

**Claude E. Shannon (1916-2001),**  
*'de vader van de informatietheorie'*

# Alles wat u wilde weten over comprimeren (maar nooit durfde te vragen)

Inleiding tot de informatietheorie voor investeerders

# **Alles wat u wilde weten over comprimeren (maar nooit durfde te vragen)**

*Inleiding tot de informatietheorie voor investeerders*

## **Voorwoord**

In de zomer van 1999 hoorde ik op de radio het nieuws dat de uitvinder Jan Sloot was overleden. Het geheim van zijn geniale datacompressietechniek nam hij mee in zijn graf. Dit nieuws gaf direct een deuk in mijn Oranje-gevoel. Had Nederland eindelijk weer eens kunnen schitteren op het wereldtoneel met nota bene een fantastische IT-vinding, legt de uitvinder plotseling het loodje. Uit de informatie die in de maanden daarna in de media doorsijpelde kreeg ik steeds meer de indruk dat we hier echt een geweldige kans hadden gemist. Immers, met onze nationale high-tech celebrity Roel Pieper aan het roer van de firma die de vinding zou uitbaten, kon het niet anders of ons land was een echte player in de internationale ICT-wereld geworden.

Het is eind 1999 en de wereld is volledig in de ban van internet en IT. In Amerika worden dagelijks nieuwe internetbedrijven naar de beurs gebracht. Grote IT-legendes als Larry Ellison, Scott McNealy en Bill Gates zijn niet van het TV-scherm te branden. In Nederland is weliswaar net als in alle andere westerse landen de beursgekte uitgebroken, maar een echt aansprekende *high-tech* beursintroductie heeft er nog niet plaatsgevonden. Net bekomen van de deconfiture van het Veluwe softwarebedrijf Baan, is het de vooravond van beursintroducties als World Online en Newconomy. Veel beleggers hebben nog een wild ritje door de polder voor de boeg.

De impact van de claim om een speelfilm van negentig minuten op een chipcard te kunnen zetten was mij in 1999 al gelijk duidelijk. Het was in die tijd al een hele klus om een speelfilm met een beetje redelijke kwaliteit op een CD-ROM te krijgen. Laat staan te comprimeren tot slechts 1000 Bytes, grofweg een verkleiningsfactor van een miljoen. Als dat toch eens waar was!

Pas 5 jaar later lees ik in het boek *De Broncode* van Eric Smit het hele verhaal. Ik word heen en weer geslingerd tussen geloof en ongelooft. Zou het algoritme bestaan of niet? Wat zat er nou eigenlijk in dat geheimzinnige kastje? Is het fundamenteel eigenlijk wel mogelijk om eindeloos te comprimeren?

Als zoon van een radioamateur ben ik opgegroeid in een milieu van radiofanaten en elektronicafreaks. Ik heb met eigen ogen

gezien dat sommige van deze lieden blind een zender-ontvanger in elkaar konden solderen. Deze technische helden uit mijn jeugd bezaten een mengeling van een ongelooflijke hoeveelheid praktische kennis gecombineerd met een goede theoretische ondergrond. En waren in staat om letterlijk alles werkend aan elkaar te knopen of om te bouwen. Dat een begaafde TV-reparateur als Jan Sloot in staat was om een kastje te bouwen dat 16 speelfilms tegelijk afspeelde verbaasde mij dan ook niet. Maar had het iets te maken met compressie? Stond er werkelijk een *keycode* op de chipkaart die werd gedecodeerd tot een complete speelfilm, kortom was het geniaal of gewoon oplichting? In het boek van Eric Smit worden vele vragen beantwoord. Mijn twijfels of er theoretisch iets mis is met de claims van Jan Sloot bleven echter bestaan.

Gefascineerd door het verhaal over deze uitvinder begon ik te grasduinen op internet. Veel forums stonden vol met complottheorieën, beschuldigingen en onsamenhangende betogen. Veel wijzer werd ik er niet van.

Op de serieuze forums werd veel gespeculeerd, maar ook hier was het niveau niet om over naar huis te schrijven. Informatie uit het boek werd eindeloos herkauwd zonder nieuwe gezichtspunten toe te voegen. De enige manier om echt wijzer te worden was om in de studieboeken te duiken en me in de informatietheorie te verdiepen.

Aangezien er na 1999 in Amerika en Australië opnieuw “uitvinders” opdoken, welke geld ophaalden met vergelijkbare verhalen, blijkt dat investeerders en beleggers nog steeds gevoelig zijn voor de verleidingen van dergelijke claims. Een kleine nascholing in de informatietheorie kan u dus wellicht geld besparen. Het onderschrift “inleiding tot de informatietheorie voor investeerders” dient niet al te serieus genomen te worden. Dit boekje is bedoeld voor iedereen die geïnteresseerd is in het verwerken en comprimeren van informatie. En mocht u een nieuwe uitvinder van super-compressie tegen het lijf lopen, dan heeft u in elk geval volop gespreksstof.

*Rombout Kerstens*

# Inleiding

*Energie kan niet verloren gaan noch kan het uit het niets ontstaan.*

Eerste wet van de thermodynamica

*Een perpetuum mobile (Latijn: eeuwig bewegend) is een hypothetische machine die in staat is om energie op te wekken uit niets. De thermodynamica heeft al meer dan een eeuw geleden de onmogelijkheid hiervan aangetoond. Patenten op een perpetuum mobile worden niet verleend.*

Gevolg van de eerste wet van de thermodynamica

## **Informatiemaatschappij**

De rol van informatie wordt in onze maatschappij steeds belangrijker. Op alle niveaus binnen de samenleving hebben systemen hun intrede gedaan die informatie opslaan, transporteren en beveiligen. Digitale en analoge informatie, vastgelegd in signalen of symbolen. Van telefoongesprekken tot televisie en van boeken tot internet. Aan de basis van deze ontwikkelingen staat de informatietheorie. Pas de afgelopen 20 jaar kon deze theorie worden vertaald naar de praktijk door de geweldige vooruitgang van de informatietechnologie. De opkomst van steeds snellere en krachtiger computers heeft er voor gezorgd dat de denkbeelden van de grondleggers van de informatietheorie ook in het dagelijks leven zichtbaar werden. Zonder hen waren zaken als internet, mobiele telefonie en kantoorautomatisering nu ondenkbaar geweest.

## **Informatietheorie**

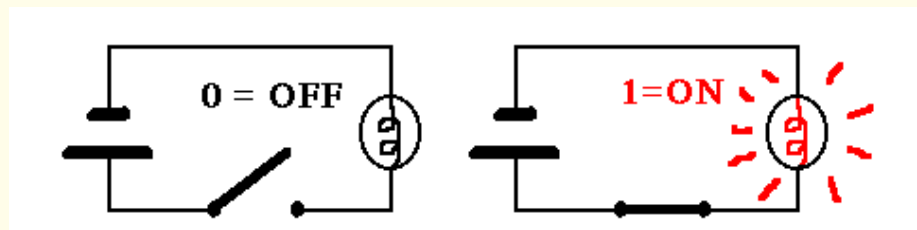
Toen investeerder Marcel Boekhoorn warm gemaakt moest worden voor de uitvinding van Jan Sloot gebeurde dat met de ronkende slogan: “Deze vinding is de brandstofloze motor van de ICT-wereld”. Bewust of onbewust bevatte deze *sales pitch* zowel een aanmoediging als een waarschuwing. Inderdaad zou het comprimeren van een speelfilm tot 1 kilobyte een technologische revolutie veroorzaken. Aan de andere kant: hoeveel ‘uitvinders’ hadden al niet het evangelie van de brandstofloze motor gepredikt? En toch moeten we nog elke dag onze auto voltanken. En er is nog een overeenkomst tussen fabelachtige compressietechniek en de brandstofloze motor. Namelijk die tussen de informatietheorie en de thermodynamica. Thermodynamische systemen (bijvoorbeeld een stoommachine of een automotor) en informatie blijken in hun wiskundige basisprincipes meer met elkaar te

maken te hebben dan je op het eerste gezicht zou denken. Helaas is de informatietheorie, net als de thermodynamica, geen lichte kost. Menig technisch student gaat met lood in de schoenen naar deze tentamens. Toch is een beter begrip van de informatietheorie zeer behulpzaam bij het oplossen van de raadsels rondom het dossier Jan Sloot. Ook wordt dan duidelijker hoe zelfs begaafde technici 'in de ban van Jan Sloot' konden raken.

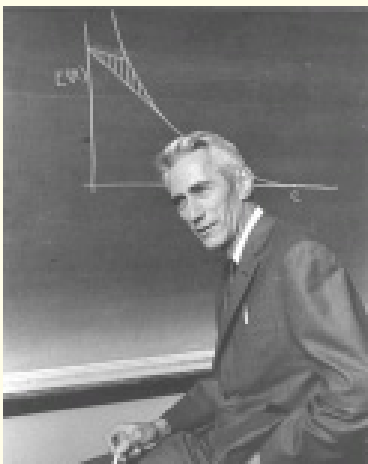
*Uitvinder Jan Sloot stierf op 11 juli 1999 thuis in zijn tuin aan een hartaanval, één dag voordat hij de broncode van zijn vinding definitief zou overdragen aan een groep investeerders rond Roel Pieper. Sloot zou een methode hebben ontwikkeld om tientallen speelfilms op te bergen op een chipkaartje van 64 kB. De specificaties van zijn vinding zijn na zijn dood nooit teruggevonden. In de kluis van Jan Sloot werden slechts wat verouderde beschrijvingen van een eerdere vinding, enkele persoonlijke documenten en een boek van Ludlum aangetroffen. Het bedrijf dat van Sloots vinding een miljardenindustrie zou maken is in de slaapstand gezet. Pieper en zijn mede-investeerders zijn hun miljoenen kwijt.*

## Bits en Bytes

De fundamentele eenheid voor informatie is een ja-nee situatie: iets is waar of niet waar. Dit kan eenvoudig worden uitgedrukt in binaire algebra door 1 en 0, waarbij 1 betekent dat de schakelaar open is. 0 betekent “off”, de schakelaar is gesloten.



Onder deze omstandigheden zijn 1 and 0 “binary digits” (binaire getallen), afgekort “bits”. De eenheid voor informatie is dus de bit. Een samenstelling van meerdere bits noemen we een Byte. Meestal wordt met een Byte een samenstelling van 8 bits bedoeld. Genoeg om 256 mogelijke waarden aan te geven. Die waarden kunnen worden gecodeerd als getallen en letters met een karakertabel, waarvan ASCII de bekendste en meest gebruikte is.



### ***De informatiemaat van Shannon***

De Amerikaanse wiskundige en elektrotechnicus Claude E. Shannon (1916-2001) legde in 1948 met zijn standaardwerk “The Mathematical Theory of Communication” de basis voor de informatietheorie. En wordt dan ook wel ‘de vader van de informatietheorie’ genoemd. Zijn werk borduurde voort op eerdere publicaties van Harry Nyquist en Ralph Hartley. In dit beroemde werk introduceerde hij het begrip entropie uit de thermodynamica in de informatietheorie. Het begrip entropie is een maat voor de hoeveelheid wanorde of chaos in een systeem. Elk systeem zal zonder invloed van buiten uit zichzelf streven naar een gelijkblijvende of grotere wanorde. Mensen die een hekel hebben aan het opruimen van hun bureau zullen dit kunnen beamen. Sneeuw op de TV en ruis op de radio zijn andere vormen van hoge entropie of wanorde.

Shannon gebruikte het begrip entropie als maat voor de hoeveelheid data die maximaal over een bepaald informatiekanaal verzonden kan worden. De ideeën van Shannon hadden vooral betrekking op het transporteren van data. In dit proces onderscheidde hij een zender van de informatie, het informatiekanaal en de ontvanger. Shannon was niet geïnteresseerd in de betekenis of het belang van de informatie, maar zag het als een grootheid

op zich die met waarschijnlijkheids-rekening en wiskunde kon worden beschreven. Shannon was de eerste die het begrip kans in verband bracht met informatie. Om dit te kunnen begrijpen geef ik het volgende voorbeeld. Stel u wilt informatie overbrengen met lichtsignalen. U gebruikt twee signalen: een lang en een kort signaal (gelijk de punt en de streep in morsecode). Vervolgens geeft u (de zender) iedere tien seconden een kort signaal met een zaklantaarn (het informatiekanaal). Laten we zeggen gedurende een uur. U heeft dan 360 identieke signalen gegeven. De kans dat ieder opeenvolgend lichtsignaal kort was is voor de ontvanger 100% (kans = 1). Hoewel er dus een uur is geflitst is er geen informatie overgedragen. De hoeveelheid overgedragen informatie is gelijk aan nul. Al dat geflits was eigenlijk volledig overbodig, redundant in informatietermen.

Het informatiekanaal draagt pas daadwerkelijk informatie over als er tussen de korte flitsen ook lange flitsen voorkomen. Het informatiekanaal bevat dan wél informatie, aangezien het onzeker is geworden wat het volgende lichtsignaal zal zijn. De kans dat een bepaald signaal optreedt is dan ook een belangrijk gegeven bij het berekenen van de maat van de hoeveelheid informatie die wordt doorgegeven. Bij een kans dat een van de twee signalen optreedt van een half (50%) bereikt het kanaal zijn optimum. Het kanaal heeft dan zijn maximale capaciteit of entropie bereikt. Er is geen redundante informatie meer. Deze staat van entropie is ook de ondergrens voor datacompressie. De informatie kan fundamenteel niet verder worden gecomprimeerd.

Worden er bijvoorbeeld tien verschillende signalen gebruikt dan zal de maximale entropie optreden bij een individuele kans per signaal van 10%.

Shannon houdt zich niet bezig met de inhoudelijke aspecten van de signalen. Het gaat alleen om de kans dat een signaal optreedt, niet om het signaal zelf.

## Compressie en compressie-algoritmen

Compressie is het coderen van de totale hoeveelheid bits van een bestand teneinde de bestandsgrootte te reduceren. Doel is meestal geheugenruimte te besparen of de overdrachtsnelheid te vergroten.

Compressie-algoritmen zijn algoritmen die gebruikt worden voor datacompressie. Ze zijn te verdelen in verliesvrije (lossless) en niet-verliesvrije (lossy) algoritmen.

Uit een verliesvrij gecomprimeerd bestand kan het originele bestand weer identiek gedecomprimeerd ('geëxpandeerd') worden, er gaat daarbij geen informatie verloren.

In tegenstelling tot lossless algoritmen gooit een lossy algoritme informatie weg. De informatie die bij een lossy algoritme wordt weggegooid is idealiter niet zichtbaar (bij plaatjes JPEG of PNG) of hoorbaar (bij muziek MP3). Echter, hoe groter de compressie hoe meer van de beeld- of geluidskwaliteit verloren gaat.



### ***Kolmogorov-complexiteit***

De Russische wiskundige Andrey N. Kolmogorov (1903 – 1987) ontwikkelde in de jaren zestig een theorie over de complexiteit van objecten. Hij nam de complexiteit van een object als maat voor de mate van compressie welke mogelijk is.

De Kolmogorov-complexiteit definieert de complexiteit van een reeks karakters (string) als de lengte van het kortste programma (in bits) dat deze string kan reproduceren. Strings welke door korte programma's kunnen worden gereproduceerd worden als niet erg complex beschouwd. Het doorgronden van deze stelling leidt tot een diep inzicht in het wezen van data en datacompressie. De Kolmogorov-complexiteit heeft grote overeenkomsten met het begrip entropie van Shannon.

In de praktijk kunnen we goed zien wat het bovenstaande betekent voor het comprimeren van data. Door een voorspelling te doen over de complexiteit van een object kunnen we vaststellen in hoeverre dit goed te comprimeren is. Van onderstaande reeks van nullen kunnen we al gelijk vaststellen dat dit niet erg complex is.

*0 et cetera.*

In elke programmeertaal zal een programma om een string van uitsluitend nullen te produceren bijzonder kort zijn. Moeilijker



wordt het al bij het doorgronden van de complexiteit van de onderstaande string:

0 1 1 2 3 5 8 13 21 34 55 89 144 233 377 610 987 1597 2584 4181  
6765 10946 17711 28657 46368 et cetera.

De bèta's onder ons zien al snel dat elk getal in deze string bestaat uit de som van zijn twee voorgangers, de zogenaamde Fibonacci-reeks. Dit betekent dat een programma om deze string te reproduceren bijna net zo kort zal zijn als een string van enkel nullen. Toch hebben vrijwel alle moderne compressieprogramma's moeite om deze string te comprimeren, terwijl deze de reeks nullen wel moeiteloos tot bijna nul kunnen comprimeren. Dat komt omdat dergelijke programma's (bijvoorbeeld Zip, Gzip, Arc) geen rekening houden met de logica die in een bepaalde string zit.

Dit is praktisch ook onhaalbaar, aangezien dit een waanzinnig rekenintensieve exercitie zou worden. Dit voorbeeld geeft al aan dat er niet zoiets is als het compressie-algoritme dat voor iedere toepassing het optimale resultaat geeft. Voor de meeste standaard compressieprogramma's lijkt de Fibonacci-reeks op een reeks random getallen en zal derhalve een vergelijkbare compressie-ratio geven. In tegenstelling tot de Fibonacci-reeks heeft een reeks random getallen normaal gesproken wel een grote Kolmogorov-complexiteit.

Uiteraard kan het principe van de Kolmogorov-complexiteit worden toegepast op alle data. Dus ook op boeken, spraak, tekstbestanden en films. Hoe complexer iets is hoe lastiger het te comprimeren valt. Een film die uitsluitend bestaat uit een egaal rood beeld gedurende 90 minuten is niet complex en eenvoudig te comprimeren. Enkele bits volstaan, voor de kleur en voor de duur. Anderhalf uur (random) sneeuw op het televisiescherm is weer bijna niet te comprimeren, althans niet zonder verlies van informatie.

*In het boek De Broncode vertelt Jan Sloot aan iedere potentiële investeerder die het horen wil dat hij in staat is een film van negentig minuten verliesloos (lossless) te comprimeren tot 1 kilobyte. Op de chipkaarten die hij gebruikte voor opslag paste 64 kB, dus totaal 64 films. Dat klinkt fabelachtig en het is niet vreemd dat menig*

*geldschietter het hoofd op hol is gebracht met deze claim. Inderdaad de brandstofloze motor van de ICT!*

*En hier wreekt zich het feit dat de informatietheorie geen verplicht vak is op de middelbare school. Hoewel dit laatste kennelijk evenmin garanties biedt, want vele technische high-potentials die er voor gestudeerd hadden vergaloppeerden zich ook. Blijkbaar was de demonstratie zo overtuigend dat een nadere beschouwing van de feiten overbodig werd geacht. Laat staan dat de studieboeken weer van zolder werden gehaald.*

# **Alternatieve methoden voor compressie**

## ***Fractale beeldcodering***

Fractale beeldcompressie is een vrij nieuw fenomeen. De methode berust op de ontdekking in de jaren tachtig dat door middel van herhaalde transformaties (iteraties) van een eenvoudig basispatroon bedrieglijk echte beelden zijn te genereren. Het idee is nu dat als er met een beperkt aantal iteraties een boom kan worden gevormd die net echt lijkt, het omgekeerde ook mogelijk moet zijn. Dus een foto terugrekenen naar een beperkt aantal fractals inclusief het algoritme om het oorspronkelijke beeld weer terug te krijgen. Deze techniek staat momenteel nog in de kinderschoenen. In special effects van films wordt al veel gebruik gemaakt van fractals om bijvoorbeeld realistische wateroppervlakken te generen uit een simpel basispatroon. Het voordeel van deze methode is dat het met de computer genereren van beelden nu veel minder rekentijd en opslagcapaciteit vergt.

## ***Neurale netwerken met geheugen***

Een neuraal netwerk is een kunstmatige simulatie van de werking van menselijke hersenen. Een neuraal netwerk bestaat uit softwarematig nagebootste hersencellen, neuronen, die net als in onze hersenen met elkaar in verbinding staan. Als een neuron signalen krijgt die een bepaalde drempel overschrijden dan geeft deze een signaal door aan de omliggende neuronen. Ieder neuraal netwerk heeft inputneuronen, outputneuronen en lagen met neuronen daar tussenin. Een signaal aanbieden aan de inputkant van het netwerk zal aan de outputkant een resultaat opleveren. Door de verbindingen tussen de verschillende neuronen meer of minder belangrijk te maken ontstaat een neuraal netwerk dat een bepaalde taak kan verrichten of informatie kan bevatten. Neurale netwerken die worden gesimuleerd met software of hardware onderscheiden zich van traditionele programmatuur doordat zij het vermogen hebben om te leren. Zo kan een netwerk worden getraind om bepaalde patronen te herkennen. Telkens worden de patronen aan het netwerk aangeboden en wordt er vervolgens gekeken of het netwerk deze als goed of fout herkent. Door nu de waardes van de verbindingen tussen de neuronen tijdens het trainen telkens iets bij te stellen, wordt een netwerk gemaakt dat steeds beter wordt in het herkennen.

Als je het netwerk maar groot genoeg maakt, dus veel verschillende lagen met veel neuronen, kun je een netwerk ook trainen om complete beelden te onthouden. Als je verschillende beelden

maar vaak genoeg aanbiedt zal op een gegeven moment een netwerk aan de hand van bijvoorbeeld het beeldnummer een beeld kunnen reproduceren. Hoe langer je het netwerk traint hoe scherper de beelden zullen worden. Het neurale netwerk is dan feitelijk veranderd in een groot virtueel geheugen. Denkend aan ons menselijk geheugen dat in staat is vele beelden op te slaan moet het mogelijk zijn hiermee een zeer efficiënte opslag te bereiken. Helaas tonen experimenten aan dat de tot nu toe bereikte compressiefactoren verre van spectaculair zijn. Bovendien is het resultaat van nature lossy. Het blijkt dat onze hersenen toch nog steeds veel ingewikkelder werken dan de computermatig gesimuleerde versie.

*Jan Sloot beweerde dat zijn vinding niets met compressie van doen had. Hij had het zelf liever over codering. Ook zou hij geen gebruik maken van bits en bytes.*

*In plaats hiervan zou hij gebruik maken van een eigen opslagmethode die veel efficiënter was. Toch kon zijn vinding weer wel gebruik maken van de bestaande bits en bytes van standaard computerbesturingssystemen. Getuige ook zijn demonstraties met een gewone laptop PC. Bovendien gebruikte hij een standaard chipkaart met een geheugen van 64kB.*

## Toepassingen in de praktijk

### ***De industriestandaarden***

Tot nu toe hebben we alleen gekeken naar de theorie van het kwantificeren en comprimeren van data. Maar je hebt weinig aan een mooie theorie als deze niet in de praktijk kan worden gebruikt in browsers, tekenpakketten en tekstverwerkers. Voorts is een mooi compressieformaat waardeloos als deze door de ontvanger niet kan worden gedecodeerd of niet geschikt is voor bepaalde transport- en opslagmedia. Vandaar dat er in de IT-industrie een aantal formaten is dat in de loop der tijd tot standaard is geëvolueerd. Dit is overigens een dynamisch proces, dagelijks worden er nieuwe standaarden verzonden en verdwijnen er evenzoveel weer in de vergetelheid. Omdat ze niet meer voldoen aan de eisen van deze tijd of simpelweg nooit zijn aangeslagen.

### ***Geluid en video***

Geluids- en videoformaten boeken hun compressiewinst door het weglaten van informatie en zijn dus per definitie lossy. Het blijkt dat zowel het menselijk oor als oog voor de gek te houden is door informatie weg te laten die toch niet hoorbaar of zichtbaar is. Nu is het uiteraard niet moeilijk om bewegende beelden te comprimeren door de resolutie omlaag te brengen en/of kleurinformatie weg te gooien. De kunst van het comprimeren is nu juist om zoveel mogelijk weg te gooien, zonder dat menselijke ogen en oren dat in de gaten hebben.

In de audiovisuele wereld wordt vaak gewerkt met zogenaamde *streams*. Dit betekent dat niet het gehele bestand op de computer wordt afgespeeld, maar bit voor bit direct wordt beluisterd of bekeken. De datatransportsnelheden van deze streams worden weergegeven in eenheden van 1000 bits per seconde, hetgeen wordt geschreven als kbps.

### ***MPEG audio layer 3 (MP3)***

*MPEG-1 Layer 3* (de audiolaag van de bewegende beelden compressie MPEG) is een manier om geluid te comprimeren. De veel gebruikte afkorting is *MP3*. Met MP3 is het mogelijk de hoeveelheid gegevens die nodig is voor het opslaan van geluid sterk te verminderen. Dit geschiedt door elementen uit het geluid die een mens toch niet echt waarneemt weg te laten. Bijvoorbeeld als een bepaalde toon luid is, zal een toon die er vlak naast ligt en minder luid is niet hoorbaar zijn. Deze kan men er dus uitfilteren.

### **Sample rate**

Een factor die speciaal van belang is voor zowel de grootte van het bestand als de kwaliteit van geluid en muziek is de zogeheten sample rate. Geluid is voor te stellen als een golvende grafiek met een verloop in tijd. Om zo'n analoge golf te digitaliseren moet de curve in nullen en enen worden vertaald. Daartoe wordt de hoogte van de curve duizenden keren per seconde gemeten en omgezet in digitale informatie. Het aantal keren dat zo'n analoge geluidsgolf per seconde wordt gemeten is uitgedrukt in de sample rate. Op een normale muziek-cd wordt het geluid 44.100 keer per seconde gemeten.

We spreken dan van een monsterfrequentie ofwel sample rate van 44.1 kHz. Hoe vaker de frequentie wordt gemeten, hoe hoger de kwaliteit van het geluid, maar ook hoe groter het bestand waarin de muziek wordt opgeslagen. Op de huidige muziek-DVD's wordt bijvoorbeeld al een sample rate van 192 kHz toegepast.

MP3 doet juist het tegenovergestelde. Door de sample rate te verlagen wordt de kwaliteit weliswaar negatief beïnvloed, maar neemt de bestandsgrootte ook flink in omvang af. De huidige MP3's hebben meestal dezelfde sample rate als audio-cd's. Sommige oudere MP3-bestanden hebben nog sample rates van 22 kHz of lager.

### **Lossy compression**

De MP3-compressie is *lossy*, dat wil zeggen dat er gegevens bij het comprimeren verloren gaan. Hierdoor kan de oorspronkelijke vorm niet meer worden teruggewonnen, alleen benaderd.

De manier waarop dit gebeurt heet Discrete Cosine Transformation (DCT). Deze wiskundige bewerking die ook goed werkt voor beeldinformatie (bijvoorbeeld JPEG) zorgt ervoor dat telkens stukjes van de audio datastroom worden omgezet in een matrix, welke daarna gemakkelijk kan worden verkleind. Daarnaast wordt ook extra compressie bereikt door informatie die zowel op het linker- als rechterkanaal voorkomt slechts eenmaal op te slaan. Dit laatste noemt men "joint stereo".

Geluid op 112 kbps neemt als MP3-bestand tot ongeveer 12 keer minder schijfruimte in dan op de CD. Hierdoor zijn MP3-bestanden bijzonder geschikt om muziek over *netwerken* met een relatief lage bandbreedte uit te wisselen. Gangbare formaten zijn

128, 160, 192 kbps (meest gebruikt op internet), 256 kbps en 320 kbps. Deze laatste levert een geluidswaergave op die vrijwel identiek is aan het origineel.

De compressie en de decompressie algoritmen van MP3 zijn gepatenteerd door de eigenaar, het Fraunhofer-instituut, en dus niet vrij beschikbaar.

*Zou het werkelijk zo zijn dat hij geheel van de bits and bytes was afgestapt dan zou Sloot een eigen besturingssysteem voor zijn laptop moeten hebben geschreven. Voorwaar geen eenvoudige opgave. Voor zijn latere demonstratiemodel, ingebouwd in een zwart kastje, zou hij dan misschien wel eigen chips hebben moeten vervaardigen. Maar in de inboedel is geen apparatuur aangetroffen die dit kon. Dus is het waarschijnlijker dat er een soort PC is nagebouwd in het kastje zelf. Volgens Ben Sloot, de zoon van de uitvinder, was het kastje ongeveer 25 \* 40 centimeter. En werd het gevoed met netstroom. Binnen deze afmetingen was het in 1999 zeker mogelijk om een PC te bouwen.*

### **Bewegende beelden compressie: MPEG**

Voor de compressie van bewegende beelden is MPEG de algemene standaard. MPEG-1 was oorspronkelijk bedoeld voor een beperkt aantal pixels, voldoende voor ongeveer een kwart televisiebeeld. Dat is inmiddels uitgebreid, waardoor bijvoorbeeld de nu verouderde CD-I standaard toch beeldvullend kon werken middels MPEG-1. In MPEG-2 is een aantal nieuwe voorzieningen aangebracht. MPEG-4 betekent een radicaal nieuwe aanpak waarbij objecten in het bewegende beeld apart worden behandeld en de kijker deze interactief kan manipuleren.

Hoe werkt MPEG in grote lijnen? MPEG garandeert een behoorlijke (lossy) compressie, waardoor de gecomprimeerde data soms maar 1/20 van het origineel in beslag nemen. Een video-cd van 600 Mbyte bevat ongeveer een uur televisie (MPEG-1 kwaliteit) inclusief geluid. Het zou voor de hand liggen om 25 TV-beelden per seconde volgens de JPEG-norm te comprimeren en dit op slaan. Zo eenvoudig is het echter niet, aangezien dit teveel rekentijd zou vergen en de resulterende datareductie ook nog eens onvoldoende is.

Bij MPEG worden opeenvolgende televisiebeelden met elkaar vergeleken en wordt alleen het verschil overgesleurd. De informatie-inhoud van dat verschil is heel klein en vraagt dus weinig bits. In feite wordt een voorspelling gemaakt van één

beeldje en die raming wordt afgetrokken van het origineel; alleen het verschil wordt verzonden. De compressie van dat verschil vindt plaats op dezelfde manier als bij JPEG. Er wordt begonnen met een volledig beeld. Vervolgens worden de verschillen door- gestuurd. Af en toe wordt weer een volledig beeld, een zogenaamd intrabeeld, ingevoegd anders accumuleren de fouten. MPEG bevat ook nog een soort bewegingscompensatie middels vectoren, waarmee snelle beeldveranderingen beter te coderen zijn. Als een blok pixels alleen van plaats verschoven is hoeft dan alleen de verplaatsing (vector) te worden doorgegeven.

MPEG-compressie is niet volmaakt. Vooral de codering is nogal rekenintensief: het is niet ongebruikelijk dat de verwerking van een één uur durende film op een krachtige computer enkele uren in beslag neemt. Het resultaat moet visueel gecontroleerd worden op zogenaamde artefacts, kunstmatige verschijningen in het beeld. Om deze weg te halen zijn extra bits nodig, die weer op andere plaatsen kunnen worden weggelaten, zodat de gemid- delde signaalsnelheid constant blijft. Een elektronische buffer vangt de korte variaties in snelheid op.

De MPEG-4 standaard speelt in op twee belangrijke trends: interactiviteit en personalisatie. Interactiviteit betekent dat de kijker het beeld kan manipuleren; hij kan bepaalde details weglaten en andere naar voren halen. De personen en voor- werpen in het beeld zijn, gecodeerd als objecten, over te zenden. Hierbij hoeft alleen het object te worden gecodeerd, niet het hele plaatje. Het is vergelijkbaar met een tekenprogramma waarmee in verschillende lagen objecten en kleuren worden gecreëerd die samengevoegd een totaalbeeld opleveren. Op die manier zijn objecten dus gemakkelijk in te voegen en weg te laten. Hetzelfde geldt overigens voor het geluid: men kan een reporter die in een studio zit plaatsen in een drukke winkelstraat of in de stille natuur. MPEG-4 werkt net als MPEG-1 en 2 op basis van DCT en JPEG.

### ***Digitale televisie***

De moderne digitale televisie, vastgelegd in de DVB-norm, maakt gebruik van MPEG-2. Hiermee is het mogelijk over de band- breedte van een traditioneel analoog TV-kanaal maar liefst 7 digitale TV-kanalen door te sturen. Hiermee kan de capaciteit van bestaande TV-kabelnetten dus behoorlijk worden uitgebreid.



Uiteraard heeft de consument wel een speciale decoder nodig om het digitale MPEG-2 signaal te decoderen naar een standaard, analoog TV-signaal. Moderne digitale TV-toestellen hebben een ingebouwde decoder.

### **MIDI**

MIDI is een manier van coderen van muziek, waarbij alleen de noten worden doorgegeven. De instrumenten of klanken zitten in een databank. MIDI is een zeer compact formaat. MIDI is dus geen compressie, maar een vorm van codering.

Een heel muziekstuk kan in enkele kilobytes worden opgeslagen. Maar er is altijd een externe databank nodig om deze af te spelen. Of een *interface* naar echte of nagebootste instrumenten. Een afgespeelde MIDI-file heeft dus altijd iet kunstmatigs. Zelfs als een MIDI-file met een *interface* wordt afgespeeld op een echte piano (bijvoorbeeld met een pneumatische toetsaanslag) wordt nog altijd de *touch* van een pianist van vlees en bloed gemist.

*Het moeilijkst te weerleggen zijn de claims dat Sloots vinding geheel losstond van de bestaande informatietheorie en “moedig ging waar nog geen mens geweest was” (vrij vertaald naar Star Trek). Vaak wordt het dan al snel metafysisch. Zwarte gaten komen voorbij, wormgaten en voor je het weet valt de term “snaartheorie”. Maar al deze bespiegelingen brengen volgens mij de broncode niet echt dichterbij.*

## Het scenario van Bit by Bit

Stel een alternatieve videokunstenaar exposeert een videofilm die uitsluitend bestaat uit egaal witte en zwarte beelden. Elke seconde verschijnt er random een nieuw wit of een zwart beeld. Met twee en een half uur speeltijd betekent dit dat er 9000 beelden zijn. Er bestaan dus twee tot de macht 9000 mogelijke varianten (permutaties) van de Bit by Bit film. Hiervoor zijn al minimaal 1125 Bytes opslag nodig om iedere variant onder een unieke code te kunnen opslaan. Bij het decoderen moet immers iedere unieke code terugverwijzen naar slechts één van alle mogelijke versies. Kan dit niet dan is er geen sprake van verliesloze compressie.

Het bovenstaande is dus in tegenspraak met de claim van van Jan Sloot dat hij iedere willekeurige speelfilm (of ander bestand) verliesloos kon comprimeren naar 1000 Bytes. Het bijbehorende geluid laten we dan nog buiten beschouwing. En zelfs de grootste pulpfilm aller tijden zal hoogstwaarschijnlijk nog iets ingewikkelder zijn dan Bit by Bit.

We moeten bij compressie ook even onderscheid maken tussen het begrip van de meeteenheid van informatie, de bits, en de fysieke verschijningsvorm (geheugenchips, DVD's, cd's en harddisks). Deze fysieke vorm is de afgelopen jaren steeds kleiner geworden. Ook de transportnetwerken van data kunnen steeds meer informatie verwerken. Dit alles heeft echter niets met datacompressie te maken. De eenheid van informatie blijft een bit, hoe die wordt opgeslagen en getransporteerd is een andere zaak. In de discussies over het dossier Jan Sloot worden deze zaken nogal eens op een hoop gegooid. Zeker met de huidige ontwikkelingen in transport en opslag wordt al snel gewezen op de mogelijkheid dat de vinding van Jan Sloot toch is opgedoken, terwijl het dan gaat over de fysieke verschijningsvorm van de opslag van de bits en Bytes. Zo heeft het Philips NatLab ontdekt dat met het gebruik van Blauwe Laser nog veel meer informatie op een nog kleiner schijfje kan. Maar alles blijft gewoon binnen de context van de informatietheorie.

### Het patent van Jan Sloot

Voor iemand die de zaak wilde oplichten had Jan Sloot er wel werk van gemaakt. Zijn vermeende coderingstechnologie is bij het Nederlands Octrooibureau in 1998 ingediend onder de rubriek: *Werkwijze en Inrichting voor het opslaan van gegevens*. Wat hier

feitelijk wordt beschreven is een werkwijze welke grote overeenkomsten heeft met het Lempel – Ziv principe, maar dan speciaal toegesneden op TV beeldinformatie. In het octrooi wordt gesproken over een compressiefactor van 8. Wel iets minder dan de latere claim van een factor 2 miljoen, hetgeen vergelijkbaar is met het comprimeren van al het water in een olympisch zwembad in een pak melk.

*Prof.dr.ir. Kees Schouhamer Immink (1946), ontwerper van coderingstechnieken en lange tijd werkzaam geweest bij het Philips NatLab, zei in een interview met het Eindhovens dagblad over de zaak Sloot:*

*De octrooien - verleend in 1997 en 1998 - zijn 'het papier niet waard waarop ze geschreven zijn. Voor een octrooi moet aan drie belangrijke voorwaarden worden voldaan: nieuwheid, uitvindersgehalte en industriële toepasbaarheid. Deze octrooien bewijzen dat in Nederland nauwelijks nog een toetsing plaatsvindt en dat octrooi-aanvragen veel te gemakkelijk worden gehonoreerd.' De supercompressie van Jan Sloot noemde hij 'Absolute onzin' en 'Een factor 2 zou nog geloofwaardig zijn, maar een factor 2 miljoen is onmogelijk.'*

## De bewijsvoering van Jan Sloot

Vele betrokkenen die een demonstratie van Jan Sloot hebben bijgewoond waren ervan overtuigd dat hij geen oplichter was. Hij gaf demonstraties met een kastje dat videofilms vertoonde als hij er een chipkaart instak. De bijeengeroepen zakenlieden en technici lagen zelfs met hun oren op het kastje om zich ervan te overtuigen dat daar geen harde schijf, CD-ROM of DVD inzat. Bij eerdere demonstraties werd een laptop gebruikt, en ze keken of de harddisk daarvan niet actief was. Zo wisten ze zeker: de films stonden op het chipkaartje en Sloot had 'de uitvinding van de eeuw'.

Dat de demonstratie van Jan Sloot overtuigend was staat vast. In het boek *De Broncode* komen tientallen getuigen aan het woord die verklaren dat ze onder de indruk waren van wat ze zagen. Zestien films werden simultaan op een beeldscherm afgespeeld. Alhoewel de beeldkwaliteit wisselde per demonstratie en er nooit sprake is geweest van echt goede kwaliteit. Alle betrokkenen verklaren dat de films begonnen te spelen op het moment dat de chipkaart in de gleuf van het magische kastje werd gestoken. De suggestie die hiervan uitging is natuurlijk behoorlijk sterk. Alle feiten die je tegelijkertijd waarneemt doen je al snel geloven dat het niet anders kan dan dat alle films daadwerkelijk op de chipkaart staan. Uiteraard was het de taak van de aanwezige technici om door de waargenomen feiten heen te prikken. En er werden ook kritische vragen gesteld, maar ze waren al half overtuigd door de kracht van de demonstratie.

Hoe kun je bewijzen, zonder het kastje open te schroeven, dat zijn claim op waarheid berust? Dus dat de chipkaart het enige medium is dat de film kan bevatten en dat er geen aanvullende data in het kastje zelf zit, dat noodzakelijk is om de film af te spelen? Dit kan alleen met de volgende methode. Laat in twee kamers een opstelling maken. In de ene kamer een opstelling om films te coderen en in de andere kamer om de films te decoderen. Geef een onbekende speelfilm of testvideo vlak voor de demonstratie om te coderen en ga vervolgens zelf met de chipkaart naar de andere ruimte om de film of video te decoderen. Uiteraard mogen de ruimtes niet met elkaar in verbinding staan en mag er geen radiocontact tussen de ruimtes mogelijk zijn (anechoïsche ruimte). Deze test werd ook door verschillende investeerders als voorwaarde gesteld. Ook de ingenieurs van Philips NatLab stelde deze test als voorwaarde voor verdere samenwerking. Door allerlei politieke spelletjes

tussen diverse investeerders en Philips heeft deze test echter nooit plaatsgevonden. In een aantal andere gevallen hield Jan Sloot zelf de boot af met allerlei uitvluchten.

Vreemd genoeg wordt in het boek *De Broncode* (pagina 42) wel melding gemaakt van een demonstratie die erop lijkt. Een stukje van de vergadering wordt opgenomen met een digitale camera, vervolgens vanaf de camera direct gecodeerd en met behulp van een cardreader gekoppeld aan een laptop weer afgespeeld. Deze demonstratie was echter een thuiswedstrijd op het kantoor van Jan Sloot zelf en er waren geen andere technici aanwezig. Uitsluitend investeerders en een accountant. Er waren dan ook mogelijkheden om deze hele demonstratie te falsificeren, zoals verborgen draden of een radioverbinding. Bovendien ontbrak een kritische toeschouwer.

Naast de ingenieurs van Philips was ook de benaderde investeerder Eckart Wintzen niet te vermurwen om er een cent in te steken voordat het bewijs was geleverd dat het werkte. Maar ook benaderde Amerikaanse investeerders verlangden bewijzen. Het blijft echter de vraag waarom diverse andere investeerders toch miljoenen wilden steken in de vinding van Jan Sloot zonder dat bovenstaande test was uitgevoerd.

*Wat zat er in dat magische kastje?*

Het zou natuurlijk het mooiste zijn om het kastje zelf open te schroeven. Dit kan in principe ook want er is nog een exemplaar in bezit van Roel Pieper. Alleen: hij laat het aan niemand zien. Desgevraagd ook niet aan de auteur dezes. Dus het blijft helaas bij speculeren.

***Het kastje volgens de beschrijvingen van getuigen:***

*Afmetingen: +/- 25 X 40 cm (ongeveer een grote koektrommel), kleur zwart. Met een lichtgevende tuimelschakelaar bovenop. En een chipcard reader aan de voorkant ingebouwd. De voeding was netspanning. Aangezien het systeem ook in Amerika heeft gefunctioneerd moet het voltage omschakelbaar zijn geweest.*

Bij het speculeren over de inhoud van het kastje ga ik uit van drie scenario's:

### *1. De meesteroplichter*

Dit is de meest brutale mogelijkheid. Het hele optreden van Jan Sloot als de neurotische uitvinder met paranoia was een act. Daaronder zat een berekenende oplichter, die zijn technische gaven misbruikte om investeerders geld uit de zak te kloppen. De techniek in het kastje zou dan uit een kleine cassette videospeler kunnen bestaan. Dit stelt echter wel hoge eisen aan de demonstrateur. Aangezien alleen een vast programma kan worden afgespeeld. De interactiviteit moet geheel worden gesuggereerd. Alleen een demonstratie met passieve toeschouwers is goed uitvoerbaar. Een videorecorder heeft een trage opzoektijd (access time) en kan niet wat de getuigen van de demonstraties vertelden dat ze zagen (en ze hoorden geen loopwerk). Zoals het interactief van begin tot einde doorspoelen van een hele film in een paar seconden, en het op aanvraag heen en weer springen tussen het mozaïek beeld van 16 films en een full screen TV beeld.

Een betere mogelijkheid om de zaak te flessen is een klein PC-moederbord in het kastje te bouwen met een zogenaamde statische RAM-Disk. Dit is een met RAM-geheugen nagebootste harddisk. De RAM-Disk heb je nodig voor de performance van het doorspoelen van de films en het feit dat getuigen verklaarden dat er geen harddisk te horen was. Hierop wordt, naast het besturingssysteem van de PC, een aantal MPEG-files met de gewenste inhoud geplaatst.

In principe kan deze oplossing alles wat de getuigen hebben aanschouwd: de muisbesturing, het mozaïek beeld en het razendsnel doorspoelen van de MPEG-files. Er zijn wat detailproblemen, zoals performance en het gebruikte PC-besturingssysteem. Het grootste probleem is echter dat niet 16 MPEG-files van 90 minuten speelfilms volledig op de RAM-drive konden passen. Dit was met de techniek in 1999 eigenlijk nog niet mogelijk.

Maar 100% noodzakelijk voor een overtuigende demonstratie is dit ook niet. Niemand controleert tijdens een demo of bij het razendsnel heen- en terugspoelen door een film ook de volle 2 uur en 54 minuten van de Sound of Music voorbij komt! Van het merendeel van de films heb je alleen clips nodig, en van een paar films een behoorlijk deel. Afhankelijk van hoe begaafd de demonstrateur is in het manipuleren van zijn publiek, heb je minder minuten film nodig.

De chipkaart wordt in dit scenario alleen gebruikt om een switch aan of uit te zetten. De informatie die erop staat doet geheel niet terzake. Dit alles vereist voor de uitvoering wel een goede elektronicus die ook praktisch zeer handig is. Hetgeen ook overeenkomt met het profiel van Jan Sloot met zijn verleden als begaafd TV-reparateur.

## *2. De mislukte uitvinder*

In dit scenario heeft Jan sloot werkelijk een stuk software geschreven waarmee hij zoals beschreven in zijn patent video kan comprimeren. Dit is zeker niet onmogelijk. Wat de software dan doet is de film beeld voor beeld in stukjes hakken en die stukjes die veel voorkomen oproepen met een korte code. Dit is in feite het LZW-algoritme direct toegepast op videobeelden. Op de chipkaart staat dan als het ware alleen de gebruiksaanwijzing hoe al die puzzelstukjes weer één film worden. Deze gebruiksaanwijzing is relatief klein. Het is vergelijkbaar met het MIDI-principe. Het probleem is echter dat de database, welke zeer groot moet zijn, slechts geschikt is voor één film. De kans dat de puzzelstukjes beeld geschikt zijn voor andere films is immers zeer klein. Wellicht dat sommige films in een database gecombineerd kunnen worden, als je genoeg neemt met een mindere beeldkwaliteit, maar Jan Sloot claimde verliesloze compressie. Het is onmogelijk dat er met een database met beeldmateriaal, volgens Jan Sloot zelf was dit geheugen 370 Megabyte, alle bekende en toekomstige films weer kunnen worden weergegeven. Daarvoor is een gemiddelde speelfilm een te complex geheel (zie ook Kolmogorov). Dit is ook op een andere manier in te zien. Stel dat het mogelijk zou zijn uit een database van 370 MB alle speelfilms op te roepen. Het aantal gebruiksaanwijzingen die op de chipkaart passen zijn eindig (zie Bit by Bit). Door een groot aantal willekeurige gebruiksaanwijzingen in te voeren zou het dan mogelijk moeten zijn een bioscoopkraker op het scherm te krijgen die nog opgenomen moet worden!

Scenario twee is feitelijk samen te vatten als het op een bepaalde manier versleutelen van een DVD tot een database. De code op de chipkaart is dan alleen een soort database ontgrendelingscode om de DVD te kunnen bekijken. Met als neveneffect dat er wellicht enige –niet spectaculaire- compressie plaatsvindt. Roel Pieper (RP) geeft zelf in een interview met Eric Smit (ES) voeding aan deze theorie, getuige onderstaand fragment:

ES: *Het gaat in potentie om een Nederlands Microsoft...*

RP: Het was een mooie uitvinding die zeker heel veel mensen de kop om had kunnen draaien. Van: wat is dat nou. Dat hebben we ook bij Kleiner Perkins neergelegd. Die zeiden...als je dit kan waarmaken...perfect.

Maar als je het niet kan waarmaken, dan heb je het niet. Ik verwijs weer even naar het voorbeeld van de DNA-structuur. Als je het niet hebt, dan heb je het niet.

ES: *Bedoel je dat de uitvinding niet werkt?*

RP: Nee, dat wat er is, werkt. Maar dat stukje dat niet af was...Daarom heeft het ook geen patent...Dat stukje is het onderdeel dat het echt in bedrijf brengt in een netwerk. Hij kon een stukje laten zien - 'proven concept' - maar...

ES: *Sloot kon video, audio en data versleutelen?*

RP: Maar hij kon het niet versturen op een netwerk, daar zat het probleem.

ES: *Maar een sleutelcode zoals Sloot die kon maken, kun je toch gewoon over een netwerk versturen?*

RP: Maar daar ging het niet om, je moest de database versturen en dat lukte hem niet.

ES: *Maar een database kun je ook als een stuk electronica op de markt brengen, verpakt als een soort settopbox.*

RP: Nu gaan we er even te diep op in. Daar kunnen we beter een andere keer verder over praten, maar dan wel op basis van fair play.

Hier wordt duidelijk dat Sloot waarschijnlijk ook aan Roel Pieper uiteindelijk heeft verteld dat alleen de code versturen niet genoeg was. Er moest ook een database worden verstuurd. En in die database bevond zich de eigenlijke data. En niet in de code. Waarschijnlijk moest voor iedere film een nieuwe database worden verstuurd. En dan kan je net zo goed de originele DVD-datafile sturen. Het idee van een universele database van honderden MB's inblikken in een soort settopbox en deze als warme broodjes verkopen is hiermee feitelijk van de baan.



Het beste wat er met het patent van Jan Sloot bereikt kan worden is waarschijnlijk nooit beter dan de bestaande industrie-standaarden. De inhoud van het demokastje van Sloot kan in dit scenario van alles bevatten. Het is zelfs mogelijk dat hij zelf een kleine computer heeft gebouwd op een eigen ontwerp printplaat met een aanzienlijk intern en extern geheugen (RAM-Drive). Met gewone standaardprocessoren (CPU's) en geheugenchips. Hopelijk geeft Roel Pieper nog eens opening van zaken. Samen met mij zullen een hoop nieuwsgierige elektronica knutselaars hem dankbaar zijn.

### *3. De geniale Nobelprijswinnaar*

Dit had de grande finale van dit werkje moeten worden. Als ik er althans in geslaagd zou zijn het geheim van de supercompressie te onthullen. Mysterie opgelost. Helaas kan dit gedeelte naar mijn mening alleen als pure fictie geschreven worden. Alle beschikbare feiten duiden erop dat de geclaimde compressiefactoren onmogelijk zijn en dat de benodigde hard- en software om dit te kunnen bereiken nooit heeft bestaan. Elke speculatie over de inhoud van het kastje zal onvermijdelijk het scenario voor een science-fiction film opleveren. Of een boek van Ludlum.

## **Demonstraties Jan Sloot zoals beschreven in het boek *De Broncode***

### **1995**

- zomer* Jan Sloot heeft naar eigen zeggen een doorbraak bereikt in zijn technologie om bestanden te verkleinen en laat dit thuis zien aan zijn vrouw.
- nazomer* Jan Sloot demonstreert zijn videocompressie op een laptop-PC aan investeerder Jos van Rossem.

### **1998**

- januari* Demonstratie met laptop en cardreader aan investeerder Mark Herweijer.
- januari* Eerste en enige demonstratie waarin getuigen verklaren dat er daadwerkelijk overdracht via chipkaart van beelden (live gefilmd tijdens vergadering) plaatvindt. Aanwezig: investeerder Leon Sterk, diens vriendin en accountant Balthasar Kakisina. Jan Sloot maakt gebruik van laptop met cardreader.
- voorjaar* Investeerder Eckhart Wintzen uit zijn twijfels met de woorden ‘demo or die’ en stelt aan een verdere demonstratie de eis dat bewezen wordt dat de chipkaart de daadwerkelijke drager is van informatie. Hier komt verder geen vervolg op.
- november* Mislukte grootschalige demonstratie in Stork-fabriek. Nu voor het eerst niet met een laptop, maar voor het eerst met een “solid state” kastje, dat echter net voor de demonstratie kapot valt.

### **1999**

- januari* Demonstratie, nu weer met laptop en cardreader, voor ondernemer Bart van Rheenen, aanwezig onder andere Dick Vesters, database expert van Oracle Nederland.
- 4 maart* Demonstratie met kastje (geen harddisk hoorbaar) en door Jan Sloot meegebrachte 24-inch monitor op hoofdkantoor Philips, aanwezig Roel Pieper, eerdere investeerders in het project en NatLab-ingenieur Carel-Jan van Driel.
- 17 maart* Tweede demonstratie voor Philips in kantoor van de Picus-fabriek. Getoonde beelden op het scherm zijn afkomstig van een DVD die eerder door Philips is opgestuurd en door Jan Sloot moet zijn gecodeerd.
- 2 april* NatLab doet schriftelijk demonstratievoorstel om te bewijzen dat chipkaart het werkelijke overdrachtsmedium is. Aan voorstel wordt geen gehoor gegeven
- 9 juni* Demonstratie bij Computer Associates in New York met kastje.
- juni* Demonstratie met kastje op kantoor venture capitalist Kleiner & Perkins in Silicon Valley. Eerste demonstratie zonder

aanwezigheid Jan Sloot. Kleiner & Perkins doet voorstel voor Due Diligence om technologie te bewijzen.

*9 juli* Demonstratie in Bloemendaal aan investeerder Tom Perkins. Nu bevindt zich in het kastje plotseling wel een harddisk, volgens Jan Sloot wegens technische problemen.

*11 juli* Uitvinder Jan Sloot overlijdt in Nieuwegein. Alleen kastje met harddisk wordt teruggevonden. Geen spoor van programmatuur en broncode van Jan Sloots vinding.

## Nog meer compressie hypes

In 1992 claimde het bedrijf *WEB Technologies* uit de Amerikaanse staat Georgia dat het een compressie-algoritme had ontwikkeld dat ieder bestand kon terugbrengen tot 1024 Bytes. Een reporter van het Amerikaanse computerblad *Byte* ontving een betaversie van het WEB-programma *DataFiles/16*, welke in elk geval gegarandeerd een compressiefactor 16 zou halen. Het programma comprimeerde wel, maar de gedecomprimeerde file vertoonde helaas geen gelijkenis met het origineel. WEB zegde toe een nieuwe goed werkende versie te sturen, maar deze is nooit aangekomen. Na drie maanden belde de reporter nog eens. Helaas was de telefoon afgesloten. Diverse pogingen om iemand van het bedrijf te spreken liepen op niets uit. Het bedrijf had zichzelf kennelijk succesvol tot nul gereduceerd.

Andere compressie hypes zoals die van het bedrijf *ZeoSync* in Amerika (2002) en *Adams Platform Technology* (2003) in Australië vertonen gelijkenis met de affaire Sloot. Eerst worden investeerders en publiek warm gemaakt met mooie claims, maar de uiteindelijke tests die moeten uitwijzen dat het werkelijk werkt worden keer op keer uitgesteld. In 2002 probeerde het Amerikaanse bedrijfje *ZeoSync* 40 miljoen dollar op te halen met een compressietechnologie die een compressieratio van 100 beloofde. Ook hier waren er weer talrijke weldenkende lieden die de aandelen warm aanbevalen.

Het Australische bedrijf *Adams Platform Technology* wist een onafhankelijke test drie jaar voor zich uit te schuiven. Totdat de door ontevreden aandeelhouders afgedwongen test aantoonde dat het hier om gewone standaard compressietechnieken ging, welke de geclaimde factor van 200 uiteraard nooit konden halen.

De bovenstaande bedrijven wilden nooit iets loslaten over de details van hun techniek. Aandeelhouders werden warm gemaakt met teksten als: “*ZeoSync* intentionally randomises naturally occurring patterns” en wat dacht u van “the manipulation of binary information and translation to complex multidimensional mathematical entities”. Een gewaarschuwd mens telt voor twee (1 bit).

# Woordenlijst

<i>A/D-converter</i>	Zet analoog naar digitaal om. Bijvoorbeeld: een digitale microfoon zet een analoog geluidsignaal om naar een digitaal signaal van nullen en enen.
<i>Analoog</i>	Werkend met continue (traploze) fysische grootheden.
<i>Artefact</i>	Zichtbare beeldfouten, bijvoorbeeld aliasing of moiré.
<i>ASCII</i>	American Standard Code for Information Interchange. Internationaal tekencode die toetsenbord-input in digitale signalen omzet. De ASCII-code omvat alle letter- en getalstekens.
<i>Baud rate</i>	Snelheid van de data-overdracht via een modem. Gebruikelijke snelheden zijn 1200, 2400, 9600, 14.400, 28.800 Baud. Om de overdracht in bytes per seconde te berekenen moet het Baud-getal door acht gedeeld worden.
<i>Binary</i>	Getalsysteem dat gebaseerd is op twee waarden: het kent bijvoorbeeld alleen de toestanden 0 en 1.
<i>Bit</i>	Binary digIT. Kleinste informatie-eenheid in een computer, weergegeven door een 0 of een 1. Acht bits vormen één Byte.
<i>Bitmap</i>	Beeld dat uit pixels/beeldpunten bestaat die in een rechthoekig raster gelijkmatig verdeeld zijn.
<i>Buffer</i>	Geheugen in een computer om onregelmatige datastromen op te vangen. De inhoud gaat bij het uitschakelen verloren.
<i>Byte</i>	Standaardeenheid voor computerdata. Een Byte bestaat uit acht bits.
<i>CAD</i>	Computer Aided Design. Ontwerpprogramma dat onder andere gebruikt wordt door ingenieurs, architecten en grafisch ontwerpers. Beroemd voorbeeld is het AutoCAD-programma.
<i>CD-ROM</i>	Compact Disc Read Only Memory. Elektronisch opslagmedium met een capaciteit van ruim 600 MB.
<i>CMYK</i>	Subtractief kleursysteem dat op vier kleuren is gebaseerd: cyaan, magenta, geel en zwart.
<i>Compression</i>	Compressie. Het reduceren van de totale hoeveelheid bits door codering om de beschikbare geheugenruimte te vergroten of om de overdrachtssnelheid te verbeteren.
<i>Computer graphics</i>	Methoden en technieken voor het construeren, manipuleren, opslaan en weergeven van door een computer gegenereerde beelden.
<i>CPU</i>	Central Processing Unit. Centrale verwerkingseenheid. Het hart van de computer.

<i>CT</i>	Continuous Tone = halftoonbeeld. Een halftoonbeeld gebruikt één Byte voor elke RGB-waarde en kan dus 256 schakeringen per kleur tonen en meer dan 16 miljoen kleuren.
<i>D/A-converter</i>	Omzetten van digitaal naar analoog.
<i>Digital</i>	Numeriek. Data zijn gecodeerd in het binaire getsysteem.
<i>Digital Imaging</i>	Het creëren van een digitale kopie van een illustratie of foto.
<i>Digital Photography</i>	Digitale fotografie. Het fotograferen van beelden met een digitale camera, waardoor de beelden rechtstreeks digitaal opgeslagen worden in de camera.
<i>Dither</i>	Techniek die gebruikt wordt bij reductie van het aantal kleuren van een afbeelding. Door, op een vooraf in te stellen wijze, bepaalde kleurwaarden toe te kennen aan pixels, wordt getracht de afbeelding een natuurlijker aanzicht te geven.
<i>DOS</i>	Disk Operating System. Besturingssysteem voor een computer. Gebruikelijker is tegenwoordig OS, Operating System.
<i>Dot</i>	Kleurpunt. Een dot heeft een vaste kleurdichtheid bij variabele grootte.
<i>DPI</i>	Dots per inch. Maateenheid voor de resolutie van gedrukt materiaal.
<i>DTP</i>	Desk Top Publishing. Het prepareren en zetten van tekst en beeld op de computer.
<i>EPS</i>	Encapsulated Post Script (File). Een vectorgeoriënteerd grafisch bestandsformaat ontwikkeld door Adobe Systems. EPS heeft vaak de voorkeur voor vele computer-illustraties, vanwege zijn efficiënt gebruik van het geheugen en goede kleurbeheersing.
<i>Ethernet</i>	Net voor snel datatransport via coaxkabels.
<i>Font</i>	Een compleet assortiment van een bepaalde grootte van een lettertype. Hierin inbegrepen zijn meestal hoofdletters, interpunctietekens et cetera.
<i>Frame</i>	Eén beeldje van een video of TV-uitzending.
<i>GIF</i>	Graphic Interchange Format. Een 8-bits (256 kleuren of grijswaarden) of minder computerbestandsformaat.
<i>Huffmancodering</i>	De Huffmancodering is een verliesloos algoritme voor datacompressie, genoemd naar David Huffman die hem in 1952 voor het eerst beschreef. Het principe is eenvoudig: gegeven een serie symbolen, waarvan er sommige meer kans hebben voor te komen dan andere, is het mogelijk om voor de veel voorkomende symbolen een kortere code te kiezen dan voor de weinig voorkomende.

<i>Interface</i>	In principe alles wat de communicatie tussen onderdelen van een computer of tussen de gebruiker en de computer bepaalt of vorm geeft. Een apparaat kan met het moederbord verbonden zijn door middel van een interfacekaart, maar ook het bureaublad van bijvoorbeeld Windows XP is een interface tussen gebruiker en computer.
<i>JPEG</i>	Joint Photographic Electronic Group. Standaard voor het comprimeren van afbeeldingen. Bij het comprimeren gaat informatie verloren.
<i>Kilobyte</i>	KB, 1024 bytes
<i>Laser</i>	Acroniem voor "Light Amplification by Stimulated Emission of Radiation". Een laser is een intense straal licht met een zeer kleine bandbreedte.
<i>LZW Algoritme</i>	Het LZW of Lempel-Ziv-Welch algoritme is een exact omkeerbaar (lossless) compressie-algoritme dat door Abraham Lempel, Jacob Ziv en Terry Welch is uitgevonden. Het algoritme werkt volgens het principe dat alle veelvoorkomende tekenreeksen worden vervangen door een aparte code. Deze bibliotheek van veelvoorkomende reeksen levert een flinke ruimtebesparing op. Het ZIP compressie-programma is ongetwijfeld een van de bekendste toepassingen.
<i>Mac</i>	Afkorting voor Macintosh, veel gebruikte computer in de grafische wereld.
<i>Mass storage</i>	Opslagmedia bedoeld voor de opslag van grote hoeveelheden gegevens.
<i>Matrix</i>	Hulpmiddel voor het vereenvoudigen van weergave en berekening van meerdimensionale verhoudingen tussen data die in veelvoudige afhankelijkheid van elkaar staan.
<i>MB</i>	Megabyte, 1024 KB
<i>Modem</i>	Modulator-demodulator. Apparaat dat de overdracht van data tussen twee computers over een telefoonlijn mogelijk maakt.
<i>Monochrome</i>	Eenkleurig, zwart-wit
<i>MPEG</i>	"Moving Picture Experts Group". Compressiemethode voor bewegend beeld.
<i>PAL</i>	"Phase Alternation Line". Standaard TV-systeem in Europa (met uitzondering van Frankrijk) en Australië met 625 lijnen en op 50Hz.
<i>PDF</i>	"Portable Document File". Een bestandsformaat ontwikkeld door Adobe voor het overzetten van ontwerpen van het ene naar het andere computerplatform.
<i>Pixel</i>	Het kleinste weer te geven element op een computerscherm dat onafhankelijk gecodeerd kan worden voor kleur en intensiteit

<i>Pixeldiepte</i>	Hoeveelheid data die gebruikt wordt om een pixel op een computerscherm te beschrijven. Monochroom is 1 bit, grijswaarden zijn 8 bits, RGB is 24 bits en CMYK dient 32 bits diep te zijn.
<i>Postscript</i>	Adobe's paginabeschrijvingstaal, die afbeeldingen en teksten zo vertaald dat ze op verschillende outputapparaten uitgevoerd kunnen worden.
<i>RAM</i>	"Random Access Memory". Vluchtig werkgeheugen van een computer.
<i>Run Length Encoding (RLE)</i>	Een van de simpelste algoritmen van lossless datacompressie is de RLE methode. Het is gebaseerd op de een simpele datacodering die wel zeer intuïtief is. De string DAAAAAREEEEE wordt bijv. : D*A5R*E5
<i>TIFF</i>	"Tagged Image File Format". Veelgebruikt beeldbestandsformaat.
<i>Vector</i>	Een geordende set van reële getallen waarvan ieder getal een afstand weergeeft op een coördinatenas.
<i>ZIP formaat</i>	Zie LZW algoritme



*Alle conclusies en stellingen zijn persoonlijke gevolgtrekkingen op basis van de eigen kennis en ervaring van de auteur. En hebben niet de pretentie wetenschappelijke bewijzen te zijn. De auteur aanvaardt geen enkele aansprakelijkheid voor de juistheid of volledigheid van de verstrekte informatie.*

Document Versie 1.2

Copyright © 2004

**Keyword Info Systems BV.**

Postbus 677  
2600 AR Delft

ISBN 90 77553 02 9