelijk te kunnen opvragen. Zoeken naar een bestand was eigenlijk alleen mogelijk door je vraag te stellen op een mailinglist en te hopen op een antwoord met daarin gegevens om het bestand te kunnen opvragen. Het opzoeken en opvragen van bestanden was kortgezegd behoorlijk omslachtig, en er ontstond een verlangen om dit makkelijker te maken.

Archie was het eerste prototype van een zoekmachine zoals we die tegenwoordig kennen. Het programma bezocht anonieme FTP-servers op het internet en indexeerde alle bestanden die daarop te

vinden waren. Het resultaat werd weggeschreven in een centrale database die te doorzoeken was. Zo kon men door te zoeken in één database bestanden over de hele wereld vinden; een hele verbetering ten opzichte van de mailinglist. Een ander voorbeeld van een zoekmachine uit die tijd is Gopher, dat de mogelijkheden van Telnet en een FTP-cliënt combineerde om gemakkelijk, door middel van klikken, door serverdirectories te navigeren.

Beide oplossingen waren echter op hun eigen manier vrij gelimiteerd in hun mogelijkheid om het gehele web makkelijk te doorzoeken. Als antwoord op deze tekortkomingen bedacht Tim Berners-Lee het world wide web, een systeem dat ons in staat stelt om gelinkte informatie op een webserver op te vragen door gebruik te maken van een webbrowser.

Met de komst van het world wide web ontstonden ook de eerste zoekmachines. Een voorbeeld hiervan is WebCrawler, dat gemaakt werd door Brian Pinkerton. WebCrawler struinde autonoom het WWW af op zoek naar websites. De gevonden pagina's werden opgeslagen in een database met bijbehorende sleutelwoorden waarop gezocht kon worden. In april 1994 maakte Pinkerton de database via het web toegankelijk en de eerste echte zoekmachine was geboren. In de jaren die volgden kwamen er nog talloze andere zoekmachines bij, waaronder Google en MSN Search (tegenwoordig Bing) in 1998.

De ontstane zoekmachines werken zoals gezegd door middel van het afstruinen van het web en worden daarom ook wel crawlers genoemd. Links op een pagina worden gevolgd naar andere pagina's en zo kan door een groot deel van de beschikbare informatie genavigeerd worden door de crawlers. Tegenwoordig is echter in steeds grotere mate informatie in andere formaten dan tekst beschikbaar, zoals de gigantische hoeveelheden videomateriaal op YouTube. Een zoekmachine kan in beginsel niet veel met het videomateriaal zelf, maar is afhankelijk van de door mensen toegevoegde informatie zoals de titel en beschrijving van de video om deze correct te kunnen indexeren.

Voor de huidige zoekmachines is het dus niet mogelijk om actief te indexeren op materiaal dat niet bestaat uit tekst, en zo blijft een groot deel van de op het web beschikbare data ondoorzoekbaar. De indruk dat je alles op het web kan vinden door een query uit te voeren bij bijvoorbeeld Google is dan ook ongegrond. Volgens BrightPlanet was de hoeveelheid informatie die zich in 2000 op het onzichtbare web bevond enkele honderden malen groter dan de hoeveelheid informatie die zich op het zichtbare web bevindt. [THE INVISIBLE WEB 2001] schat echter dat dit

'slechts' 2 tot 50 keer zoveel is. Wat wel zeker is, is dat er een zeer grote hoeveelheid informatie aanwezig is op het onzichtbare web.

Wat is het onzichtbare web?

Een precieze definitie voor het onzichtbare web is moeilijk te geven. Makkelijker is om te beschrijven wat er op het onzichtbare web te vinden is. Kortgezegd bestaat het onzichtbare web uit (hyper)tekstpagina's en bestanden die beschikbaar zijn via het world wide web maar die zoekmachines niet kunnen of willen toevoegen aan hun index. Deze beperkingen kunnen technisch van aard zijn, maar kunnen ook ontstaan door de afwegingen die een zoekmachine maakt tijdens het indexeren.

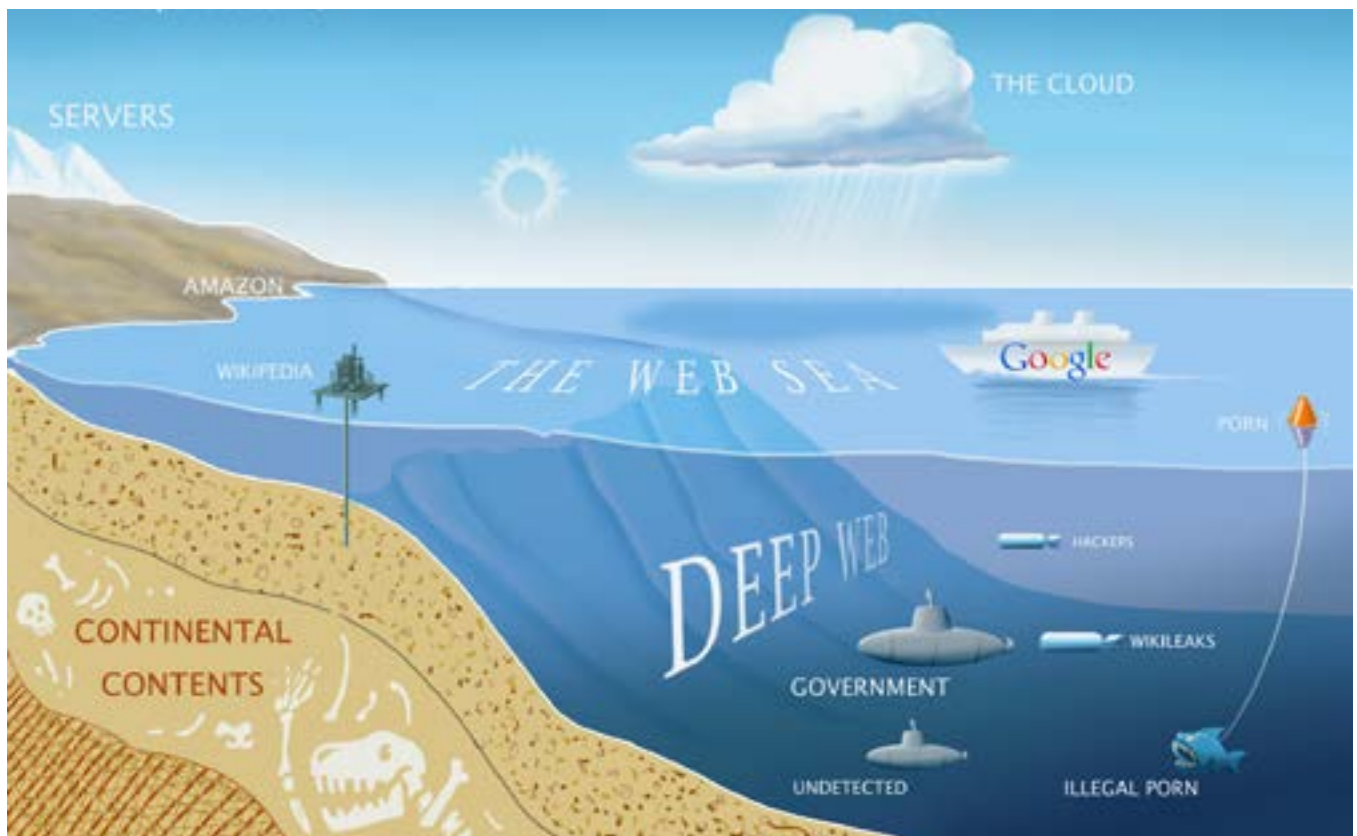
Content waar niet naar gelinkt wordt en non-tekstuele content zijn voorbeelden van content die zich op het onzichtbare web bevinden. Zoekmachines kunnen deze niet indexeren omdat er simpelweg geen manier is om er te geraken of omdat er niet kan worden vastgesteld welke informatie de content bevat. Content die niet in HTML is opgemaakt bevindt zich dan meestal ook op het onzichtbare web. Daarnaast kunnen webmasters door middel van robots.txt aangeven dat bepaalde pagina's niet geïndexeerd

mogen worden. Of een zoekmachine zich daaraan houdt is natuurlijk een tweede.

Dat is uiteraard niet alles wat er te vinden is op het onzichtbare web. Een groot deel van de data bevindt zich namelijk in databases. Deze databases zijn via het web toegankelijk door gebruik te maken van speciaal hiervoor ingerichte pagina's waarop query's uitgevoerd kunnen worden. Zoekmachines voeren deze query's echter vaak niet uit, en zo blijft de data onontsloten achter op het onzichtbare web. Het onzichtbare web komt hier echter wel in aanraking met het zichtbare web, want de speciale zoekpagina's worden wel door zoekmachines gevonden. Het gaat hier uiteraard niet om de databases die de content leveren voor dynamische sites, aangezien deze inhoud normaliter met HTML opgemaakt wordt op het moment dat een pagina opgevraagd wordt.

Ontwikkelingen in zoekmachines

De grote spelers op de zoekmachine-markt zijn er vanuit hun bedrijfs perspectief uiteraard op uit om zo veel mogelijk informatie op het web te indexeren, om zo hun gebruikers beter van dienst te kunnen zijn. Dit is dan ook de voornaamste reden dat er voort-



Figuur 1: Een impressie van 'The Deep Web' © brandpowder.com

durend onderzoek naar verbetering van zoekmachines heeft plaatsgevonden om het onzichtbare web (beter) te kunnen indexereren.

Een voorbeeld hiervan is het Sitemap protocol dat een webmaster in staat stelt om aan een zoekmachine aan te geven welke pagina's er op zijn site beschikbaar zijn. Een webcrawler kan gebruik maken van de lijst van URLs om op intelligente wijze, de normaal gesproken niet-indexeerbare site, toch te kunnen indexereren. Mod_oai is een Apache module die het OAI-PMH protocol implementeert. Met mod_aoi kunnen webcrawlers snel vaststellen of er nieuwe of gewijzigde pagina's op een website aanwezig zijn door de metadata van content te bekijken.

Tegenwoordig gebruikt men en masse social media en blogs om met elkaar te communiceren of een bedrijf te profileren. Dit zorgt voor een gigantische hoeveelheid data die veel bruikbare informatie kan bevatten. Dit is de reden dat Google PubSubHubbub (PuSH) ontwikkelde. PuSH is een syndicatiesysteem dat gebaseerd is op Atom. In tegenstelling tot normale Atom- of RSS-feeds wordt er echter niet periodiek gepolld of er nieuwe berichten zijn; deze

worden vergeleken. Dit stelt informatiezoekers in staat om een archief te verkennen. Dit kan van pas komen op het moment dat een informatiezoeker niet precies weet welke en hoeveel informatie er over een bepaald onderwerp beschikbaar is in een archief. Ook kunnen verschillende termen voor eenzelfde voorwerp vergeleken worden, wat van pas komt als beheerders van een archief een ander jargon gebruiken dan informatiezoekers.

Het web van het hier en nu

Recente ontwikkelingen hebben ervoor gezorgd dat mensen steeds vaker op zoek zijn naar informatie over zaken die op dit moment plaatsvinden in plaats van statische content op webpagina's. Een groot deel hiervan is 'sociale informatie', zoals die aanwezig is op sociale netwerken als Twitter en Facebook. Zoekmachines hebben reeds geprobeerd om deze zeer dynamische data, met wisselend succes, te indexeren. Zo kondigde Google in 2009 aan dat ze in zee gingen met Twitter om tweets in de zoekresultaten te kunnen gebruiken. Deze samenwerking werd echter vorig jaar verbroken, waardoor Google nu minder goed de real-time informatie van Twitter kan indexeren.

overzichtelijke graaf.

Conclusie

Het web biedt ons een bijna onuitputtelijke bron van informatie over eigenlijk alle thema's die er te bedenken zijn. Zoekmachines helpen ons om de informatie te vinden die we nodig hebben. Om ons goed van dienst te kunnen zijn, moeten zoekmachines grote delen van het web indexereren. Er bestaat echter een deel van het web dat om technische redenen of bepaalde voorkeuren niet geïndexeerd wordt. Dit wordt het onzichtbare web genoemd, omdat de informatie die daar beschikbaar is, niet te vinden is door standaard zoekmachines. De afgelopen jaren zijn er ontwikkelingen geweest om zoekmachines in staat te stellen delen van het onzichtbare web te ontsluiten. Er is bijvoorbeeld vooruitgang geboekt bij het indexereren van real-time informatie. Ondanks dat er al veel verbeteringen aan zoekmachines zijn doorgevoerd, en deze steeds beter in staat zijn om ons de informatie voor te schotelen die we zoeken, en ons zelfs real-time resultaten opleveren, is nog niet alle informatie op het onzichtbare web ontsloten. Zo zijn er nog steeds databases die niet door zoekmachines geïndexeerd worden en kunnen afbeeldingen, geluid en video's vaak nog niet optimaal geïndexeerd worden. Kijk dus bij je volgende Google-query langer dan je neus lang is, en ga daar waar geen zoekmachine geweest is.



worden namelijk gepusht. Op deze manier kunnen nieuwe blogs en statusupdates direct gepusht worden naar zoekmachines die zich geabonneerd hebben op de feed, wat hen in staat stelt om de nieuwe content te indexereren.

Een ander voorbeeld van een ontwikkeling op zoekmachinegebied is het onderzoek dat Marc Bron, PhD student bij Intelligent Systems Lab Amsterdam (ISLA), heeft uitgevoerd. Samen met collega's van het ISLA, het Centre for Television in Transition van de Universiteit Utrecht en het Nederlands Instituut voor Beeld en Geluid ontwikkelde hij een nieuwe interface waarmee zoekopdrachten in archieven kunnen

Een veelbelovende start-up op dit gebied is Bottlenose. Bottlenose is naar eigen zeggen geen zoekmachine, maar een 'discovery engine', die in real-time de social media datastroom scant en gebruikers in staat stelt om direct te weten te komen wat er op het moment speelt of belangrijk is. StreamOS is het platform dat aan Bottlenose ten grondslag ligt. Dit platform gebruikt verschillende technieken en algoritmen om sociale netwerken in de gaten te houden, te analyseren en trends te detecteren en maakt hierbij onder andere gebruik van natural-language processing en semantische classificatie. Een mooie feature is 'Sonar', die de huidige zoekopdracht aan andere trending topics relateert in een

Bronnen

Dark Web - Exploring, Data Mining the Dark Side of the Web (Springer, 2012)

H. Chen, et. al.

The Invisible Web - Uncovering Information Sources Search Engines Can't See (Information Today, 2001)
C. Sherman, G. Price

A Subjunctive Exploratory Search Interface to Support Media Studies Researchers (SIGIR '12)
M. Bron, et. al.

Van Rom

Curricula en gastronomie



Door: Rom Langerak
Opleidingsdirecteur Informatica

Nu de U-raad eindelijk overstag is kan een begin gemaakt worden met het uitrollen van het Twents Onderwijsmodel of TOM, voorheen ook NOM (Nieuw Onderwijsmodel) geheten.

Een van de issues daarbij is: wat moet er precies in het curriculum zitten?

Voor de bachelor Technische Informatica hebben we nu ongeveer zeven modules (van elk 15 EC) gepland die de kern van het curriculum bevatten. Er is van elke module een abstracte beschrijving, maar welke onderwerpen er precies behandeld gaan worden zal bekeken moeten worden door ontwerpteams onder leiding van zogeheten modulecoördinatoren. Overigens worden zes van de zeven modules gedeeld met andere opleidingen (BIT, EE en TW). Met de eerste vier modules, dus de modules van het eerste jaar, gaan de ontwerpteams binnenkort aan de slag. Aangezien het nieuwe onderwijs van start moet gaan in september 2013, zullen de roosters en vakbeschrijvingen rond februari 2013 af moeten zijn – dat betekent dat de ontwerpteams krap een half jaar hebben! En in dat half jaar moeten ook nieuwe onderwijsvormen en manieren van toetsen ontwikkeld worden.... Al met al een gigantische ontwerpklus, spannend en opwindend, en ook een klein beetje griezelig – gaat dat allemaal lukken?

De vraag naar wat er precies in het curriculum moet, kan niet eenvoudig beantwoord worden door naar het huidige curriculum te kijken. Eén van de pijlers van het Twentse model is namelijk dat we gaan reduceren in vakinhoud. In het Old School universitaire onderwijs was het doel om maar zoveel mogelijk kennis in de hoofden van de studen-

ten te gieten, en dat ging wel eens ten koste van vaardigheden als academisch denken en kritisch redeneren. In het internettijdperk staat veel informatie op het web, en is het veel belangrijker om te leren met die informatie om te gaan. Gesprekken met alumni bevestigen dit beeld: alumni kijken soms geamuseerd terug op de hoeveelheid detailonderwerpen die ze tijdens hun studie tot zich moesten nemen, en die ze na het verlaten van de UT-poort nooit meer gebruikt hebben. En tegelijkertijd geven ze aan dat de opleiding meer aandacht zou moeten geven aan hoe een academisch informaticus dient te denken en te werken: met meer aandacht voor het “waarom” dan het “hoe”, en meer gericht op het verkrijgen van de brede blik die vaak met “helicoptervisie” wordt omschreven (een slechte term als je weet dat veel ongelukken gebeuren doordat helicopterpiloten onvoldoende om zich heen kijken, gedwongen als ze zijn om naar hun instrumentenpaneel te turen).

Ook referentiecurricula zoals het nieuwe ACM 2013 curriculum, kunnen maar in beperkte mate helpen bij wat er wel en niet in het curriculum moet. Het ACM 2013 is zeer gedetailleerd en bevat zo’n beetje elk denkbaar onderwerp uit de informatica – niet direct een richtsnoer om met beleid te gaan snoeien. Het kan wel goede diensten bewijzen bij het documenteren van onze keuzes: wat doen we wel, wat doen we niet? Maar welke keuzes, daar moeten we toch echt zelf zien uit te komen. Wat er precies in een curriculum moet, blijft dus een lastige kwestie!

Peter Sloterdijk, de filosoof en tevens rector van de Karlsruhe University of Arts and Design, geeft in zijn duizelingwekkende boek “Du musst dein Leben

ändern” (in het Nederlands verschenen onder de titel “Je moet je leven veranderen”) een opzetje voor een zogenaamde Algemene Disciplineleer, die zich onder andere zou moeten bezighouden met de vraag welke disciplines een universitair curriculum zou moeten omvatten. Hij geeft een opsomming van zo’n twaalf disciplines, een opsomming die naast te verwachten zaken als techniek, administratie en geneeskunde ook verrassingen bevat als meditatie, ritualistiek, “Sexualpraxiskunde” (wat dat ook mag inhouden), en.... gastronomie! Zou onze eigen rector niet voor dat laatste te porren zijn?

Sinds april 1992 is dr. ir. Rom Langerak universitair docent bij de Formal Methods and Tools groep van de faculteit EWI. Romanus (Rom) werd op 1 februari geboren in Dordrecht en ging naar het Christelijk Lyceum aldaar. Hij haalde op de Universiteit Twente met lof zijn studie Toegepaste Wiskunde, waar hij afstudeerde op een onderwerp over Databases. Het is dan ook niet vreemd dat hij na zijn afstuderen ging promoveren bij de toenmalige faculteit Informatica. Na zijn promoveren in 1992 bleef hij bij de faculteit werkzaam.

Rom houdt van literatuur, filosofie, gitaar spelen, biljarten en Taekwondo. Sinds september 2009 is hij de opleidingsdirecteur Informatica, een taak die hij met liefde zal gaan uitvoeren om zo het onderwijs voor zowel studenten als docenten

Marvin Minsky

een van de grondleggers van Artificial Intelligence



Door: Niek Tax
Redacteur I/O Vivat

Sinds 1966 rijkt de Association for Computer Machinery ieder jaar de Turing Award uit aan iemand die substantiële bijdragen van technische aard heeft geleverd aan de IT-gemeenschap. In dit vaste katern “Turing-award winnaar in beeld” zal elke editie van de I/O Vivat een andere winnaar van de Turing Award worden uitgelicht en zal een blik op zijn of haar leven en bijdragen aan de computerwetenschap worden geworpen. Ditmaal zal in dit katern Marvin Minsky, geroemd om zijn ontwikkeling aan neurale netwerken en bijdragen aan het vakgebied Artificial Intelligence, centraal staan.

Marvin Minsky

Op 9 augustus 1927 kwam Marvin Minsky ter aarde in de stad New York. In deze stad groeide Minsky ook op en ging hij naar de prestigieuze Bronx High School of Science, waar hij zijn middelbare schoolperiode tegelijkertijd doorliep met latere winnaars van de Nobelprijs in de Natuurkunde Melvin Schwartz en Leon N. Cooper (waarnaar Sheldon Cooper uit The Big Bang Theory is vernoemd). Na zijn middelbare school diende hij in de Amerikaanse marine gedurende het laatste jaar van de tweede wereldoorlog.

Na de oorlog begon Minsky zijn studie aan Harvard. Aanvankelijk volgde hij vakken in werktuigbouwkunde, wiskunde, neuroanatomie, neurofysiologie

en psychologie. Uiteindelijk behaalde Minsky in 1950 zijn Bachelor of Arts in de wiskunde. Op Harvard werkte Minsky voor het eerst concepten uit voor een machine die zelfstandig kan leren. Na zijn bachelor haalt Minsky zijn master en PhD-graad in de wiskunde aan Princeton. Gedurende zijn tijd aan de Princeton universiteit bouwt Minsky het eerste hardwarematige neurale netwerk, genaamd de SNARC (Stochastic Neural Analog Reinforcement Calculator). SNARC was in staat zich zelfstandig aan te leren door een doolhof te navigeren.

In 1958 gaat Minsky als onderzoeker in dienst bij het Massachusetts Institute of Technology, waar hij een jaar later samen met collega-onderzoeker en latere Turing-award winnaar John McCarthy de MIT Computer Science and Artificial Intelligence Laboratory opricht. Dit onderzoeksinstituut zal uiteindelijk uitgroeien tot een van de meest vooraanstaande onderzoeksinstituten in het Artificial Intelligence onderzoeksgebied in de wereld.

Perceptrons en controverse



Figuur 1: Marvin Minsky

In 1969 publiceert Minsky, samen met collega-onderzoeker Seymour Papert, het boek *Perceptrons: an introduction to computational geometry*. In dit boek beschouwen Papert en Minsky perceptrons, een soort neurale netwerk die met name in de jaren '60 een erg grote rol speelde binnen het vakgebied van machine learning.

In het boek *perceptrons* formuleren Papert en Minsky enkele wiskundige bewijzen omtrent perceptrons, maar zijn daarnaast ook erg kritisch over de mogelijkheden van de perceptron als machine learning techniek. Papert en Minsky's kritiek op perceptrons zat hem in het feit dat het volgens hen onmogelijk was om een logische XOR-operator in een artificieel neuron te implementeren, waarvan grotere perceptron-netwerken ernstig geschaad zouden worden in hun prestaties. Deze kritiek op perceptrons had dermate veel impact op het onderzoeksgebied van neurale netwerken dat sommigen het als directe aanleiding zien voor een recessie in het Artificial Intelligence gebied waarin interesse in en onderzoeksbudgetten voor dit onderzoeksgebied

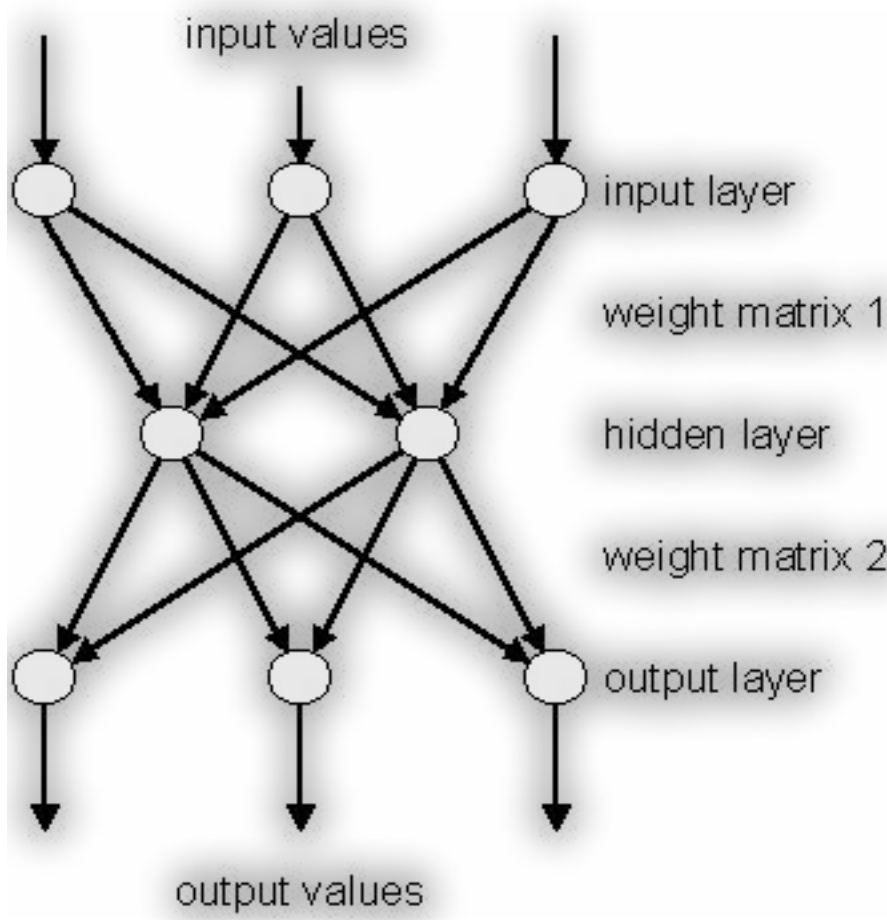
sterk afnamen. Deze recessie, ook wel de AI winter genoemd, hield gedurende de gehele jaren '70 aan. Omtrent Papert en Minsky's kritiek op perceptrons heerste vanwege haar implicaties veel controverse. In de jaren '80 bleek de goede kritiek op perceptrons ook zeer ongelukkig, toen nieuw onderzoek aan het licht bracht dat een logische XOR-operator wel in een artificieel neuron geïmplementeerd kon worden.

Cognitie

Marvin Minsky heeft naast zijn werk in de Artificial Intelligence ook werk geleverd in het aanverwante vakgebied van de cognitiewetenschappen, het wetenschapsgebied wat zich bezig houdt met de aard van alle cognitieve functies van de mens. Het voornaamste werk van Minsky binnen de cognitiewetenschap is de *Society of the Mind* theorie, waarin Minsky stelt dat menselijke intelligentie niet is toe te schrijven aan één enkel proces maar daarentegen een netwerk van communicerende individuele agents aan zich ten grondslag heeft liggen. Minsky onderbouwt deze theorie met het argument dat de verschillende

taken van het menselijk denkproces zo verschillen dat ze wel fundamenteel andere mechanismen moeten vereisen. Minsky ontwikkelde deze ideeën door inzichten uit de ontwikkelings-kindpsychologie te combineren met zijn ervaringen op het gebied van Artificial Intelligence.

Recent, in 2006, publiceerde Minsky een vervolg op de *Society of the Mind* theorie in zijn boek *The Emotion Machine*. In *The Emotion Machine* stelt Minsky dat menselijke emoties niets anders zijn dan manieren van het brein om na te denken over verschillende typen problemen in de wereld. Het brein heeft, zo stelt Minsky, regel-gebaseerde mechanismen die emoties aansturen om met verscheidene problemen om te gaan. Minsky stelt dat het menselijk denkproces opgesplitst kan worden in reeksen van specifieke acties en betoogd dat wanneer we deze onderliggende acties begrijpen in staat zullen zijn om machines te bouwen die deze onderliggende acties kunnen nabootsen. Hoewel zijn ideeën zeker niet onomstreden zijn, gelooft Minsky erin dat het op deze manier mogelijk is om emotionele machines te bouwen – net zoals de mens.



Figuur 2: Schematische weergave van een perceptron

ACM

De Turing Award wordt jaarlijks uitgereikt door de Association for Computer Machinery (ACM). De ACM is de grootste en meest prestigieuze wetenschappelijke IT gemeenschap, waar meer dan 100,000 IT onderzoekers bij aangesloten zijn. De ACM is onderverdeeld in vijftig Special Interest Groups (SIGs) die elk evenementen voor een deelgebied van de informatica organiseren.

Bronnen

Academic Biography of Marvin Minsky (2011)
MIT Library

In Honor of Marvin Minsky's Contributions on his 80th Birthday (2007)
AI Magazine Volume 28, number 4



Jouw Artikel hier?

Neem contact op met de
I/O Vivat-redactie:

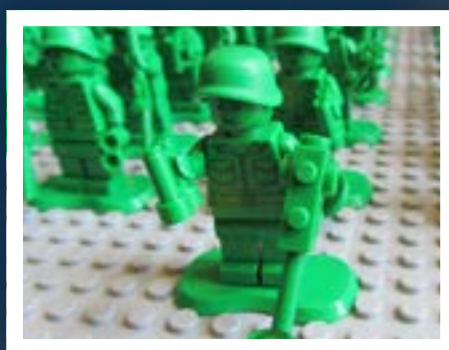
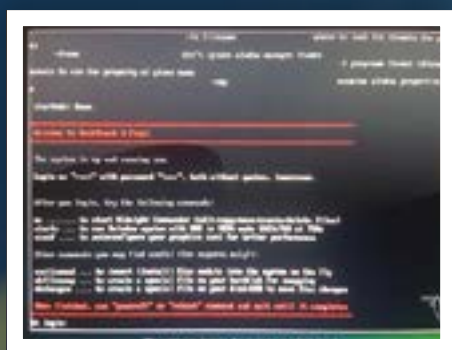
vivat@inter-actief.net





VOLGENDE KEER IN I/O VIVAT

- Anatomie van een Pentest
- Cyberreservistenleger
- IT in Azië



Advertentie

TECHNO-

ADV-A4-CS.

pdf