

## DRIE SCENARIO'S VOOR HET 'VOLGENDE' INTERNET

# Derde- generatie web

Na de web2.0-hype met zijn social networks, tagclouds, AJAX-sites en user generated content, gonst het nu op internet over het 'intelligente' semantische web. Ofwel: web 3.0, de Machine of het zelfdenkend web. Slim, slimmer, slimst?

**G**ewend aan web 2.0? Precies op tijd, want hoewel niemand precies weet waar het voor staat, klopt web 3.0 al op de deur. Vooralsnog lijkt het de grote vergaarbak voor alles wat nieuw is op het web. De aanjager van wéér een nieuwe versie van het world wide web is het 'slimme' semantische web. Maar er gebeurt nog veel meer, zoals apparaten die met internet gaan communiceren, waardoor alles tot één grote Machine samensmelt. En die zou ook nog wel eens superintelligent kunnen zijn. Waar gaat het heen met het web? Emerce zet drie scenario's op een rij

## Scenario 1: Semantisch web

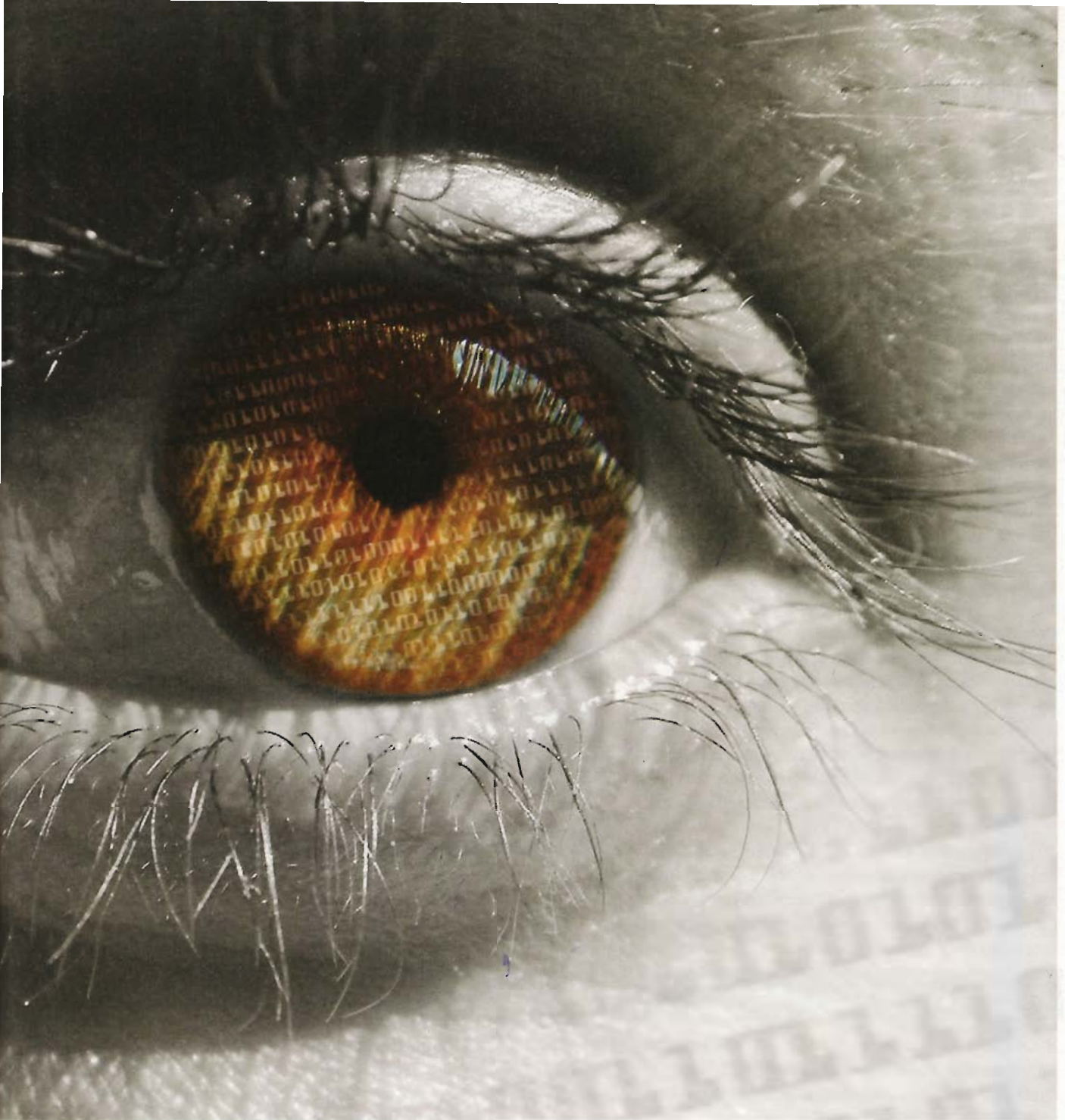
Werken met slimme data en daardoor alles kunnen vinden en betere zoekresultaten boeken. Dat is in een notendop de essentie van het semantische web. Kern van de semantische webdroom: surfende webagenten of 'widgets' die automatisch problemen analyseren en vragen beantwoorden.

Tim Berners-Lee, de grondlegger van het World Wide Web (zie kader), schrijft in 2001 een inmiddels beroemd artikel in Scientific American: 'The Semantic Web'. Het gaat over het dusdanig organiseren van webcontent dat het web data zélf kan verwerken. In het artikel beschrijft hij de volgende case. Pete wordt gebeld door zijn zus Lucy met de mededeling dat



hun moeder meerdere therapeutische sessies van een specialist nodig heeft. Lucy instrueert haar semantische webagent, die informatie ophaalt over de voorgeschreven therapie, een lijst specialisten bekijkt die binnen moeders verzekering vallen, binnen 20 kilometer afstand van de woonadressen van Pete en Lucy praktijk houden en bovendien goed staan aangeschreven. Vervolgens vergelijkt de agent de beschikbare spreekuren met de elektronische agenda's van broer en zus, in verband met het wegbrengen van mama. De agent komt met een plan dat Pete niet bevalt, hij zou door de spits naar de andere kant van de stad moeten rijden. Met een striktere opdracht voor 'locatie' en 'tijd' komt het semantische web vervolgens met een goed voorstel. Ten slotte leggen Pete en Lucy de afspraken vast.

In dit voorbeeld gaat het semantische web zelf aan de slag met data. Dit is mogelijk omdat de webcontent niet



alleen gepresenteerd is om door mensen gelezen te worden, maar ook om computerprogramma's er zelf betekenis aan te kunnen laten geven. Het semantische web weet dat een specialist in de kliniek werkt op bepaalde dagen, dat er dagen in

## Het semantische web is een 'world wide database'

een jaar zitten, dat er op die dagen afspraken zijn te maken, et cetera. Bovendien moet zo'n systeem bijvoorbeeld wel iets ogenschijnlijk eenvoudigs als post- en woonadres kunnen onderscheiden. En weten dat een plaatsnaam iets anders is dan een achternaam, voor het geval deze overeenkomen (zoals bij mevr. van Putten).

Berners-Lee schrijft, samen met twee collega's, verder:

"Tot op heden heeft het web zich ontwikkeld als een medium van documenten voor mensen, meer dan een medium voor data en informatie die automatisch verwerkt kunnen worden." Het semantische web gaat dus intelligenter met kennis om, het is een 'world wide database'. Eén van de basisingrediënten van het semantische web is een ontologie, een systeem voor kennisrepresentatie.

Een eenvoudige definitie van kennis is: informatie plus context. Toegevoegde context aan informatie of data wordt ook wel metadata genoemd. En met deze metadata worden data slimmer, krijgen ze betekenis (semantiek) voor machines. De grootste uitdaging voor het semantische web is momenteel het formuleren van ontologieën die uitwisselbaar zijn tussen systemen (zie kader 'Ontologie en owl').

De semantische webgemeenschap groeit met de dag. Op Google scoort 'semantic web' inmiddels een kleine twintig

miljoen hits. Voor zover dit een goede graadmeter is, duidt het in ieder geval op een enorme interesse voor het onderwerp. Nova Spivack, internetondernemer en semantisch-webvisionair, plaatst in zijn essay 'Minding the Planet' (november 2006) het nieuwe web in een historisch perspectief: "De drukpers en het web veranderden de economie van het reproduceren, distribueren en het toegankelijk maken van kennis. Het semantische web verandert de economie van informatieverwerking." Spivack voorspelt dat binnen twee jaar semantisch-webproducten en -diensten de massamarkt bereiken. Binnen vijf jaar zijn er een paar 'killer apps'. Welke dat zouden kunnen zijn, verklapt hij helaas niet.

Wat is er concreet semantisch op het internet? Voorlopig weinig. Het bekendst is Ontoworld.org, het semantische zusje van Wikipedia, dat op 15 juli 2006 startte.

Ontoworld.org maakt gebruik van Semantic MediaWiki (SMW), een wiki-systeem waarmee gebruikers semantische annotaties (gestructureerde data) aan wiki-teksten kunnen toevoegen. Bij 'Semantic MediaWiki' staat: "Hoewel het (annoteren) op het eerste gezicht dingen meer complex lijkt te maken, kan het enorm helpen de structuur van de wiki te vereenvoudigen, de gebruikers meer informatie te laten vinden in minder tijd en de algehele kwaliteit en consistentie van de wiki te verbeteren."

Op commerciële basis zijn er inmiddels meerdere bedrijven die ontologieën en browsers voor het semantische web ontwerpen en deze verkopen aan bijvoorbeeld de medische industrie. Via de website van het W3C is een aantal semantische demo's en portals te vinden.

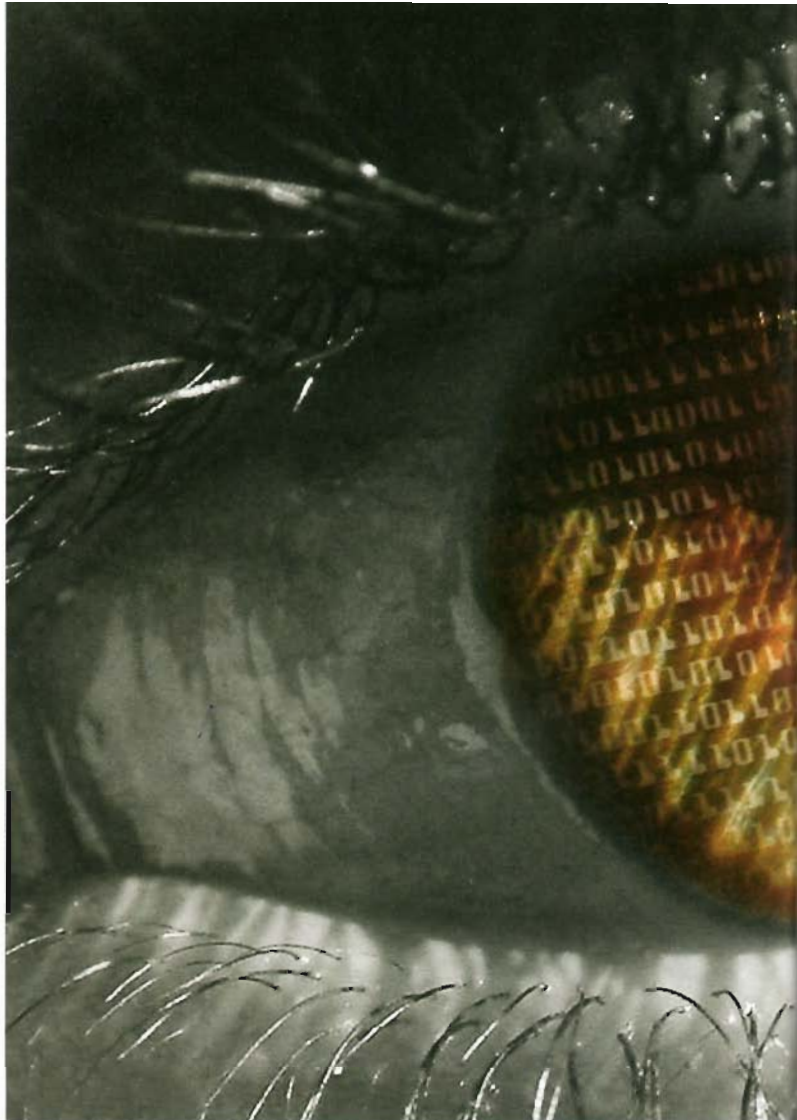
### Scenario 2: Web 3.0

Het semantische web wordt ook vaak web 3.0 genoemd. De Britse krant The Guardian berichtte bijvoorbeeld in maart dat de BBC en IBM een 'web 3.0'-deal hebben gesloten.

#### Tim Berners-Lee

Tim Berners-Lee (1955) is directeur van het World Wide Web Consortium (W3C) dat hij zelf oprichtte in 1994. Hij is ook onderzoeker bij het Computer Science and Artificial Intelligence Laboratory van het Massachusetts Institute of Technology (MIT) en professor aan de faculteit Department of Electronics and Computer Science van de Southampton University. Hij wordt gezien als de grondlegger van het World Wide Web.

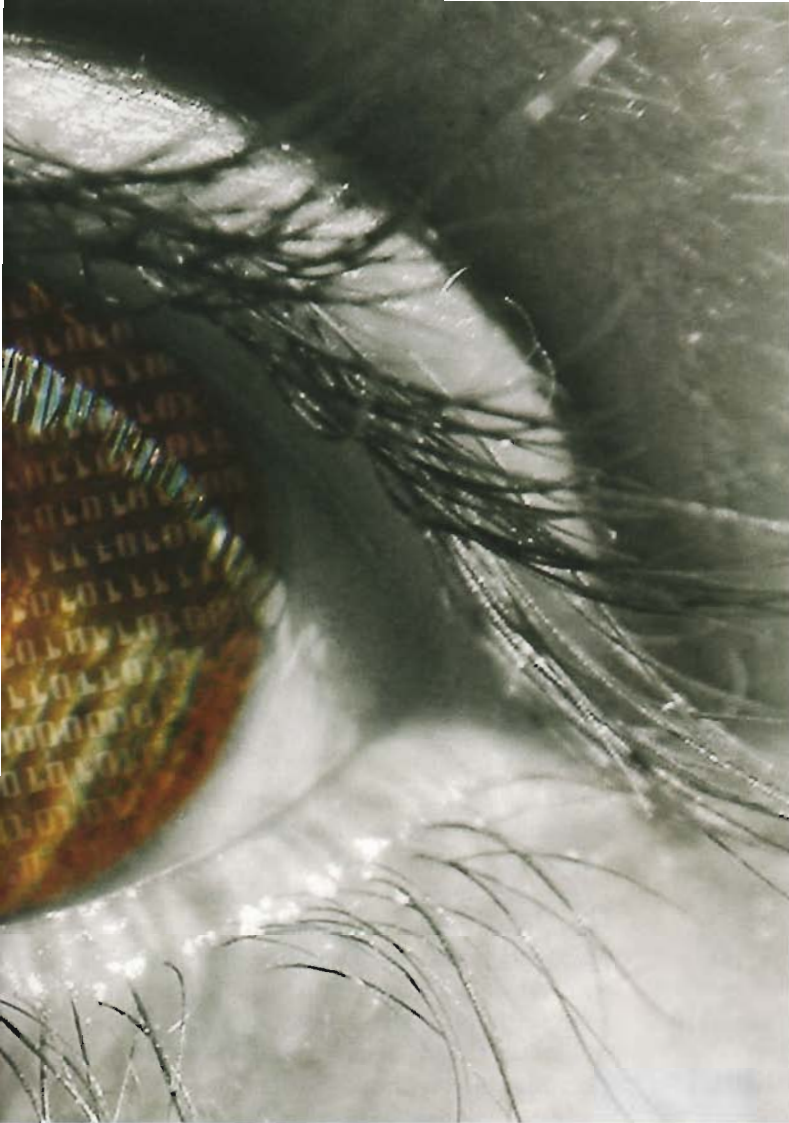
Heeft Tim Berners-Lee het internet uitgevonden? Op de website van W3C antwoordt hij op deze vraag: "Nee, nee, nee! Ik verbond alleen de ideeën over hypertext met tcp en dns en - tadaah! - het World Wide Web." In de jaren zestig van de vorige eeuw had Ted Nelson al de mogelijkheden van hypertextlinks uitgewerkt. En in de jaren tachtig stuurden mensen al e-mails met ip (internetprotocol), tcp (transfer control protocol) en dns (domain name system). In 1980 diende Berners-Lee bij CERN, de Europese organisatie voor nucleair onderzoek in Genève, een voorstel in voor informatiemanagement op basis van hypertext. Naar eigen zeggen verbond hij uiteindelijk in 1989 dus verschillende ideeën en bestaande technieken tot het World Wide Web. Ook maakt hij de eerste internetbrowser, WorldWideWeb, en zette hij in 1991 de eerste website online. Momenteel is hij een actieve ontwikkelaar en promotor van het semantische web.



Volgens Ashley Highfield, BBC-directeur van de afdeling 'Future Media and Technology', begrijpt web3.0- of semantisch-webtechnologie de waarde van content. Alle huidige videozoektechnologie werkt alleen met metatags of termen toegevoegd aan de video's. Dat werkt prima, zegt Highfield, als alle content voorzien is van metatags. Maar de BBC wil haar archief met meer dan 1,4 miljoen uur televisie en audio digitaliseren en doorzoekbaar maken. En handmatig alles omzetten en 'taggen' is onbegonnen werk. Om het inventariseren van al die content te automatiseren, gaat de BBC gebruikmaken van het systeem Marvel van IBM, dat afbeeldingen en video's kan doorzoeken en indexereren. Wel zegt Highfield dat het nog de grote vraag is of het allemaal gaat werken.

Dit voorbeeld is exemplarisch voor het op één hoop vegen van het semantische web en web 3.0. Het verbaast ook niet dat de derde editie van de conferentie Les Blogs, LeWeb3 in december 2006 in Parijs, op internet vaak een web3.0-event heet, terwijl het een conferentie over web 2.0 is. Of heeft de organisatie dat bewust gedaan?

Verder is opvallend dat in blogs vaak commentaar te lezen is op het gebruik van de term web 3.0. Mensen zijn net gewend aan web 2.0 en nu dient de opvolger zich al aan. Het lijkt of er een bepaalde 'webX.0-moeheid' ontstaat. Is het wel zinnig om in versies van het web te spreken? Wat zou een eventueel web 3.0 onderscheidend maken van web 2.0 en web 1.0? Is web 1.0 de eerste generatie van commercieel



internet met weinig interactiviteit? Is web 2.0 dan de interactieve en sociale versie met blogs, wiki's, podcasting, social communities, p2p-netwerken en sites als Digg, Technorati, Google Maps, Del.icio.us, Flickr, Wikipedia? En web 3.0 dan? Daarover doen uiteenlopende opvattingen de ronde. In een veelgelezen artikel in de New York Times (november 2006) beschrijft John Markoff web 3.0 als een verzameling nieuwe technologieën die efficiënte nieuwe manieren bieden om computers te helpen conclusies te trekken uit online data. Dit is min of meer een beknopte beschrijving van het semantische web.

Nova Spivack geeft in 'The Third-Generation Web is Coming' een meer uitgebreide definitie van web 3.0, waarin het semantische web één van meerdere kenmerkende eigenschappen is. Hij noemt onder andere de volgende eigenschappen van web 3.0: open dataformats, meer verbondenheid tussen web en andere apparaten, mobiel toegang tot internet, draagbare identiteit en persoonlijke data, meer intelligent met semantische webtechnologieën, verspreide databases, verwerking van natuurlijke taal en lerende en denkende machines.

Het lijkt er sterk op dat hier alles wat nieuw is of niet past in bijvoorbeeld de categorie web 2.0, bij elkaar wordt geraapt en een nieuwe naam krijgt. Web 3.0 als containerbegrip. Maar ontstaat hiermee een consistent of helder beeld? Net als in het begin bij web 2.0, inmiddels een redelijk ingeburgerd begrip, zal over de exacte inhoud en betekenis van

## Het lijkt of er een bepaalde 'webX.0-moeheid' ontstaat

### Ontologie en owl

Voor een systeem van verrijkte data zijn verzamelingen definities nodig die begrippen en de relaties daartussen verklaren. Daarom staan ontologieën aan de basis van het semantische web, het web moet immers begrijpen welke termen of data waar naar verwijzen. Op Wikipedia staat 'ontologie (informatica)' als volgt beschreven: "In de informatica is een ontologie het product van een poging een uitputtend en strikt conceptueel schema te formuleren over een bepaald domein. Een ontologie is typisch een hiërarchische datastructuur die alle relevante entiteiten en hun onderlinge relaties en regels binnen dat domein bevat, zoals bij een domeinontologie het geval is. Het gebruik van het woord binnen de informatica is afgeleid van het veel oudere gebruik van het woord ontologie binnen de filosofie." Binnen de filosofie betekent ontologie zijnsleer, dat het geheel der dingen of de zijnden beschrijft.

Om websites semantisch te organiseren is onder andere owl (web ontology language) ontworpen. Door het extra vocabulaire gecombineerd met formele semantiek (gedefinieerde relaties tussen de termen) biedt owl aan machines of toepassingen meer interpretatiemogelijkheden van webcontent dan de semantische technologieën xml (extensible markup language), rdf (resource description framework) en rdf-schema (rdf-s). Owl, sinds 2004 formeel aanbevolen door het World Wide Web Consortium, wordt gezien als dé technologie of standaardtaal voor het semantische web.

Het semantische web kent vele critici en sceptici die menen dat breed bruikbare ontologieën niet, of zeer moeilijk te maken zijn. Het commentaar van een Emerce-lezer op een nieuwsbericht over het semantische web is hiervoor illustratief: "De opvatting dat er zoiets bestaat als een universele ontologie getuigt van een zeer simplistische visie op het fenomeen menselijke taal en de betekenis ervan.

Elke poging om zo'n ontologie te vormen is per definitie gedoemd te mislukken omdat het resultaat aantoonbaar arbitrair en ambigu is; arbitrair omdat het woord 'hart' voor een slager een andere betekenis zal hebben dan voor een chirurg; ambigu omdat taal per definitie ambigu is. Neem bijvoorbeeld de zin 'time flies like an arrow' die voor mensen maar één interpretatie heeft ('de tijd vliegt') maar voor computers ten minste twee betekenissen heeft (waaronder 'de tijdvliegies houden van een pijl')."

Toch zijn ontwikkelaars van semantische websites positief en wijzen ze op de vooruitgang die wordt geboekt en op werkende toepassingen op beperkte domeinen. Het integreren van verschillende ontologieën in één overkoepelende ontologie, ofwel de integratie van alle semantische websites in één wereldwijd semantisch superweb, is echter nog een fors aantal stappen verder.

[www.w3.org/TR/owl-features](http://www.w3.org/TR/owl-features)

web 3.0 ook nog wel het nodige gediscussieerd worden. Als de naam überhaupt al bestaansrecht krijgt... En vervolgens kunnen we wachten op web 4.0, 5.0 en verder.

### Scenario 3: De Machine

In het verlengde van het semantische web liggen ideeën over communicatie tussen web en apparaten. Bijvoorbeeld een magnetron die op de website van de fabrikant checkt hoe je iets optimaal ontdooit en dit vervolgens zo doet. Of chips in kleding met informatie over wasprogramma's die met zowel de wasmachine als het internet kunnen communiceren. En dat alles verbonden in één grote machine - van mensen tot data en apparaten. Dat is het favoriete thema van internetfilosoof Kevin Kelly, oprichter en oud-hoofdredacteur van het Amerikaanse magazine Wired en auteur van onder andere de bestseller *New Rules for the New Economy*. In het artikel 'We are the Web' zegt hij: "Het gebeurt maar één keer in de geschiedenis van elke planeet dat de bewoners voor de eerste keer alle ontelbare delen met elkaar verbinden tot één grote machine. Daarna kan die machine harder gaan draaien, maar het wordt slechts één keer geboren."

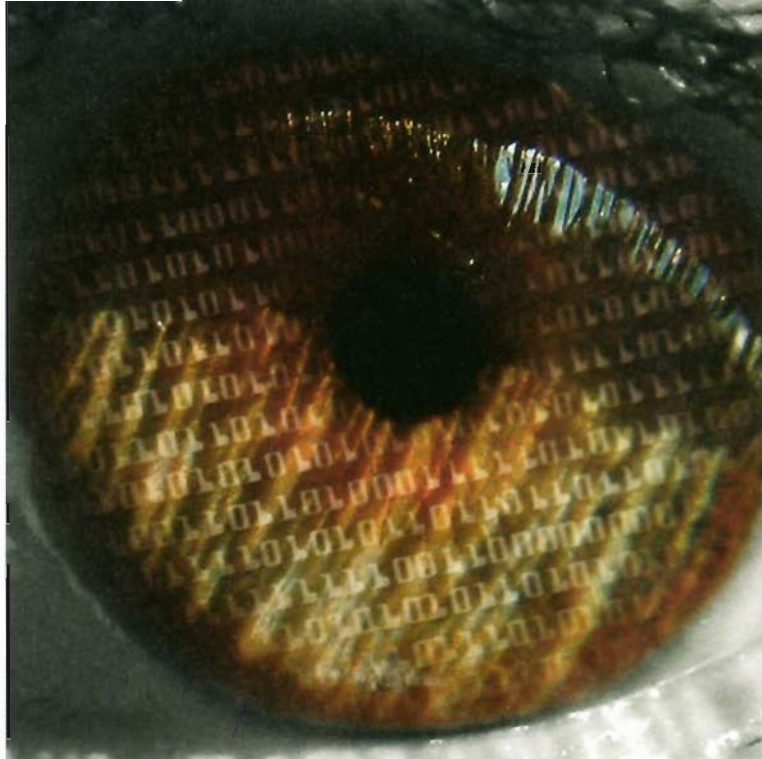
## Google is met kunstmatige intelligentie bezig, maar zegt niet wat dat inhoudt

Nog meer begripsverwarring kan optreden als er gesproken wordt over kunstmatige intelligentie (AI) in verband met het semantisch web of web 3.0. Google-oprichter Larry Page kondigde afgelopen februari tijdens een lezing voor Amerikaanse wetenschappers aan dat Google met AI bezig is, zonder erbij te zeggen wat dat precies inhoudt. Hij zei wel dat AI minder ver van ons afstaat dan mensen denken.

Het creëren van machines die autonoom kunnen 'denken', informatie analyseren en deze creatief synthetiseren of verwerken, is het domein van kunstmatige intelligentie. Maar wat is intelligent? Een rekenmachine heeft ook een bepaalde vorm van intelligentie. In die zin is het semantische web intelligent te noemen. Maar dit web heeft geen enkel benul van wat het analyseert of 'weet'. Menselijke intelligentie is van een andere orde. Ook al legt het semantische, of intelligente, web verbindingen en komt het terug met zoekresultaten waarvan mensen zeggen dat ze 'erg slim' zijn gevonden, het blijft een set toepassingen die 'domweg' doet wat mensen hebben geprogrammeerd.

Of toch niet? Sinds de begindagen van internet, en zelfs daarvoor, zijn er speculaties over de stap van machines met brute rekenkracht naar een intelligent denkend, creatief web. De lijst met visionairs of fantasten over dit onderwerp is lang, met onder meer Kevin Kelly, toekomstfilosoof en uitvinder Ray Kurzweil, sf-auteur William Gibson, robotonderzoeker Hans Moravec en Nova Spivack.

Het verst gaan theorieën over een 'zelfbewust' web. Deze gedachte komt vaak voort uit een analogie met het evolutio-



naire ontstaan van bewustzijn bij de mens door de miljarden verbindingen en communicatie tussen neuronen in de hersenen. De wereldwijd aan elkaar gekoppelde computers met een semantische of neurale organisatie van data, zullen op een gegeven moment de complexiteit, en daarmee naar verwachting de capaciteit, van de menselijke hersenen overtreffen. Dat zou over grofweg tien tot dertig jaar het geval kunnen zijn. Dergelijke ideeën zijn onder andere bekend als 'The Global Mind', of Het Superbrein. Kevin Kelly meent dat we het ontstaan hiervan misschien nooit zullen merken. Misschien is het zelfs al zover, want wie zegt dat dit superweb behoefte heeft te communiceren met mensen? Als we er überhaupt al iets van zouden begrijpen, want deze vorm intelligentie zou voor mensen wel eens geheel onbegrijpelijk kunnen zijn. ■

### Filmpjes over toekomst internet

Uitgelezen? Bekijk dan de volgende video's over de toekomst van het internet:

1. [boss.streamos.com/wmedia/energycommerce/030107.ti.hrgwww.wvx](http://boss.streamos.com/wmedia/energycommerce/030107.ti.hrgwww.wvx)  
Webcast van lezing door Tim Berners-Lee aan leden van de Energie- en Handelscommissie van het Amerikaanse Huis van Afgevaardigden over de toekomst van het World Wide Web. Hier spreekt hij onder andere over het semantische web (maart 2007).
2. [www.youtube.com/watch?v=6gmP4nk0EOE](http://www.youtube.com/watch?v=6gmP4nk0EOE)

Titel: 'Web 2.0 ... The Machine is Us/ing Us'. Onder andere geïnspireerd door een citaat van Kevin Kelly uit 'We are the Web': "De honderd miljard keren per dag dat mensen op webpagina's klikken, is te zien als een manier om de Machine te leren wat wij belangrijk vinden."

3. [www.youtube.com/watch?v=nsa5ZTRJQ5w](http://www.youtube.com/watch?v=nsa5ZTRJQ5w)

Hierin wordt web 2.0 de basis genoemd voor een semantische omgeving. De services en technologieën die samen web 2.0 vormen, zijn volgens de commentaarstem veel minder krachtig dan het semantische web waarbij machines betekenis kunnen 'begrijpen'.

### Websites

[ontoworld.org](http://ontoworld.org)

[www.semanticweb.org](http://www.semanticweb.org)

[www.w3.org/2001/sw/](http://www.w3.org/2001/sw/)

[semanticweb.startpagina.nl](http://semanticweb.startpagina.nl)