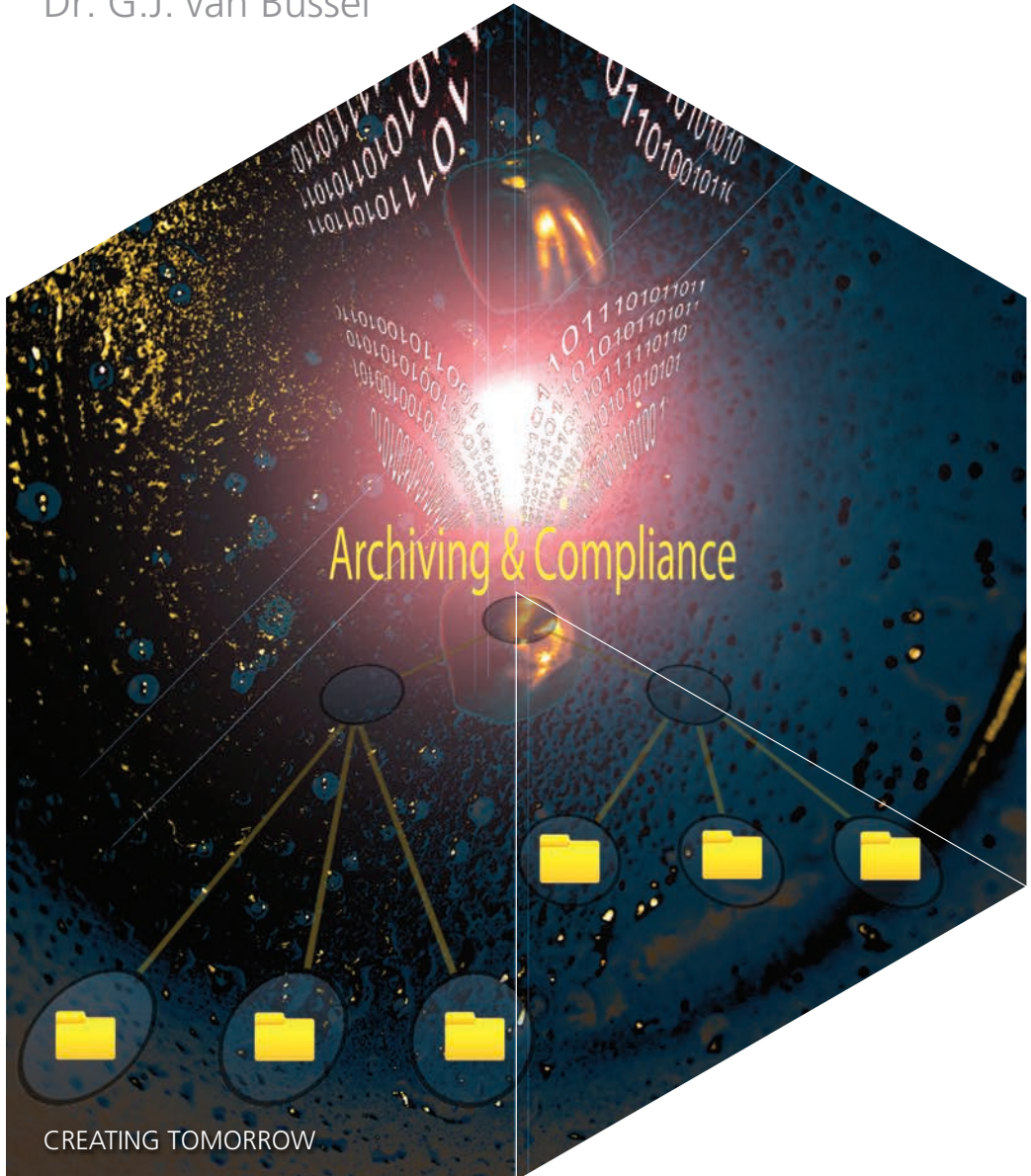


'ARCHIVING SHOULD BE JUST LIKE AN APPLE™' EN ACHT ANDERE, NUTTIGE (?) STELLINGEN

Dr. G.J. van Bussel



‘Archiving should be just like an
Apple^(tm)’, en acht andere,
nuttige (?) stellingen

‘Archiving should be just like an
Apple^(tm)’, en acht andere,
nuttige (?) stellingen

Lectorale Rede

in verkorte vorm uitgesproken op dinsdag 16 oktober 2012

door

dr. G.J. van Bussel

bijzonder lector Digital Archiving & Compliance
aan de Hogeschool van Amsterdam



Hogeschool van Amsterdam

Omslagillustratie: foto en collage van Bert Zuiderveen.nl
Vormgeving omslag: Kok Korpershoek, Amsterdam
Opmaak binnenwerk: JAPES, Amsterdam

ISBN 978 90 5629 721 3
e-ISBN 978 90 4851 826 5 (pdf)
e-ISBN 978 90 4851 827 2 (ePub)

© G.J. van Bussel / HvA Publicaties, Amsterdam 2012

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikel 16B Auteurswet 1912 j° het Besluit van 20 juni 1974, Stb. 351, zoals gewijzigd bij het Besluit van 23 augustus 1985, Stb. 471 en artikel 17 Auteurswet 1912, dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) dient men zich tot de uitgever te wenden.

Mevrouw de rector, collega's uit praktijk en wetenschap, studenten, vrienden, familie en andere belangstellenden,

Het wordingsproces van deze rede was kort en intensief. Het was een worsteling met een structuur die niet de expressie wilde worden van het denkbeeld dat ik had. Totdat ik mij afvroeg of wat ik wilde vertellen niet samen te vatten was in een aantal stellingen. Dat bleek een werkbare en interessante benadering, waarmee het mogelijk was mijn ideeën en wensen op een schijnbaar ongestructureerde manier toch te structureren.

Het lectoraat dat ik hier mag presenteren is erg jong. Dat betekent dat mijn kenniskring en ik nog druk doende zijn om onze onderzoeksagenda te definiëren, projecten te initiëren en geïnteresseerde partijen aan ons te verbinden. We zijn druk bezig om onze aanwezigheid in en beschikbaarheid voor de opleidingen van (vooral) het domein Media, Creatie en Informatie (MCI) vorm te geven. Om onze activiteiten te volgen of om daarin te participeren als sponsor of als een geassocieerd lid van mijn kenniskring, is er een website in het leven geroepen, te vinden onder www.digitalarchiving.nl. In de komende maanden zullen we ons presenteren op een aantal sociale media, zoals LinkedIn, Twitter en Facebook. Enkele studenten gaan een strategie daarvoor vormgeven.

Verschillende van mijn collega-lectoren hebben in hun redes aandacht besteed aan het onderzoek binnen een hogeschool en de daarbij gebruikte methodologie. Ze hebben alles gezegd wat ik daarover zou kunnen zeggen en ik ga dat dan ook niet herhalen.¹ Ik richt mij hier volledig op het themagebied van mijn bijzonder lectoraat: *Digital Archiving & Compliance*. Ik wil vooral laten zien dat het een breed en zeer uitdagend werkveld is, dat mogelijkheden biedt voor vele onderzoeks- en onderwijsactiviteiten. In de komende onderzoeksagenda zal duidelijk worden waarop wij ons binnen deze mogelijkheden vooral gaan richten.

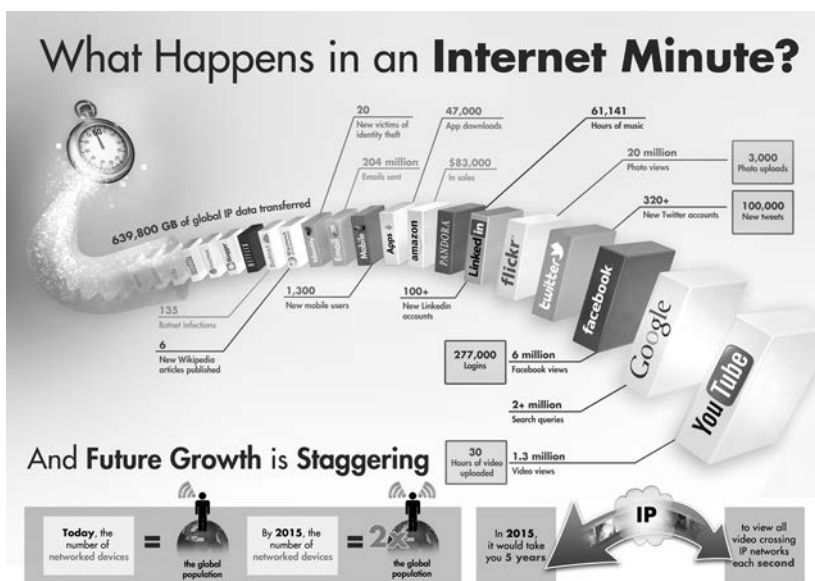
1. 'Het maatschappelijke belang van Archiving & Compliance is enorm'

*Digital archiving*² richt zich, kortweg, op het ontstaan, de bewerking, het gebruik, de terugvindbaarheid en het behoud van betrouwbare informatie.³ Aan-

gezien bijna alle informatie vandaag de dag digitaal is gegenereerd (*born digital*) of digitaal is gemaakt (*made digital*), zal ik in het vervolg van deze rede de aanduiding *digital* veelal niet gebruiken. Het tweede thema, *compliance*⁴, wil de naleving van wet- en regelgeving, normen en richtlijnen aantonen met betrouwbare informatie, zodanig dat verantwoording afgelegd kan worden aan 'legitieme fora' en dat bewijs geleverd kan worden van rechten en plichten en van de uitvoering en afhandeling van bedrijfsprocessen, taken en transacties.⁵ De kern van zowel *archiving* als *compliance* is de informatiewaardeketen: de procesketen die de creatie, de vastlegging, de opslag, de bewerking, de structurering, de publicatie, het gebruik, het behoud, de toegankelijkheid, de distributie, de waardering, de vernietiging, de beveiliging en de toetsing van betrouwbare informatie realiseert.⁶ De processen van deze waardeketen zijn gericht op het *besturen* van de informatiestromen in organisaties met de bedoeling de *betrouwbaarheid* van informatie door juridische, organisatorische en technische maatregelen te waarborgen in de tijd. Ook het digitale culturele erfgoed, dat aangewezen is om ten eeuwigden dage bewaard te blijven, valt onder de werking van deze processen. De informatiewaardeketen probeert in de realisatie van betrouwbare informatie vier dimensies optimaal te regelen: de kwaliteit, de context, de relevantie en het voortbestaan van informatie. Alle processen van de waardeketen zijn daarop gericht.⁷

Het maatschappelijke belang van beide thema's is groot. Ik wijs daarvoor op drie ontwikkelingen. De eerste ontwikkeling is *de toename van de hoeveelheid informatie* in de maatschappij, door de futurist Alvin Toffler al in 1970 voorspeld.⁸ Deze 'informatie-explosie' bedreigt (waarschijnlijk) de opslagmogelijkheden, (zeker) de toegankelijkheid en de vindbaarheid, en (mogelijkerwijze) de begrijpelijkheid van informatie. Er wordt grote druk gelegd op de beschikbare IT-mogelijkheden om betrouwbare en onbetrouwbare informatie van elkaar te kunnen scheiden, om informatie te beveiligen en om privacy te bewaren.⁹ De informatiemassa groeit door de toenemende publicatiemogelijkheden die het internet biedt, de exploderende elektronische communicatie door de inmiddels revolutionaire impact van *social media*, de mobiele revolutie, de toename van *Big Data*, de onomkeerbare trend naar volledig digitaal werken (het 'Nieuwe Werken'), de grootschalige digitalisering van cultureel erfgoed zoals film, muziek, kunst, beeld, kaartmateriaal, etc., etc.¹⁰ Het kwantificeren van die groei is niet zo eenvoudig. Er zijn verschillende analyses uitgevoerd, op basis van verschillende definities en onderzoeksmethoden. Al deze analyses zijn het (ondanks verschillende resultaten) eens over één fundamenteel feit: de hoeveelheid informatie in de wereld groeit in een verbazingwekkend tempo. Zo stelden Lyman en Varian vast dat er in 2002 wereldwijd 5 exabytes (5 miljard gigabytes) aan nieuwe data werd opgeslagen en dat er ongeveer 18 exabytes

werden gecommuniceerd (maar niet noodzakelijkerwijze opgeslagen) via elektronische kanalen als telefoon, radio, televisie en internet. De hoeveelheid nieuwe, opgeslagen data groeide met gemiddeld 25% per jaar.¹¹ Het onderzoeksbureau IDC stelde in 2007 vast dat de jaarlijks gecreëerde hoeveelheid data de wereldwijd beschikbare opslagcapaciteit te boven ging; het verschil tussen beide grootheden werd in de jaren daarna steeds groter. In 2009 werd er volgens hun onderzoek 800 exabytes aan data gecreëerd en gerepliceerd. Tot 2020, zo schat IDC, is een jaarlijks groeipercentage van 40% te verwachten.¹² Hilbert en López constateerden in 2011 dat de beschikbare opslagcapaciteit vanaf 1986 met een jaarlijks gemiddelde van 23% groeide (tot plusminus 290 exabytes in 2007), maar dat de verwerkingscapaciteit met 58% per jaar groeide. Die constatering verklaart de snelle opmars van de digitalisering: in 2000 was nog maar ongeveer 25% van de data opgeslagen in digitale vorm, in 2007 was dat al gestegen tot 94%.¹³ Ondanks de huidige economische crisis stijgen de behoeften en uitgaven aan ‘archiveringssoftware’.¹⁴ Als voorbeeld laat ik in figuur 1 zien wat er in één minuut op internet gebeurt. Het toegankelijk houden van de informatiemassa is maatschappelijk van belang; de informatiewaardeketen speelt hierbij een grote rol.



Figuur 1 Infographic over de informatie-explosie. Bron: Intel Free Press Tech News, Creative Commons, 2011

Een tweede ontwikkeling is de *wereldwijd toenemende regulering ter vergroting van organisatorische transparantie*. Organisaties worden door steeds meer wet- en regelgeving opgedragen publiek verantwoording af te leggen over hun handelen. Dit toenemend aantal regels, richtlijnen en normen heeft diverse oorzaken: schandalen bij beursgenoteerde bedrijven (Enron, WorldCom, WorldOnline, Barings, Parmalat, etc.), wangedrag in het bancaire en financiële systeem, de (bijna) ineenstorting daarvan, toenemende fraude en (grotendeels veroorzaakt door al het voorgaande) de wens naar een grotere transparantie van het handelen van publieke en private organisaties.¹⁵ Implementatie van nieuwe regelgeving brengt (hoge) investeringskosten met zich mee. Onderzoek van het Ponemon Institute uit 2011 onder 46 verschillende multinationals toonde dat de totale *compliance*-gerelateerde kosten bij deze bedrijven ongeveer 162 miljoen dollar bedroegen. Daartegenover: de kosten van *non-compliance* zouden ongeveer 431 miljoen dollar bedragen hebben.¹⁶ De investeringen worden mede veroorzaakt door de groeiende hoeveelheid informatie, waardoor het moeilijker is die informatie te filteren die nodig is om *compliance* aan deze toenemende regulering aan te tonen.¹⁷ Het blijkt ook dat het zich concreet kunnen verantwoorden voor veel organisaties problematisch is, wat afbreuk doet aan het vertrouwen van de klant en de burger.¹⁸ Het is verbazend, gezien het verschil tussen de kosten van *compliance* en *non-compliance*, dat er nauwelijks onderzoek is naar mechanismen die organisaties helpen bij het nakomen van wetten, regels, richtlijnen en normen en in het vergroten van organisatorische transparantie.¹⁹ Een van die mechanismen is de informatiewaardeketen. Deze keten zorgt er immers voor dat de voor aantoonbare *compliance* noodzakelijke informatie betrouwbaar, opgeslagen, beveiligd, toegankelijk en beschikbaar is, kan worden geïdentificeerd, kan worden gereconstrueerd in de toekomst, en wordt verwijderd als dat volgens de bestaande wet- en regelgeving mag.²⁰ *Archiving* speelt hier een maatschappelijke rol van betekenis.

De derde ontwikkeling is *de maatschappelijke verwachting* (tot uitdrukking gebracht overigens in veel regelgeving) *dat organisaties zo 'groen' of milieuvriendelijk mogelijk functioneren*. In een van mijn columns op mijn eigen website besteedde ik in april 2008 al aandacht aan deze problematiek.²¹ Informatietechnologie heeft op verschillende manieren gevolgen voor onze leefomgeving. Iedere fase in het leven van een computer schept milieuproblemen. De productie van computers en hun (niet-)elektronische componenten vergt elektriciteit, grondstoffen, chemische materialen en grote hoeveelheden water, en levert daarnaast (vaak giftige) afvalstoffen op. Computers en randapparatuur worden twee of drie jaar na aankoop vervangen. Deze afgedankte apparaten worden nog steeds (zij het ver van huis) gedumpt op afvalbergen, waar de aarde wordt vergiftigd en het water ernstig wordt vervuild.²² Daar-

naast groeit het totale energieverbruik voortdurend, met en zonder gebruik van 'groene' energiebronnen, ondanks alle maatregelen die IT-leveranciers genomen hebben om het energieverbruik in te dammen, zoals de consolidatie en virtualisatie van servers. Als voorbeeld: van 2000 tot 2005 verdubbelde het elektriciteitsverbruik van datacenters, terwijl wereldwijd het *totale* elektriciteitsverbruik slechts met 16,7% per jaar groeide.²³ Van 2005 tot 2010 steeg het verbruik in datacenters met 56%; in 2012 alleen wordt een stijging van 19% verwacht.²⁴ Dat is een bevestiging van het feit dat pas na 2000 de digitalisering werkelijk is ingezet, zoals Hilbert en López stellen.²⁵ Het is tevens de reden waarom elke computer ieder jaar meer dan een ton koolzuurgas genereert.²⁶ Opslagnetwerken waren in 2008 verantwoordelijk voor 15% van de totale IT-energiekosten.²⁷ Mijn inschatting is dat dit percentage vier jaar later ongeveer verdubbeld is, gezien de stijgende behoefte aan opslag de afgelopen jaren door de verveelvoudiging van informatie, de neiging van gebruikers om meer informatie nuttig te achten en de angst niet *compliant* te zijn aan allerlei wetten en regels. Door de consolidatie van opslagsservers en door gebruik te maken van andere, minder elektriciteit vergende opslagmedia zoals tape, kan het energieverbruik worden teruggebracht.²⁸ De informatiewaardeketen kan bij de inrichting van dit soort 'groene' opslagnetwerken behulpzaam zijn (door de processen 'opslag' en 'behoud'), aangevuld met het managen van de opgeslagen hoeveelheid informatie zelf (via 'waardering' en 'vernietiging'). *Archiving* speelt een rol in het terugdringen van opslagcapaciteit door de informatiemassa rigoureuus te saneren.

Consolidatie en virtualisatie

Vaak worden de termen consolidatie en virtualisatie vermengd.

Met *consolidatie* worden servers en opslagsystemen waar mogelijk fysiek bijeengebracht en als één geheel beheerd. De applicaties worden efficiënt verdeeld. Hiermee is een optimale benutting van capaciteit en een hoge beschikbaarheid te realiseren. Het leidt tot minder systemen, minder vervanging, minder afschrijvingen en eenvoudiger beheer. Deze benadering zorgt ervoor dat het aantal opslagmogelijkheden wordt beperkt, net als de hoeveelheid benodigde energie, de geproduceerde hitte en (als een interessant neveneffect) de operationele en beheerskosten van back-ups, archiefopslag, e.d.

Virtualisatie leidt tot het beheren van minder fysieke machines en minder besturingssystemen. Door het inzetten van virtualisatie kunnen meerdere virtuele omgevingen met verschillende besturingssystemen geïsoleerd van elkaar werken op een fysieke machine. Elke virtuele omgeving heeft zijn eigen set aan virtuele hardware (schijfruimte, geheugen, processoren, netwerkkaart etc.), waar een besturingssysteem en applicaties op werken. Volledige omgevingen kunnen zonder onderbreking verplaatst worden van de ene fysieke machine naar de andere.

Het maatschappelijke belang van *Archiving & Compliance* is evident. De eerste twee ontwikkelingen dwingen organisaties tot het (volledig) geautomatiseerd uitvoeren van de informatiewaardeketen. De derde ontwikkeling dwingt ertoe dat zo milieuvriendelijk mogelijk te doen. De competenties van studenten dienen daarop te worden afgestemd. Studenten moeten zich ervan bewust zijn dat voor alle organisaties waarvoor zij uiteindelijk gaan werken, de hele informatiewaardeketen cruciaal is. Dat betekent dat zowel in onderwijs als in onderzoek deze waardeketen ‘rode draad’ dient te zijn. Het onderzoek van het bijzonder lectoraat richt zich op de realisatie van *fully automated archiving environments*, omgevingen waarin de informatiewaardeketen volledig geautomatiseerd wordt uitgevoerd. Daarbinnen zal het thema *Green Archiving* een belangrijke plek innemen, gezien de relevantie daarvan voor maatschappelijk verantwoord ondernemen. In het informatie- en archiefwetenschappelijk onderzoek zijn deze thema’s grotendeels genegeerd.²⁹

2. ‘Archiving is een doelgerichte organisatorische asset’

Het is een enerverende, dynamische tijd. Gangbare opvattingen, methoden, oplossingen en regels worden ter discussie gesteld. Nieuwe mogelijkheden voor kennisverspreiding, productcreatie en communicatie manifesteren zich.³⁰ De concepten ‘informatiemaatschappij’ en ‘kennismaatschappij’ zijn gemeengoed geworden om die dynamiek te karakteriseren. Beide concepten zijn in definiëring tegenstrijdig en vaag.³¹ De essentiële kern van deze concepten – de transformatie naar een nieuw sociaal systeem door de toename en de veranderende rol van informatie – lijkt zich niet te manifesteren. Er doen zich geen radicale wijzigingen voor in de wijze waarop onze postindustriële maatschappij is georganiseerd en in de richting waarin deze maatschappij zich ontwikkelt.³² Desalniettemin doen zich wél grote veranderingen voor in wet- en regelgeving, in de verschijningsvorm en de impact van informatietechnologische

structuren en mogelijkheden, in de wijze van communicatie en de daarvoor gebruikte technologische middelen, in de manieren waarop organisaties functioneren, opereren, samenwerken en zichzelf structureren, in de wijzen waarop informatie moet worden beveiligd en in de manieren waarop mensen economisch en sociaal functioneren.³³

Door alle veranderingen die zich voordoen, reorganiseren organisaties zichzelf continu. Het afgelopen decennium zijn ze druk in de weer geweest met de *re-engineering* van hun werkprocessen. Ze hebben (meer of minder succesvol) hun vaak niet-geïntegreerde software vervangen door Enterprise Information Systems (EIS), systemen die de sleutelprocessen van een organisatie integreren in één softwareomgeving. Hierdoor kunnen gestructureerde gegevens in *real time* en zonder organisatorische blokkades in die sleutelprocessen worden verwerkt. Coördinatie, efficiency en besluitvorming zijn hierdoor verbeterd, net als de documentatie van afgehandelde procestransacties en de informatiekwaliteit.³⁴ Informatie is nodig voor de verbetering van de performance van organisaties en is een onmisbare *asset* voor het behalen van concurrentievoordelen.³⁵ De tijdige beschikbaarheid van betrouwbare informatie is van belang om te kunnen reageren op de onzekerheid die het gevolg is van onverwachte gebeurtenissen en veranderingen in de hoog-complexe maatschappelijke omgeving van organisaties.³⁶

Integratie brengt echter ook problemen die op zijn minst de positieve effecten temperen, zoals het feit dat databases niet goed in staat zijn ongestructureerde informatie op te nemen.³⁷ Ongestructureerde informatie, zoals tekst, geluid en beeld, beslaat echter wel ongeveer 90% van alle informatie die in organisaties wordt gegenereerd.³⁸ En dat terwijl gestructureerde én ongestructureerde informatie gebruikt worden in beleid, besluitvorming, producten, acties en transacties, en als zodanig cruciaal zijn voor de performance van bedrijfsprocessen. Zonder deze informatie wordt uitvoering van bedrijfsprocessen of het bereiken van doelstellingen volstrekt onmogelijk.³⁹ Daarnaast: informatie is bewijs voor gevormd en uitgevoerd beleid, gedane besluitvorming, vervaardigde producten, nagestreefde doelstellingen, uitgevoerde acties en afgehandelde transacties. Zoals al is benadrukt eisen publieke en private autoriteiten steeds meer dat deze informatie betrouwbaar is en dat op basis daarvan beleid, besluitvorming, producten, acties en transacties kunnen worden gereconstrueerd. De informatiewaardeketen wordt van grotere organisatorische betekenis.

Dat informatie van organisatiestrategische betekenis is, wordt niet betwijfeld.⁴⁰ Juist daarom ook hebben organisaties controlesystemen in het leven geroepen die waarborgen dat het management van informatie bijdraagt aan de gestelde doelstellingen.⁴¹ Deze controlesystemen blijken echter onevenwichtig,

doordat ze zich vooral concentreren op gestructureerde informatie. Dat heeft gevolgen voor het bereiken van de gestelde doelstellingen en voor de verbetering van de performance van organisaties, zowel qua doelmatigheid als qua rechtmatigheid.⁴² Barata en Cain hebben onmiskenbaar aangetoond dat zonder betrouwbare informatie als bewijs voor beleid, besluitvorming, producten, acties en transacties het bereiken van de doelstelling 'rechtmatigheid' (door middel van het afleggen van verantwoording en het aantonen van *compliance*) onmogelijk is.⁴³ Ik durf wel de stelling aan dat betrouwbare informatie ook onmisbaar is om de doelstelling 'doelmatigheid' te realiseren. Het is vreemd dat de informatiewaardeketen in veel organisaties slechts marginaal, onevenwichtig of onvolledig is geïmplementeerd, wat ertoe heeft geleid dat van een groot deel van de in een organisatie beschikbare informatie niet bekend is of zij betrouwbaar is. Als het ongestructureerde informatie betreft, is de kwaliteit vaak zeer twijfelachtig: niet toegankelijk, niet (tijdig) beschikbaar, niet compleet, niet relevant, niet accuraat en/of niet begrijpelijk. Van zowel gestructureerde als ongestructureerde informatie zijn de oorsprong en de contextuele omgeving niet of nauwelijks bekend.⁴⁴ Het feit dat informatie over het algemeen de informatietechnologie waarmee zij is gegenereerd en waarin zij wordt gemanaged overleeft, stelt (zeker met het toenemen van de hoeveelheid informatie) de betrouwbaarheid van die informatie verder ter discussie.⁴⁵ Corrupte informatie maakt het bereiken van organisatorische doelstellingen problematisch.

Problemen in informatiekwaliteit

Bedrijven zijn afhankelijk van systemen. In veel gevallen klopt er weinig van de gegevens waarmee deze systemen werken. Het gevolg is dat ondernemingen sturen op foutieve informatie. '80% van onze klanten heeft problemen met hun datakwaliteit', zegt Vincent Wylenzek, Business Intelligence Consultant bij Ordina, op basis van zijn ervaring met het opschonen van gegevens.

Volgens Wylenzek is niet ieder geval even ernstig, maar fouten in administraties kunnen leiden tot verkeerde financiële verslaglegging. Uiteindelijk kan dit zelfs leiden tot problemen met de fiscus of verkeerde bedrijfsvoering, omdat conclusies niet deugen. De consultant zag ooit een systeem niet gebruikt worden, omdat het personeel geen enkel vertrouwen in de geleverde informatie had.

‘Er zijn diverse soorten vervuiling te bedenken: regels die dubbel opgeslagen zijn, datatypes die niet overeenkomen of percentages waar een komma op de verkeerde plek staat’, weet Wylenzek. Dat komt ook, omdat door personeelsverloop kennis van systemen onbereikbaar is geworden voor het bedrijf. Een van de belangrijkste problemen is het gebruik van spreadsheets voor dataverwerking. Spreadsheets zijn zelden voldoende beschermd tegen het maken van typefouten, die vervolgens in de organisatie over verschillende systemen worden verspreid. ‘Wat je ziet is dat negen van de tien bedrijven Excel gebruiken als primair bronsysteem voor dit soort gegevens, zoals begrotingen of targets’, verzucht Wylenzek. ‘We noemen dat “Excelites” als ziektebeschrijving’. Als oorzaak duidt hij tekortkomingen van de gebruikte systemen aan, omdat deze bepaalde informatie niet gestructureerd verwerken.

Een goede stap is om gegevens af te wijzen als ze niet correct zijn.

Bron: <http://computerworld.nl/article/1234/80-bedrijven-heeft-data-kwaliteitsproblemen.html>.

Archiving is belangrijk voor het bereiken van *alle* organisatiedoelstellingen. Het vormt een van de belangrijkste waarborgen van het *organizational memory*, het DNA van een organisatie, zoals het door Spear en Bowen is benoemd.⁴⁶ Het is een metaforisch concept dat de opslag, de representatie en het delen van informatie, cultuur, macht, praktijken en beleid beschrijft. In de organisatiekundige literatuur worden de doelen van dit ‘geheugen’ gezien als het terugbrengen van de kosten voor de afhandeling van bedrijfstransacties, het versnellen van de toegang tot praktijken en oplossingen uit het verleden, het assisteren in de besluitvormingsprocessen door het leveren van referentiemateriaal, het delen van kennis en het verminderen van de afhankelijkheid van individuele medewerkers als kennisdragers.⁴⁷ Volgens Weicks ‘*The social psychology of organizing*’ zijn het gebruik en de verspreiding van dit *organizational memory*, de accuraatheid en compleetheid ervan, en de omstandigheden waarin het is gegenereerd (de context) van groot belang voor organisaties om te kunnen ‘leren’ en zo hun doelen beter te kunnen bereiken. De betrouwbare reconstructie van het verleden is daarvoor noodzakelijk.⁴⁸ De informatiewaardeketen is hiervoor een organisatorische noodzakelijkheid.

Archiving is voor organisaties een belangrijke *asset* voor het bereiken van organisatorische doelstellingen zowel qua doelmatigheid als qua rechtmatigheid. Zonder *archiving*, zonder een goed functionerende en geïmplementeerde

informatiewaardeketen, is het bereiken van deze doelstellingen qua performance ondermaats.⁴⁹ *Archiving is goal-oriented*, het is een noodzakelijke functionaliteit voor het bereiken van organisatorische doelen. Dat is iets waarvan docenten en studenten van deze hogeschool zich bewust dienen te zijn.

3. 'De sleutel voor organisatorische informatie-infrastructuren is de informatiewaardeketen'

De organisatorische noodzakelijkheid van *Archiving & Compliance* zal duidelijk zijn. Organisaties dienen informatie, om een modewoord te gebruiken, goed te managen. Aangezien de informatiewaardeketen in vele organisaties slechts marginaal, onevenwichtig of onvolledig is geïmplementeerd, lijkt dat 'goede management' niet overal doorgedrongen te zijn. Het is onbekend of de beschikbare informatie betrouwbaar is en of betrouwbare reconstructies van het verleden mogelijk zijn.⁵⁰ De processen van de informatiewaardeketen kunnen de genoemde vier dimensies van (gestructureerde én ongestructureerde) informatie niet waarborgen: de kwaliteit, de context, de relevantie en het voortbestaan.⁵¹ Dit creëert onzekerheid bij de uitvoering van werkzaamheden en beïnvloedt het bereiken van de organisatiedoelstellingen negatief. Deze onzekerheid wordt versterkt door de effecten van de geconstateerde toevloed aan informatie en de grotere druk op organisaties om zichzelf te verantwoorden. Het falen van organisaties om de informatiewaardeketen volledig en evenwichtig te implementeren en betrouwbare informatie te genereren, leidt tot het zichtbaar worden van een veronderstelde 'onzichtbaarheid': de altijd aanwezige achtergrond van de werkzaamheden die worden uitgevoerd, de organisatorische informatie-infrastructuur. Deze infrastructuur 'becomes visible upon breakdown', aldus Star en Ruhleder, waarbij het de werkzaamheden in een organisatie (volledig) verstoort.⁵² De consequenties van een langdurige storing kunnen catastrofaal zijn voor het bereiken van de organisatorische doelstellingen doelmatigheid en rechtmatigheid.⁵³ Het is niet voor niets dat deze complexe entiteit ongeveer 50% vergt van het informatietechnologisch budget van een gemiddelde organisatie.⁵⁴ Toch is dat niet voldoende om de informatiewaardeketen te doen functioneren.

Het concept 'informatie-infrastructuur' heeft geleid tot een andere manier van denken. Niet langer stond het systeem centraal, maar het netwerk van systemen dat door organisaties gebruikt werd om doelstellingen te realiseren.⁵⁵ Vanaf 1990 is er veel onderzoek gedaan naar deze infrastructuren. Naar de technologische convergentie en de implicaties daarvan voor het strategisch management van organisaties bijvoorbeeld⁵⁶, naar de groei en dynamiek van

wetenschappelijke infrastructuren⁵⁷ en naar de effecten van de *actor-network*-theorie op informatie-infrastructuren.⁵⁸ Het is niet verwonderlijk dat er vele pogingen gedaan zijn om informatie-infrastructuren te definiëren. Deze definities zijn onder te brengen in twee verschillende groepen. De eerste groep omschrijft wat informatie-infrastructuren *zijn*. Een voorbeeld: Hanseth en Lyytinen definiëren een informatie-infrastructuur als ‘a shared, open (and unbounded), heterogeneous and evolving socio-technical system (which we call installed base) consisting of a set of IT capabilities and their user, operations and design communities’.⁵⁹ De tweede groep definities omschrijft wat informatie-infrastructuren *doen*. Een voorbeeld is Pironti, die een infrastructuur ziet als een geheel van mensen, processen, procedures, tools, middelen en technologie voor de creatie, het gebruik, de communicatie, de opslag en de vernietiging van informatie.⁶⁰ Informatie-infrastructuren zijn er volgens deze omschrijving om organisaties te faciliteren bij het realiseren van de informatiewaardeketen. Mensen, processen, procedures, tools, middelen, informatie- en communicatietechnologie, methoden en kennis (de componenten binnen de informatie-infrastructuur) worden ingezet en gebruikt met het doel het genereren van betrouwbare informatie en het maken van betrouwbare reconstructies van het verleden mogelijk te maken. Zoals ik al heb gemeld: daartoe zijn infrastructuren (nog) niet of nauwelijks in staat, zeker niet als het ongestructureerde informatie betreft. Ook in de groei en ontwikkeling van informatie-infrastructuren namelijk, lag de nadruk sterk op gestructureerde informatie.

Informatie-infrastructuren zijn in de loop der tijd ontwikkeld en verbeterd tot cumulatieve technische, menselijke en organisatorische constructies, waardoor ze als sociotechnische netwerken te beschouwen zijn.⁶¹ Ze zijn gebaseerd op een bestaand conglomeraat van infrastructurele componenten, dat de neiging heeft uiterst traag te reageren en te bewegen (‘the inertia of the installed base’).⁶² Deze sociotechnische netwerken kunnen, zoals Bygstad betoont, als een organisatorische innovatie gezien worden, als een nieuwe organisatorische vorm⁶³, als een ‘computer mediated organization’.⁶⁴ Het ontwerp van organisaties en technologie wordt daarbij volledig geïntegreerd, zoals Beniger al in 1986 stelde.⁶⁵ Een nieuwe organisatorische vorm botst altijd met de traditionele vorm, omdat de organisatorische karakteristieken niet parallel lopen. Traditionele organisatorische vormen zijn gebaseerd op eigendom, gedefinieerde grenzen en hiërarchische, verticale controle. Informatie-infrastructuren zijn als organisatorische vorm open, onbegrensd, gedecentraliseerd, gebaseerd op standaarden en horizontaal. De overeenkomsten zitten in de organisatorische structuren, coördinatiemechanismen en gedeelde doelstellingen. De reden voor het falen van informatie-infrastructuren in het genereren van betrouw-

bare informatie en het maken van betrouwbare reconstructies van het verleden zit, denk ik, vooral in deze botsing tussen oude en nieuwe organisatorische vormen, waardoor verschillende interpretaties bestaan van zoiets cruciaals als de informatiewaardeketen. De aandacht voor gestructureerde informatie en het lange tijd grotendeels negeren van de organisatiedoelstelling ‘rechtmatigheid’ heeft ertoe geleid dat de informatiewaardeketen voor (vooral) ongestructureerde informatie in informatie-infrastructuren ongecontroleerd is gelaten, met negatieve effecten op de informatiekwaliteit als gevolg.⁶⁶

De botsing tussen verschillende organisatorische vormen is er tevens de oorzaak van dat aangebrachte veranderingen zo goed als onomkeerbaar blijken⁶⁷ en dat zwakke punten, zoals de afhankelijkheid van eens gekozen technologische systeemoplossingen, slechts moeizaam zijn aan te passen.⁶⁸ Het inbrengen van nieuwe technologische componenten of diensten in de infrastructuur, bijvoorbeeld hard- en softwareoplossingen op basis van *cloud computing*, zijn uiteraard mogelijk, maar veranderen weinig aan de bestaande botsing tussen organisatorische vormen. Beide vormen beïnvloeden elkaar voortdurend en proberen over en weer nieuwe interpretaties en heroriëntaties ingang te doen vinden. Heroriëntaties zijn dus, zij het moeizaam, mogelijk.⁶⁹ Organisatorische doelstellingen vereisen die heroriëntatie; de informatie-infrastructuur zal aangepast moeten worden door ook de informatiewaardeketen te implementeren voor ongestructureerde informatie. Voor een effectieve *Archiving & Compliance* is aanpassing van de bestaande informatie-infrastructuren nodig.

Twee sociologische concepten kunnen een rol spelen bij deze heroriëntatie: inscriptie en translatie. Inscriptie verwijst naar de manier waarop gewenste werkwijzen en handelingen verankerd kunnen worden in organisatorische vormen. Translatie wijst op het realiseren van een gemeenschappelijke interpretatie van een gewenste inscriptie tussen de verschillende belanghebbenden in organisaties. Samen zorgen deze concepten ervoor dat gemeenschappelijk geaccepteerde standaarden en protocollen, gebruikersverwachtingen en -ervaringen, bureaucratische procedures, kwaliteitsnormen, en dergelijke, gebruikspatronen inscriberen. De informatiewaardeketen kan, denk ik, op deze wijze worden geïnscribeerd in informatiekwaliteit, specifiek gedrag, organisatorische functies en hard- en software, mits een gemeenschappelijk belang wordt onderkend dat breed gedragen wordt. Het veelvuldig gebruikte woord ‘draagvlak’ is een translatiebegrip. De processen van de informatiewaardeketen (creatie, vastlegging, opslag, bewerking, structurering, publicatie, gebruik, behoud, toegankelijkheid, distributie, waardering, vernietiging, beveiliging en toetsing) worden geïnscribeerd in de wijze van afhandeling van acties en transacties, in de gebruikte informatie- en communicatietechnologie, in de eisen tot toepassing, opname of deelname aan functionaliteiten, werkprocessen, transacties of syste-

men, in de systeemarchitectuur, in de services die van serviceproviders worden afgenomen en in de context en kwaliteit van de informatie zelf. Op die wijze wordt de gewenste werking van de informatiewaardeketen afgedwongen en worden het genereren van betrouwbare informatie en het maken van betrouwbare reconstructies van het verleden mogelijk gemaakt.⁷⁰

Draagvlak als translatiebegrif

Maar liefst 540.000 resultaten voor wie 'draagvlak creëren' opzoekt via Google. Maar wat is draagvlak nou precies? Waarom is draagvlak zo belangrijk?

Draagvlak heeft meerdere betekenissen. Draagvlak creëren betekent ervoor zorgen dat je ondersteuning en goedkeuring werft voor initiatieven die je wilt uitvoeren of beslissingen die je wilt nemen. Als je aan mensen vraagt wat draagvlak is, krijg je vaak te horen: een gezamenlijk doel hebben, motivatie of samen verantwoording dragen. Draagvlak in de meest ideale vorm is de back-up en support van de aanwezigen, het creëren vormt het begin van een succesvolle bijeenkomst.

Het doel van draagvlak is in veel gevallen overeenstemming hebben over verandering. Het doel is een *mindset* die leidt tot versterking en verandering van denkwijze en gedrag.

Aanpassingsvermogen is niet ieders grootste gave. Daarom roepen veranderingen vaak weerstand op. Er bestaat geen blauwdruk voor het creëren van draagvlak. Communiceren en het geven van informatie is van groot belang. Als mensen in een vroeg stadium worden betrokken bij een verandering, bijvoorbeeld doordat ze inspraak krijgen of deelnemen aan het implementatieproject, ontstaat er draagvlak voor de verandering of het project. Onderzoek bewijst dat leren en veranderen dezelfde activiteit in de hersenen is.

Wist u dat een normaal mens minstens zeven herhalingen nodig heeft om informatie op te slaan in het langetermijngeheugen? Translatie is een traag proces. Wie herhaalt, wint!

Bron: <http://www.mediaplaza.nl/?p=1831>.

De botsing tussen traditionele en nieuwe organisatorische vormen leidt tot een zeer uitdagend werkkterrein, dat aanleiding geeft tot talloze vragen waarop door onderzoek en onderwijs door deze hogeschool kan worden ingespeeld. Kern daarvan zal de informatiewaardeketen zijn. Die keten is de sleutel voor het waarborgen van betrouwbare informatie en betrouwbare reconstructies van het verleden. Daarnaast zal grote aandacht besteed moeten worden aan de informatie-infrastructuur en de wijze waarop deze is te manipuleren tot het optimaal ondersteunen van de informatiewaardeketen voor gestructureerde én ongestructureerde informatie. De Hogeschool van Amsterdam dient studenten af te leveren die deze vaardigheid in zich hebben.

4. ‘Voor vertrouwen is hoge kwaliteit van informatie noodzakelijk’

In deze tijd van ketenorganisaties, e-commerce en e-governance, intra- en extranetten, interorganisatorische *datawarehouses*, authentieke registraties en *computer mediated exchange* moeten informatiegebruikers erop kunnen vertrouwen dat informatie betrouwbaar is. Dat wil zeggen dat bekend moet zijn dat (gestructureerde én ongestructureerde) informatie juist, volledig, consistent en uit een vertrouwde bron afkomstig is. In de literatuur wordt vertrouwen (*trust*) vooral geassocieerd met ‘vertrouwde bron’, die door kennis van de oorsprong (*provenance*) kan worden achterhaald.⁷¹ Artz en Gil stellen dat ‘*provenance of information is key to support trust decisions*’: op basis van *provenance* kan bepaald worden of het daadwerkelijk een vertrouwde bron betreft.⁷² Het uitgangspunt is dat de kwaliteit van de informatie groter is als die uit een vertrouwde bron afkomstig is en dat informatiegebruikers genegen zijn daarin meer vertrouwen te hebben.⁷³ Als niet bekend is of de informatie uit een vertrouwde bron komt (wat vaak het geval is), dan moet óf de informatiewaardeketen *volledig* zijn geïmplementeerd óf de informatie *zelf* van hoge kwaliteit zijn. Dat zijn dan namelijk de enige criteria waarop vertrouwen kan worden gebaseerd.

Ik heb kwaliteit altijd een moeilijk begrip gevonden. Dat komt doordat het geen concept is maar een construct. Het begrip wordt geconstrueerd ‘in the minds of the definers, so no single definition is correct for every circumstance’.⁷⁴ Het is een subjectief begrip. Het zoeken naar een alomvattende omschrijving lijkt me dan ook zinloos, maar het is wel mogelijk om een aantal benaderingen te onderscheiden. Pirsig bijvoorbeeld omschreef kwaliteit in het onvolprezen *Zen and the Art of Motorcycle Maintenance* op een transcendente, volstrekt subjectieve en niet te operationaliseren manier: ‘Quality cannot be

defined, you know what it is'.⁷⁵ In een productgerichte benadering is kwaliteit inherent aan een product. Die kwaliteit kan door de aan- of afwezigheid van meetbare productelementen objectief worden vastgesteld. In een gebruikersgerichte benadering bepaalt de consument welke producten kwaliteit bezitten: kwaliteit 'lies in the eyes of the beholder', samen te vatten in 'fitness for use'.⁷⁶ Kwaliteit kan ook benaderd worden vanuit een producent: elke afwijking van het ontwerp of de productspecificaties betekent een verlies aan kwaliteit. Deze benadering is samen te vatten als 'conformance to specifications'.⁷⁷ En ten slotte kan kwaliteit ook nog gedefinieerd worden in termen van waarde, van kosten en prijzen. Een kwaliteitsproduct wordt geleverd tegen een acceptabel kostenniveau en een aantrekkelijke prijs ('affordable costs').⁷⁸ Kwaliteit heeft dus vele dimensies, die niet allemaal in één omschrijving te vangen zijn.

Toegepast op de kwaliteit van informatie betekent het dat het productieproces van informatie een informatieproduct levert dat voldoet aan de gestelde specificaties en aan de gebruikersverwachtingen, dat geschikt is voor gebruik en dat tegen aanvaardbare kosten produceerbaar en verkrijgbaar is. Ik gebruik zelf graag de begrippen inherente en pragmatische kwaliteit van informatie. De inherente kwaliteit wijst op de juistheid, volledigheid en correctheid van de informatie zelf. De informatie is een 'equivalent reproduction' van de werkelijkheid die in de informatie is vastgelegd.⁷⁹ De pragmatische kwaliteit wijst op het nut en de waarde die de informatie heeft voor de bedrijfsprocessen en het effectief en efficiënt bereiken van de doelstellingen van organisaties.⁸⁰ Het ontbreken van een van beide vormen van kwaliteit schept problemen. Het kwaliteitsprobleem is het grootst indien informatie pragmatische kwaliteit heeft, maar geen (of marginale) inherente kwaliteit. In dat geval is de informatie blijkbaar van belang voor het bereiken van de organisatorische doelstellingen, maar is realisatie daarvan niet mogelijk omdat de kwaliteit van de informatie zelf slecht of onzeker is. In organisaties is dit het meest voorkomende probleem.⁸¹ Dit kwaliteitsprobleem is de grootste bedreiging voor het vertrouwen van de gebruikers. Hoe je het ook wendt of keert: de kwaliteit van informatie dient dusdanig te zijn dat beleid, besluitvorming, producten, acties en transacties kunnen worden gerealiseerd en verantwoord, en dat betrouwbare reconstructies van het verleden mogelijk zijn. Ondanks het feit dat kwaliteit al jarenlang onderwerp is (geweest) van informatiekundige literatuur⁸², blijkt dit in de praktijk zeer problematisch. Informatie-infrastructuren hebben de informatie-waardeketen te lang ongecontroleerd gelaten.

Ik stel hoge eisen aan de kwaliteit van informatie, juist om betrouwbaarheid te garanderen en reconstructies mogelijk te maken. Alle informatie binnen organisaties dient over pragmatische kwaliteit te bezitten, dient aan gestelde specificaties (normen, wet- en regelgeving) en gebruikersverwachtingen te vol-

doen, dient geschikt te zijn voor gebruik en is (indien mogelijk) tegen aanvaardbare kosten te produceren. De inherente kwaliteit van informatie is onafhankelijk van de gebruikte informatietechnologie. Dit betekent dat elke gebruikte technologie *in elk geval* (als een *conditio sine qua non*) die inherente kwaliteit (van zowel gestructureerde als ongestructureerde informatie) moet kunnen realiseren en in stand houden, gestuurd door de processen van de informatiewaardeketen.⁸³ Gezien de rol die de informatiewaardeketen heeft in het borgen van de kwaliteit van informatie is het nodig om ook kwaliteitseisen te stellen aan de processen in deze keten.⁸⁴

Cruciaal voor de inherente kwaliteit van informatie is *historiciteit*. Dit betekent dat informatie kan worden gereconstrueerd zoals het ooit, op een eerder moment in de tijd, is gemaakt, ontvangen of vastgesteld. Historiciteit veronderstelt een zodanige vastlegging (conversie, bewaring en waardering) dat de blijvende juistheid en volledigheid in alle omstandigheden gewaarborgd is en dat, indien informatie in bedrijfsprocessen is gewijzigd, achteraf kan worden vastgesteld wie, wanneer, welke wijzigingen heeft aangebracht. Historiciteit wordt bepaald door:

- *integriteit*: de mate waarin informatie (en de weergave daarvan) is zoals het was, waarbij niets ten onrechte is toegevoegd, verdwenen, achtergehouden of veranderd;
- *authenticiteit*: de mate waarin, ongeacht de gebruikte compressiemethodiek, de weergave van informatie de juiste oorspronkelijke vorm en inhoud bevat;
- *controleerbaarheid*: de mate waarin informatie (en de weergave daarvan) toetsbaar is.

Samengevat: informatie moet meerdere malen, onafhankelijk van tijd, kunnen worden samengesteld, met dezelfde inhoud, presentatievorm en samenhang als op het moment van ontstaan, ontvangst of vaststelling.

De informatiewaardeketen borgt de inherente kwaliteit van informatie tijdens de hele 'levenstijd' ervan, ongeacht alle handelingen die nodig blijken te zijn bij bewerking of behoud (zoals conversie of compressie). Op de keten zijn dezelfde kwaliteitsdimensies van toepassing als op de bedrijfsprocessen, te weten levertijd, leverbetrouwbaarheid, effectiviteit, efficiëntie, productkwaliteit, behoeftenafstemming, productenbeheer en *compliance*. Dit levert het diffuse beeld op dat de processen van de informatiewaardeketen bijdragen aan het realiseren van kwaliteit binnen bedrijfsprocessen, maar dat ze zelf ook moeten voldoen aan de kwaliteitseisen die ze in de bedrijfsprocessen proberen te waarborgen. Op die kwaliteitseisen dient voortdurend te worden getoetst.

Grote investeringen in de informatie-infrastructuur garanderen niet de kwaliteit van informatie en informatiewaardeketen.⁸⁵ Daartoe zal gebruikge- maakt moeten worden van de inscriptie- en translatiemogelijkheden, zoals ik die hiervoor aan de orde gesteld heb. Gezien de enorme toename van de informatie in elke organisatie, wordt de nadruk op de kwaliteit van de informatie groter. Kwaliteit van informatie zal als een van de belangrijkste aspecten van informatiemanagement ingebed moeten worden in de opleidingen binnen de Hogeschool van Amsterdam. Studenten zullen moeten leren om te bepalen of informatie pragmatische en inherente kwaliteit heeft. Ze zullen op basis van deze bepaling moeten leren welke acties moeten worden ondernomen. Ze zullen moeten leren informatie te herkennen die pragmatische kwaliteit mist, en deze moeten durven vernietigen. Ze zullen moeten leren welke maatregelen genomen kunnen worden om de inherente kwaliteit van informatie te verbeteren en te behouden. Ze zullen moeten leren hoe de kwaliteitscriteria die aan informatie en aan de informatiewaardeketen worden gesteld kunnen worden ingebed in technologie, in software en in hardware. Ze zullen als informatie-professionals verantwoording moeten durven dragen voor de betrouwbaarheid van informatie. In het praktijkgericht onderzoek naar *fully automated archiving environments* speelt kwaliteit een belangrijke rol. Het kwaliteitssysteem dat ten grondslag ligt aan deze omgevingen zal samen met marktpartijen vormgegeven worden.

Problematische reconstructie

Binnen een verzekeringsmaatschappij is de polis- en schadeafhandelingsadministratie volgens stringente kwaliteitsprocedures ingericht. Binnenkomende correspondentie van verzekerden wordt gescand en in procesapplicatie toegekend aan het dossier van de klant. De originele documenten worden vernietigd.

De medewerkers van de polisadministratie voegen registratiegegevens toe in de database van de procesapplicatie. Aan het digitale dossier worden ook de uitgaande correspondentie en de procesformulieren toegevoegd. In deze formulieren, gemaakt met een macro in een tekstverwerkingsapplicatie, worden alle afhandelingsgegevens van de specifieke polissen en schades opgenomen. De formulieren worden aan het dossier toegevoegd in het specifieke bestandsformaat van de tekstverwerkingsapplicatie.

De automatisch toegevoegde gegevens zijn grotendeels *pointers* naar de grote, centrale klantendatabase, die alle klantgegevens bevat. In de formulieren wordt vastgelegd hoe de specifieke polisaanvragen en schademeldingen van klanten worden afgehandeld; veel van deze gegevens worden automatisch toegevoegd uit de actuele procesbeschrijvingen. De correcte afhandeling van deze procedures is door de Chief Compliance Officer verplichtend opgelegd.

Interne richtlijnen leggen de reconstructie van deze procedures gedurende een periode van vijf jaar na vervallen van de polis op. In het formulier is het niet mogelijk deze automatisch toegevoegde gegevens te veranderen. Bij een audit ontstaat op de administratie paniek als blijkt dat reconstructie van specifieke zaken niet mogelijk blijkt.

Vele formulieren zijn onbruikbaar geworden, aangezien vele automatisch toegevoegde gegevens niet meer de inhoud geven van het moment waarop ze zijn ingegeven, maar de actuele gegevens, zoals ze in de centrale database zijn ingegeven. Ook de proceduregegevens blijken de actuele situatie weer te geven, niet de situatie op het moment van de afhandeling van de specifieke procedure.

Het probleem kan op een relatief simpele wijze worden opgelost, namelijk door de ingevulde formulieren te migreren naar het PDF-formaat als het in het dossier wordt opgenomen. Herstellen van de problematische formulieren blijkt onmogelijk.

5. 'Betrouwbare informatie is een contextuele uitdaging pur sang'

Context is geen natuurlijk verschijnsel. Het is een dynamisch sociaal construct, bedoeld om betekenis te ontleen aan objecten of situaties.⁸⁶ Intuïtief weet iedereen wat ermee bedoeld wordt als het wordt gebruikt. Die intuïtiviteit komt overeen met de transcendente wijze waarop kwaliteit door Pirsig is gedefinieerd: 'Context cannot be defined, you know what it is'. Meestal is dat intuïtieve begrip voldoende om een object of een situatie te kunnen begrijpen. Dat we in staat zijn om een begrip zo intuïtief te begrijpen, geeft aan dat het een uiterst generieke betekenis heeft, die moeilijk te specificeren valt. In mijn

optiek is het begrip een van de meest fundamentele grondslagen van *Archiving & Compliance*. Het is (om informatie te kunnen begrijpen) nodig om de context ervan te kunnen bepalen, ook al is die informatie in het (verre) verleden gegenereerd.

Vijftien jaar geleden al stelde Brenda Dervin over het begrip 'context' dat 'the only possible conclusion is that there is no term that is more often used, less often defined, and when defined so variously as context'.⁸⁷ Om het wat minder netjes te zeggen: het begrip is een 'conceptual garbage can'.⁸⁸ Dat kan ook niet anders, als het begrip op detailniveau onderzocht is in vele disciplines, zoals filosofie, contextuele psychologie, socio-linguïstiek, kunstmatige intelligentie, *information retrieval*, situationele semantiek, organisatiewetenschap, informatiewetenschap, archiefwetenschap, etc.⁸⁹ Ondanks omvangrijk onderzoek naar het verschijnsel, zijn begripsomschrijvingen die verder gaan dan het benoemen van verschijnselen, situaties, objecten en omstandigheden die als context zijn te beschouwen zeldzaam.⁹⁰ Ze lijken ook op elkaar. Volgens Penco is het veel beter om over een 'family resemblance'-concept te spreken.⁹¹ Er zijn een aantal gemene delers in de verschillende benaderingen:⁹²

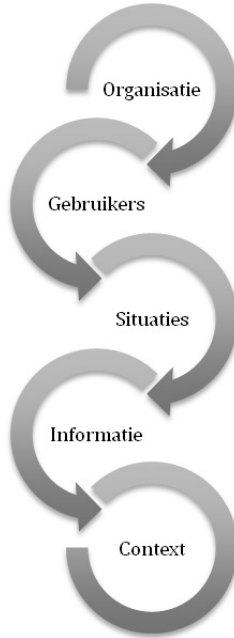
- context is een complexe, interactieve sociale realiteit, maar tegelijkertijd ook een vastgelegd, gesimplificeerd construct, dat die sociale realiteit representeert;
- context is nodig voor het ontlenen van betekenis aan het object of de situatie, waarvoor het als context fungeert. Met behulp van de context wordt een object of situatie begrijpelijk en interpreteerbaar;
- context bestaat uit de (in- en externe) eigenschappen, relaties, plaats en tijd, die een situatie of object bepalen en de regels die daaraan ten grondslag liggen;
- context is de relevante omgeving van een object of situatie en bepaalt de relevantie van de informatie, ongeacht het doel waarvoor die informatie wordt gebruikt;
- context bestaat uit de geaccepteerde interpretaties van de omstandigheden in de omgeving van situaties en objecten.

Een van de voornaamste vertolkers van het begrip context in het kader van *Archiving & Compliance*, David Bearman, stelt dat informatie (in zijn terminologie: 'records') enkel betekenis kan krijgen indien het wordt begrepen 'with reference to their origins in activity' en 'by reference to the function of which they are evidence and the record system in which they were created, stored, preserved, and accessed by the organization'.⁹³ Gedetailleerde kennis van de organisatiestructuur en de bedrijfsprocessen van organisaties maakt het mogelijk om de vormingswijze van informatie te reconstrueren en er bewijs- en in-

formatiewaarde aan te ontlenen.⁹⁴ Context geeft inzicht in de activiteiten waarin informatie is gegenereerd en gebruikt, en legt vast welke vorm en gedrag de informatie had in die setting.⁹⁵ Die vastlegging gebeurt in een speciale vorm van informatie: metadata. Deze vormen een gesimplificeerd construct, dat werkelijke gebeurtenissen op een *bepaald moment* in de tijd representeert, maar ook de omgevingsinterpretaties vastlegt die de organisatie *op dat moment* op die gebeurtenissen heeft toegepast. De metadata zijn nodig om *later in de tijd* betekenis te ontlenen aan de informatie. Ze zijn nodig om het verleden te kunnen reconstrueren. Ze vormen als het ware een model van de context.⁹⁶ Metadata en informatie zijn onlosmakelijk verbonden. Informatie in het *organizational memory* is een afbeelding van de werkelijkheid, waarin (samen met de metadata die de context gesimplificeerd weergeven) de *ontwikkeling* van organisaties *in de tijd* tot uitdrukking komt. Die afbeelding ziet er op ieder moment in de tijd anders uit vanwege de voortgaande ontwikkeling van organisaties, die voortdurend in andere omstandigheden en op basis van andere interpretaties evolueren.⁹⁷ Door deze combinatie van metadata en informatie kan de relevantie van informatie worden bepaald, de kwaliteit ervan beoordeeld, bewijs worden geleverd en verantwoording afgelegd. Context maakt het mogelijk het verleden betrouwbaar te reconstrueren en geeft inzicht in de wijze waarop informatie in bedrijfsprocessen is gegenereerd en verwerkt.⁹⁸ Het geeft zicht op de betrouwbaarheid van informatie en op de wijze waarop de informatiewaardeketen is uitgevoerd. De kwaliteitscriteria van informatie zijn dan ook *dwingend* van toepassing op metadata.

Een voorbeeld. Beleid, besluitvorming, producten, doelstellingen, acties en transacties komen tot stand, worden uitgevoerd of nagestreefd in een complexe, interactieve sociale realiteit, de organisatorische bedrijfsprocessen. De informatiewaardeketen speelt hierin een belangrijke rol. Bij de uitvoering van bedrijfsprocessen wordt informatie gegenereerd, die volgens bepaalde procedures, wet- en regelgeving, normen en maatschappelijke verwachtingen (verankerd in de processen van de informatiewaardeketen) wordt afgehandeld. Van deze informatie wordt (via het proces 'vastleggen' van de informatiewaardeketen) voortdurend in metadata vastgelegd in welke omgeving zij is gegenereerd, wat de rol ervan is, hoe zij tot stand is gekomen, welke eigenschappen zij vertoont, welke inhoud zij bezit en voor welke periode zij relevant is en bewaard moet blijven. In de metadata wordt de sociale situatie van de bedrijfsprocessen vastgelegd van dat moment: toepasselijke wet- en regelgeving, omgevingsvereisten en -verwachtingen, processchema's, *logfiles*, organisatiestructuren, autorisaties, normen, en dergelijke. De verschillende processen van de informatiewaardeketen zorgen vervolgens voor opslag, structurering, distributie, gebruik, waardering, behoud, etc. Bij raadpleging van informatie dient al-

tijd een beeld te ontstaan van de realiteit op het moment dat activiteiten plaatsvonden, in ieder geval zolang als de bewaartermijn van informatie van toepassing is. Gedurende die periode is de informatie relevant.⁹⁹



Figuur 2 Context als representatie

Dit inzicht in wat context is, geeft de uitdaging aan, waarmee *Archiving & Compliance* geconfronteerd wordt. Het ontwikkelen van metadataschema's voor het vastleggen van context is gangbare praktijk.¹⁰⁰ Het (geautomatiseerd en/of handmatig) toekennen van metadata aan informatie op basis van metadataschema's is realiteit, al is zeker geen sprake van universele en gedetailleerde toepassing.¹⁰¹ Reconstructie van context vereist een *dynamische* vastlegging (op detailniveau) van voortdurend veranderende omgevingen. Dit is verre van realiteit, aangezien de vastlegging van context in metadata voornamelijk *statisch* is. Er wordt slechts marginaal rekening gehouden met veranderingen in een dynamische omgeving. Dat is niet verbazend: metadata en metadataschema's worden gedefinieerd voor toepassing *in de toekomst* en dienen gedurende een langere periode van kracht te zijn. Automatische vastlegging van systeemmetadata (zoals *logfiles*) tracht dit te ondervangen, maar wordt niet

consistent toegepast. Het is dé uitdaging om methoden en technieken te bedenken en ontwikkelen die in staat zijn context vast te leggen op een manier die recht doet aan het dynamische karakter ervan en die in staat is tot op detailniveau voortdurend, volledig geautomatiseerd, de omgeving vast te leggen van ieder informatieobject dat in organisaties wordt gegeneerd. Ik noem dat (aansluitend bij het gangbare management-jargon) Contextual Content Management. Context is cruciaal voor het waarborgen van betrouwbare informatie, betrouwbare reconstructies van het verleden en de informatie- en bewijswaarde van informatie in de tijd. Contextual Content Management zal dan ook een belangrijk onderzoeksobject worden binnen mijn lectoraat.

6. 'Bij een informatie-explosie is het bepalen van relevantie cruciaal'

Vanaf het moment dat de informatiewetenschap tot een afzonderlijke discipline uitgroeide (de jaren 1945-1955)¹⁰² is 'relevantie' geïdentificeerd als een fundamenteel concept. Relevantie werd door informatiewetenschappers gebruikt als maatstaf bij het ontwerp en de evaluatie van informatiesystemen en bij het onderzoek naar menselijk informatiegedrag.¹⁰³ Relevantie is, net als context, een centraal concept in menselijke communicatie. Het is een term die voortdurend gebruikt wordt, maar die (ook) moeilijk te definiëren is. Zeker is dat context de relevantie van informatie bepaalt. Of, zoals Saracevic in 2007 stelde: 'For relevance, context is it.'¹⁰⁴ Iets relevant vinden is een menselijke afweging, niet die van een systeem. Menselijke afwegingen zijn complex, vaak intuïtief en moeilijk te reproduceren door een systeem. Informatie is enkel relevant voor gebruikers in de context waarin zij deze informatie raadplegen en gebruiken. Informatie is voor organisaties relevant in hun eigen context. Organisaties zullen voortdurend toetsen of 'hun' informatie pragmatische kwaliteit bezit, of informatie relevant is voor het bereiken van organisatorische doelstellingen, zeker nu de hoeveelheid informatie jaar na jaar explosief groeit.¹⁰⁵ Voor de gebruikers van deze informatie binnen de bedrijfsprocessen is informatie relevant als die antwoorden biedt op specifieke vragen over de zaken die ze aan het afhandelen zijn of die ze afgehandeld hebben. Ze verwachten 'vindtechnologie' die informatie aanlevert gebaseerd op hun persoonlijke context binnen hun organisaties.¹⁰⁶

Relevantie is een focus voor onderzoek naar zoekgedrag. Het concept heeft verschillende dimensies gekregen, die vaak als verschillende vormen van relevantie worden voorgesteld, zoals de systeem- of algoritmische relevantie, de onderwerprelevantie, de cognitieve relevantie, de gebruikersrelevantie en de

situationele relevantie.¹⁰⁷ In het kader van *Archiving & Compliance* zijn vooral de onderwerprelevantie en de situationele relevantie van belang, maar dan toegepast op organisaties, niet op personen. De onderwerprelevantie komt tot uitdrukking in de nauwkeurigheid van de zoekresultaten naar gevraagde onderwerpen of zaken in de bedrijfsprocessen van een organisatie.¹⁰⁸ De situationele relevantie heeft betrekking op de relatie tussen informatie en de taak die door procesmedewerkers moet worden uitgevoerd of de vraag die moet worden beantwoord. Deze vorm van relevantie is sterk context-afhankelijk en potentieel dynamisch van karakter.¹⁰⁹ Beide vormen van relevantie spelen een rol bij de uitvoering van bedrijfsprocessen.

Een andere vorm van organisatorische relevantie is verbonden met een belangrijk thema in organisatie- en informatiekundige literatuur: de ‘waardering’ van informatie. In die literatuur staat vooral de *business value* centraal, de waarde die informatie heeft als onderdeel van het kapitaal van organisaties.¹¹⁰ Deze benadering kijkt naar informatie in termen van *return on investment*, van strategische matching, concurrentievoordelen en risico’s. Het gaat in het kader van deze rede niet om het bepalen van die economische waarde, maar om de relevantie van deze informatie voor het handelen van organisaties en voor het ‘maatschappelijke geheugen’. Een deel van deze waarde wordt afgedekt door situationele en onderwerprelevantie. Concepten als ‘waarde’ en ‘waardering’ (*appraisal*) zijn binnen de archieftheorie ontwikkeld voordat binnen de informatiewetenschap over relevantie werd gedacht.¹¹¹ Het gaat binnen waardering (niet voor niets een van de processen binnen de informatiewaardeketen) om de bepaling van de waarde, ofwel de relevantie van informatie op korte en op lange termijn. Het betreft de complexe evaluatie van informatie ter bepaling van hun relevantie in de tijd op basis van:

- het economische perspectief, waarbij de aandacht vooral gericht wordt op waardevermeerdering voor de organisatie;
- het bedrijfskundige perspectief, waarin informatie een bedrijfsmiddel is. Informatie dient als *trigger* of product voor bedrijfsprocessen of activiteiten daarin. Informatie heeft een eigen rol en een eigen functie in de verschillende bedrijfsprocessen; zij kan een eigen vorm, een eigen structuur en een eigen relevantie hebben, die afwijkt van de relevantie die zij heeft in een ander proces;
- de financiële en fiscale perspectieven;
- het juridische perspectief, dat aan informatie vele specifieke waarden toekent;
- het cultuurhistorische of maatschappelijke perspectief, waarbij informatie ‘enduring value’¹¹² krijgt, waarbij de maatschappelijke relevantie zo groot wordt geacht dat ze blijvend dient te worden bewaard.¹¹³

Deze evaluatie van informatie wordt volgens vastgestelde criteria (waaronder wet- en regelgeving) uitgevoerd, op grond waarvan wordt bepaald hoelang specifieke informatie moet worden bewaard, waar dat moet gebeuren, wie toegang krijgt, de voorwaarden die aan die toegang verbonden zijn en de manier waarop de te vernietigen informatie wordt verwijderd. De waarderingsbeslissing is (uiteraard) niet onherroepelijk, maar herwaardering is wel aan procedures gebonden. Het proces 'waarderen' in de informatiewaardeketen voert deze evaluatie continu uit. In dit proces wordt in principe de organisatorische relevantie van informatie in de tijd bepaald.

Waarderen speelt een belangrijke rol in het verwijderen van redundante informatie om op die manier de bruikbaarheid van de te bewaren informatie te verhogen.¹¹⁴ Het verwijderen van irrelevant geworden informatie is vooral in deze tijd van *overload* belangrijk. Daar past wel de opmerking bij dat dit niet ten koste mag gaan van de inherente kwaliteit van informatie, vooral voor wat betreft de integriteit. Er mag (vanuit overwegingen van *compliance*) niets worden verwijderd dat de betekenis van informatie kan verstoren en waarbij dus de integriteit van archieven wordt aangetast.¹¹⁵ Bij het verwijderen van informatie dient er ook op te worden gelet dat het mogelijk blijft het handelen van organisaties in het verleden te kunnen reconstrueren. De uitvoering van deze waardering wordt steeds meer bepaald door het digitaler en massaler worden van de hoeveelheid informatie. Huskamp Peterson herkende al in 1984 dat de potentie voor het verloren gaan van informatie een ingebouwd aspect van informatietechnologie was.¹¹⁶ Daarom dient waarderen al te beginnen bij het ontwerp van informatiesystemen.¹¹⁷ Digitalisering leidt er ook toe dat de context van bedrijfsprocessen of organisatorische functies steeds meer de waardering van informatie gaat bepalen.¹¹⁸ Een voorbeeld van een dergelijke waarderingsmethodiek is de Nederlandse PIVOT-methode, waarbij op basis van de handelingen van de overheid waarderingsbeslissingen genomen werden.¹¹⁹ Vele alumni van de archiefopleiding van deze hogeschool hebben daar hun (vaak) eerste schreden in het beroependomein van de archivaris gezet.

Waardering leidt tot het samenstellen en implementeren van selectielijsten, de overzichten van bewaartermijnen die aangeven hoelang specifieke informatie moet worden bewaard alvorens zij mag worden vernietigd of blijvend bewaard. Waarderen triggert de processen bewaren en vernietigen. Bewaren zorgt ervoor dat informatie wordt bewaard, zolang als nodig is om reconstructies van het verleden te maken en zodanig dat de inherente kwaliteit van informatie is gewaarborgd. Vernietigen zorgt ervoor dat de vernietiging van irrelevante informatie wordt doorgevoerd. Dit dient geautoriseerd, afdoende (d.w.z. onomkeerbaar), veilig, tijdig en gedocumenteerd te gebeuren.¹²⁰ Het mag alleen plaatsvinden indien een selectielijst wordt toegepast. Voor publieke orga-

nisaties is vernietiging van informatie op basis van een selectielijst een wettelijke verplichting.¹²¹ Private organisaties lijken vaak het eindeloos bewaren van informatie verleidelijker en eenvoudiger te vinden, maar maken zich (naast de hoge kosten voor het bewaren en toegankelijk houden van irrelevante informatie) kwetsbaar voor juridische procedures via eDiscovery. Bewaarde informatie (of deze vernietigd had moeten zijn of niet) kan altijd voor bewijsgraving worden gebruikt. Een ‘information clean-up’ op basis van vastgestelde procedures en selectielijsten is (samen met een goede documentatie van het vernietigingsproces) voldoende voor ‘defensible disposal’.¹²²

Een aspect bij waarderen is het zogenaamde recht om vergeten te worden, dat vanaf 2002 onderdeel van discussie is geworden. Voorop gesteld: alle wet- en regelgeving inzake privacy wordt geëvalueerd in het waarderingsproces, waardoor persoonsgegevens onder selectielijsten vallen en worden vernietigd op het daarvoor vereiste moment. Het gaat hier vooral over de wijze waarop internetactiviteiten van webgebruikers worden vastgelegd en gemonitord door internetbedrijven, een praktijk die grensoverschrijdend is en daardoor het risico loopt (of het voordeel heeft) onder veel divergerende wet- en regelgeving te vallen. Blanchette en Johnson wezen in 2002 al op de ‘social benefits of forgetfulness’ in het kader van dataretentie.¹²³ Mayer-Schoenberger heeft in 2007 ‘forgetful computers’ geïntroduceerd.¹²⁴ Hij bepleit de ontwikkeling van systemen die regelmatig gegevens vernietigen, in zijn terminologie ‘data ecology’. ‘If whatever we do can be held against us years later, if all our impulsive comments are preserved, they can easily be combined into a composite picture of ourselves. Afraid how our words and actions may be perceived years later and taken out of context, the lack of forgetting may prompt us to speak less freely and openly’.¹²⁵ Technisch blijkt het ‘vervagen’ van dit soort gegevens niet eens zo heel erg ingewikkeld.¹²⁶ Een deel van de discussie wordt veroorzaakt door het feit dat voor veel digitale informatie de bestaande waardering niet of nauwelijks in werkelijke vernietiging heeft geresulteerd. Een ander deel vloeit voort uit bewaartermijnen die blijkbaar als te lang worden ervaren. En weer een ander deel richt zich op informatie, die een verschillend belang heeft voor gebruiker en leverancier en die ‘ongrijpbaar’ lijkt te zijn voor wet- en regelgeving. Het is op deze informatie dat het recht om vergeten te worden vooral van toepassing moet worden verklaard.¹²⁷

Het bepalen van de organisatorische relevantie van informatie wordt met het toenemen van de informatiemassa belangrijker. Het reduceren van de hoeveelheid te bewaren informatie brengt lagere kosten met zich mee voor de opslag ervan. Het verlaagt *compliance*-risico’s en eDiscovery-kosten. Het draagt bij aan maatschappelijk verantwoord ondernemen, omdat (hoe marginaal wellicht ook) bijgedragen wordt aan het terugbrengen van de belasting op ons

milieu via overmatig gebruik van informatietechnologie. *Archiving* speelt daarbij een cruciale rol. Als dat geen uitdaging voor de studenten van het domein MCI is...

7. ‘Het voortbestaan van informatie is uitermate twijfelachtig’

Digitale informatie is fragiel. Het voortbestaan van informatie is afhankelijk van technologieën die snel en voortdurend aan verandering onderhevig zijn. Deze veranderingen betekenen dat in relatief korte tijd zowel de gebruikte media als de technische formaten van de ‘oude’ informatie onbruikbaar worden. Tabel 1 toont hoe de combinatie van relatieve luchtvochtigheid en temperatuur de geschatte levensduur van media beïnvloedt.¹²⁸

Medium	25 RV 10° C	30 RV 15° C	40 RV 20° C	50 RV 25° C	50 RV 28° C
D3 magnetische tape	50 jaar	25 jaar	15 jaar	3 jaar	1 jaar
DLT magnetische tapecartridge	75 jaar	40 jaar	15 jaar	3 jaar	1 jaar
CD / DVD	75 jaar	40 jaar	20 jaar	10 jaar	2 jaar
CD-ROM	30 jaar	15 jaar	3 jaar	9 mnd.	3 mnd.

Tabel 1 Relatie medium, relatieve luchtvochtigheid en temperatuur voor geschatte levensduur

Dit voortbestaan van informatie (ook ‘digitale duurzaamheid’ of ‘*digital preservation*’ genoemd) staat al twee decennia in de belangstelling.¹²⁹ Het doel van de processen opslaan en behouden uit de informatiewaardeketen is om informatie in de tijd begrijpelijk, beschikbaar en bestuurbaar te houden, ondanks de voortdurende veranderingen, vervangingen, vergankelijkheid en ‘pensionering’ van hardware, besturingssoftware, software, bestandsformaten en informatiedragers. Aan opslag en behouden worden drie eisen gesteld:

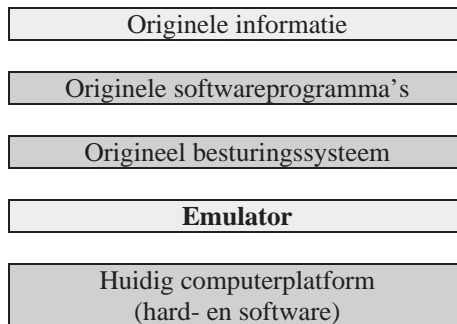
- de informatie dient reconstrueerbaar te zijn, zoals vastgesteld. Dit betekent de intacte *bit streams*, die de reconstructie van het gebruikte formulier en inhoudelijke gegevens mogelijk maken;
- de informatie dient leesbaar aan de geautoriseerde medewerkers gepresenteerd te worden;
- de informatie dient begrijpelijk te zijn in context.¹³⁰

De afhankelijkheden (en dus de risico's) zijn groot.¹³¹ De risico's worden ver-groot door compressie, back-ups en encryptie. Een *lossy* compressiemethode leidt tot verlies aan inherente kwaliteit.¹³² Back-upformaten zijn gebonden aan de oorspronkelijke IT-configuratie en niet geschikt voor het bewaren van informatie over de tijdgrenzen daarvan heen.¹³³ Encryptie vereist een volledig gewaarborgd sleutelbeheer (in de praktijk nog nergens gewaarborgd)¹³⁴, terwijl aangetoond is dat voor het behoud van authenticiteit bij langdurige bewaring encryptie niet noodzakelijk is.¹³⁵ De afhankelijkheid van de technische omge-ving maakt het voortbestaan van digitale informatie uiterst complex, en daarmee ook de mogelijkheid om het verleden te reconstrueren. 'Verversing' van media zorgt ervoor dat de *bit stream* niet verloren gaat als de dragers veroude-ren, maar het is geen strategie om het voortbestaan van informatie op lange termijn te waarborgen.¹³⁶ De migratie naar analoge *hardcopy* is dat evenmin, gezien het volledige verlies aan authenticiteit en ook de onmogelijkheid daar-voor voor de meeste vormen van digitale informatie.

Er zijn verschillende benaderingen voor het voortbestaan van informatie ontwikkeld. Ik verdeel die benaderingen in twee categorieën.¹³⁷ De eerste be-nadering richt zich op het behouden van de oorspronkelijke technologische omgeving, waarin twee strategieën ontwikkeld zijn. De eerste strategie wil wer-kende replica's van technologische platformen behouden voor toekomstig ge-bruik (de *technology preservation strategy*). De tweede strategie wil nieuwe technologische platformen programmeren om oude platformen en besturings-systemen te kunnen emuleren (de *technology emulation strategy*). De tweede hoofdbenadering richt zich op het tegengaan van de technische veroudering van bestandsformaten, waarin ook twee strategieën kunnen worden onder-kend. De eerste strategie wil het bestandsformaat van de informatie transfor-meren naar een nieuwer bestandsformaat (de *information migration strategy*). De tweede strategie wil informatieobjecten met alles wat nodig is om toegang tot dat object te verwerven inkapselen in een informatiepakket (de *information encapsulation strategy*). Ik zal de verschillende strategieën hier kort aan de orde stellen.¹³⁸

De *technology preservation strategy* impliceert het behouden van de origine-le software, besturingssysteem en hardware. Voorstanders van deze strategie benadrukken dat de originele omgeving nodig is om het gedrag en de *look and feel* van informatie te behouden. Het is echter uiterst complex, kostbaar en gevoelig voor materieel falen. Standaardisatie zou dat falen geruime tijd kunnen uitstellen, maar langer dan enkele tientallen jaren is, denk ik, ondenkbaar.¹³⁹ De strategie kan gebruikt worden als methode om 'oudere' informatie leesbaar en beschikbaar te maken en vervolgens onderdeel te maken van een behoudstrategie die op langere termijn is gericht. De *technology emulation*

strategy heeft eveneens het doel de functionaliteit van het originele systeem te behouden. De strategie is gebaseerd op de veronderstelling dat systeemgedrag onafhankelijk van de implementatie kan worden gedefinieerd en gerepliceerd met een *software-engine*.¹⁴⁰ Het is het nabootsen van een platform of programma op een ander platform of programma. Informatie kan dan worden gepresenteerd op een computer die hier eigenlijk niet voor bedoeld is. Weinig behoudstrategieën zijn zo bekritiseerd als emulatie. De meest fundamentele kritiek (naast de torenhoge kosten en een aantal andere technische aspecten)¹⁴¹ richt zich op de mogelijkheid de overweldigende taak uit te voeren om de benodigde kennis te verzamelen van een breed scala aan hard- en softwaretechnologie, besturingssystemen, databases, bestandsformaten en de functionaliteiten van elk van de te emuleren applicaties. De ervaring bij migratietrajecten heeft geleerd dat juist die informatie niet of nauwelijks wordt verzameld, terwijl dit de kosten aanzienlijk zou beperken. Het is de vraag of het bij emulatie-trajecten wél lukt. Zelfs voorstanders van de strategie twijfelen over de daadwerkelijke haalbaarheid.¹⁴² In de praktijk is de strategie beperkt tot platformen waarover veel kennis beschikbaar is (bijvoorbeeld in het Dioscuri-project).¹⁴³ Het slagen van dergelijke emulatie-trajecten is hoopgevend, maar bewijst niet dat de strategie op lange termijn valide is. Ik vraag mij ook af of het intellectueel eigendomsrecht emulatie toestaat: soft- en hardware zijn eigendom van softwareontwikkelaars en het is de vraag of het re-creëren daarvan in de toekomst toegestaan is. Emulatoren gebruikt in het kader van videogames zijn dat in ieder geval niet.¹⁴⁴



Tabel 2 Emulatie

De *information migration strategy* behelst de periodieke transformatie van informatie van de ene hard- en softwareconfiguratie naar een andere configuratie, of van de ene generatie van computertechnologie naar een opvolgende ge-

neratie. Het doel van migratie is het bewaren van de inherente kwaliteit van informatie, zodat reconstructie, representatie en gebruik mogelijk blijven, ondanks de veranderende technologie. Migratie richt zich niet op technologie, maar op informatie.¹⁴⁵ Het probleem van migraties is dat er telkens dataverlies optreedt, wat gevolgen kan hebben voor de uiteindelijke representatie. Het kan uitgevoerd worden door softwareapplicaties, die *backwards compatible* zijn of door de conversie van de bestandsformaten naar een kleiner aantal gestandaardiseerde bestandsformaten, onafhankelijk van hard- en software.¹⁴⁶ Als dat al bij de creatie van informatie gebeurt, zijn de risico's van dataverlies het kleinst. Standaardisatie is het meest veelbelovend, aangezien *backwards compatibility* voor complexe formaten minder goed werkt. Standaardformaten zijn minder aan veranderingen onderhevig, behoeven dus minder vaak te worden gemigreerd en kunnen de complexiteit van vorm en structuur van het originele bestandsformaat coderen.¹⁴⁷ Een overzicht van acceptabele bestandsformaten is opgenomen in tabel 3. Tot op dit moment is migratie de enige strategie die werkt, al kan het problematisch zijn voor complexe multimedia. Om de kosten van migratie te beperken is er een model voor automatische migratie ontwikkeld, gebaseerd op bestaande bestandsformatenregisters.¹⁴⁸ De strategie is (en wordt) op grote schaal toegepast en succesvol uitgevoerd. In mijn optiek is migratie op dit moment de meest succesvolle en meest geëigende behoudstrategie. Het succes op de heel lange termijn is echter onzeker; vandaar dat de oorspronkelijke en gemigreerde *bit streams* moeten worden bewaard.

Documenttype	Bestandsformaat
Tekst	ASCII - Unicode - TIFF - PDF/A - ODF - OOXML
Images	TIFF - PNG - CGM
Vector	SVG
Geluid	WAV
CAD	DXF
GIS	GML
Video	MXF

Tabel 3 Geschikte standaard bestandsformaten voor het behouden van informatie

De *information encapsulation strategy* probeert de problemen van de technologische veroudering en het verdwijnen van bestandsformaten te ondervangen door de gegevens over de interpretatie van het informatieobject onlosmakelijk te verbinden met het object zelf.¹⁴⁹ Inkapseling betekent 'wrapping the record to be preserved within a human readable wrapper. The wrapper contains information that supports organizational preservation, and documents the pre-

served information to allow it be decoded in the future'.¹⁵⁰ Inkapseling is in essentie niet méér dan het toekennen van metadata aan informatie, op een zodanige manier dat ze onlosmakelijk aan elkaar zijn verbonden. Het is géén methode om informatie in de toekomst te reconstrueren of toegankelijk te houden: het bevat de informatie om dat in de toekomst mogelijk te maken. Als die tenminste ook hiervoor wél verzameld wordt! Inkapseling is geen behoudstrategie, maar het kan in combinatie met zo'n strategie worden toegepast. Het is niet ten onrechte als een migratiestrategie gekenschetst.¹⁵¹

Twijfelachtig voortbestaan

'It is an oft-quoted example, but a powerful one nonetheless, that the original Doomsday Book, hand-crafted on vellum in 1086, has lasted a millennium, whereas the 1986 BBC digital replica became obsolete and inaccessible within 20 years. Perhaps a little less well-known is the case of the data received from the NASA Martian lander data generated in 1975. The magnetic tape storing the datasets became brittle, and the data that was actually retrieved was found to be in a format that was impossible to decode. Scientists were reduced to rekeying the information from old printouts.'

Bron: S. Bains, J. Toon, *'From Domesday to Digital Doomsday? Digital preservation at the National Library of Scotland* [online], geraadpleegd op 31 juli 2012), http://widwisawn.cdli.strath.ac.uk/issues/vol5/5_2_1.html.

'NASA stored data from the Viking programs landing on the planet Mars in 1976 on magnetic tapes. After 20 years the tapes were unreadable because everyone who knew the used file format was either retired from NASA or deceased. After two years of "heroic efforts" one third of the information on the tapes could be reconstructed, only because some documentation had been saved on paper.'

Bron: <http://www.ltu.se/centres/Centrum-for-langsiktigt-digitalt-bevarande-LDB/Bibliotek/Introduktion/Exempel-ur-verkligheten-1.43299?l=en>.

'In the event of a disaster that destroys the vast majority of the world, humanity's legacy will largely reside on data stored on hard drives. However, hard drives were never meant for long-term storage and no one can be sure how long they will last. The Canadian Conservation Institute's (CCI's) Joe Iraci says that although the most important data is backed up

on magnetic tapes or optical discs, these formats cannot be trusted to last even five years. Iraci has conducted accelerated aging tests by exposing different forms of media to high heat and humidity. The tests found that the most reliable data storage devices are recordable CDs with a reflective layer of gold and a phthalocyanine dye layer. Many experts believe that after a major catastrophe only information that is written on paper will survive. ... Proposals to make a paper format that can store digital data for centuries using a system similar to bar codes have been slowed due to a lack of commercial interest’.

Bron: T. Simonite, M. Le Page, ‘Digital doomsday: the end of knowledge’, *New Scientist* 2745, 2 februari 2010.

De behoudstrategieën kunnen geen van alle het voortbestaan van informatie waarborgen, zeker niet indien de inherente kwaliteit van informatie bewaard moet blijven. Dat wil niet zeggen dat informatie niet gedurende een of twee decennia in stand gehouden kan worden. Dat vereist echter systemen als eDepots of *archival vaults*,¹⁵² waarvan het doel de langdurige bewaring van informatie is. Het uitgangspunt is dat eDepots informatie bewaren tot in lengte van dagen, maar ik vraag me af of dit, zonder een werkelijk werkende behoudstrategie voorhanden te hebben, niet te hoog gegrepen is. Een eDepot is ‘het geheel van organisatie, beleid, processen en procedures, financieel beheer, personeel, databeheer, databeveiliging en aanwezige hard- en software, dat het duurzaam beheren van de te bewaren digitale archiefbescheiden mogelijk maakt’.¹⁵³ Deze omschrijving brengt mij tot de conclusie dat het gaat om de inbedding van de informatiewaardeketen in de informatie-infrastructuur van organisaties. Ik ga er hierbij van uit dat kwaliteit, context, relevantie en voortbestaan geregeld zijn, al zegt de omschrijving daar niets over. Ik vind het realiseren van eDepots erg belangrijk zolang ze gezien worden als een *middel* om het *doel* van het voortbestaan van informatie en het gebruik daarvan te bereiken. De discussie over eDepots geeft mij echter soms een ongemakkelijk gevoel. Het is net alsof het meer gaat over het eDepot als doel en niet zozeer over het voortbestaan van betrouwbare informatie zelf. Het bouwen, inrichten en organisatorisch inbedden van een eDepot waarborgt dat voortbestaan niet *vanzelf*. Het vinden van een daadwerkelijk betrouwbare behoudstrategie is in dat licht cruciaal.

Dat is voor mij ook de uitdaging in het onderzoeksprogramma van het bijzonder lectoraat aangaande *fully automated archiving environments*. Die omgevingen vragen om geautomatiseerde behoudstrategieën. Dat is nu (met uit-

zondering van de automatische migratiemogelijkheden) waarschijnlijk nog een stap te ver. Maar het is helder dat het voortbestaan van informatie (ongeacht de systeemomgeving waarin dat gebeurt) een belangrijke plek in dat onderzoek inneemt. En het liefst in combinatie met onderzoek naar manieren om dit zo milieu- en kostenvriendelijk mogelijk in te richten. Hier liggen mogelijkheden voor studenten van deze hogeschool: hun creativiteit zal hard nodig zijn om aantrekkelijke *features* van een dergelijke omgeving te ontwikkelen, waarbij vanaf het ontwerp rekening gehouden wordt met het voortbestaan van informatie in de toekomst.

8. 'Het spreekt niet vanzelf dat een digital archivist archivaris is'

In 2004 schreef Jean-Daniel Zeller een notitie over de *cyberarchivist*. In deze notitie constateert hij dat een *cyberarchivist* een archivaris is 'who develops and manages material and organizational systems capable of assuring the long-term conservation of digital documents and their communication to a large public. To this end, he possesses adequate information technology and managerial competencies'.¹⁵⁴ Een mooie, zij het wat traditionele omschrijving, die in het verloop van zijn notitie wordt onderbouwd en, uiteindelijk, gerelativeerd. Volgens Zeller kan een *cyberarchivist* niet één persoon zijn; er is zo'n groot scala aan competenties vereist, dat alleen een team van archivarissen en IT-specialisten als *cyberarchivist* zou kunnen functioneren.¹⁵⁵ In de Gartner Business Intelligence Summit 2010 werden vier beroepsprofielen gedefinieerd, die vanaf 2012 een steeds groter belang binnen organisaties gaan krijgen. Een van deze profielen is de *digital archivist*, die 'will be required to appraise, arrange and preserve digital records for legal and regulatory purposes'. Gartner verwachtte dat 15% van de bestaande bedrijven in 2012 een *digital archivist* aangesteld heeft.¹⁵⁶ Luciana Duranti definieerde negen kernactiviteiten voor een *modern records professional*. Zo'n professional:

1. positioneert zich aan het begin van de levenscyclus van informatie;
2. waarborgt de authenticiteit van informatie, vooral bij systeemupgrades en veranderingen in verantwoordelijkheden;
3. identificeert te bewaren informatie bij creatie en monitort de informatie in de tijd;
4. bepaalt (in overleg met de verantwoordelijke archiefvormer) de beste manier voor het voortbestaan van de informatie;
5. bepaalt de behoudstrategie onafhankelijk van de technologische ontwikkelingen en met nadruk op interoperabiliteit;

6. controleert de accuraatheid van de informatie na conversie of migratie;
7. ontwikkelt procedures die aangelegenheden van intellectueel eigendomsrecht en privacy regelen;
8. herkent de contextuele beschrijving van informatie als een primaire authenticatiefunctie; en
9. is voortdurend betrokken bij onderzoeksprojecten.¹⁵⁷

Alhoewel overeenstemming over wat een *digital archivist* is ver weg is, is wel duidelijk dat het beroep van steeds groter belang wordt. Het feit dat het tot de exclusieve Gartner-summits en rapporten is doorgedrongen zegt iets over de beroepspotentie voor de toekomst. Gartner positioneert nadrukkelijk de combinatie van *Archiving & Compliance*. Bij het overzicht van Duranti (dat een hoog ‘ja-gehalte’ heeft) heb ik mijn twijfels. Het lijkt mij nogal uit de losse pols geschreven. Want de activiteiten 1 en 3 en 3, 4, 5 en 6 zijn overlappend, 2 en 6 zijn identiek, evenals 4 en 5. Daarnaast zijn 1 en 8 geen activiteiten maar uitgangspunten. Met uitzondering van de notitie van Zeller, blijft onduidelijk waarop de omschrijvingen zijn gebaseerd, al zal dat vooral de bestaande taakuitoefening zijn, aangepast aan de nieuwe uitdagingen. Beroepsprofielen voor een *digital archivist* bestaan immers nog niet.¹⁵⁸

Van zowel *archiving* als *compliance* is de kern de informatiewaardeketen, de keten van processen die de creatie, de vastlegging, de opslag, de bewerking, de structurering, de publicatie, het gebruik, het behoud, de toegankelijkheid, de distributie, de waardering, de vernietiging, de beveiliging en de toetsing van betrouwbare informatie realiseert. Dit laatste wordt bepaald door de vier dimensies van informatie: kwaliteit, context, relevantie en voortbestaan. Een *digital archivist* (of een *compliance officer*) richt zich primair op het realiseren van deze vier dimensies van informatie en operationaliseert een informatiewaardeketen van hoge kwaliteit om ze te realiseren. Dit impliceert voor de verschillende processen van de informatiewaardeketen dat de *digital archivist* bij:

- *Creatie*: de kenmerken van inherente kwaliteit inbouwt in de creatie van informatieproducten, zodat vorm, structuur en inhoud van (gestructureerde en ongestructureerde) informatie zijn gewaarborgd in de tijd. Dit vereist kennis en vaardigheden betreffende (de binnen een organisatie gebruikte) applicatiesoftware (kantoorautomatisering, content management, enterprise information management, mail- en messaging systemen, e.d.), (open) standaarden, de analyse van documenten en het op basis daarvan genereren van sjablonen of formats, het inrichten en definiëren van databasemanagementsystemen, XML-talen en -databases, sociale media, wet- en regelgeving, e.d.

- *Vastlegging*: de vastlegging en beschrijving van context mogelijk maakt door de ontwikkeling en toepassing van (dynamische) metadataschema's, die onlosmakelijk worden verbonden aan informatieobjecten. Dit vereist kennis en vaardigheden betreffende contextuele en situationele analyse, het ontwerpen, automatiseren en implementeren van metadatasystemen, -methoden en -procedures, logfile-ontwerp, 'lerende systemen', XML-talen en -databases, e.d.
- *Opslag*: de (duurzame) opslag mogelijk maakt door de implementatie van opslagsystemen, gebaseerd op (open) standaarden, gericht op het behoud van de inherente kwaliteit van informatie. Dit vereist kennis en vaardigheden met de inrichting, standaardisering en implementatie van opslagsystemen (Storage Area Networks, Network Attached Storage, e.d.), bestandssystemen, bestandsformaten, metadatasystemen en -methoden, compressie, authenticatie, cloud-oplossingen, wet- en regelgeving, privacybepalingen, authentieke registraties, back-upsystemen, mirroring, e.d.
- *Bewerking*: de be- en verwerking van informatie in de afhandeling van bedrijfsprocessen faciliteert, waarbij de inherente kwaliteit van informatie wordt gewaarborgd. Hier zijn kennis en vaardigheden nodig betreffende analyse en verandering van (keten)processen en informatiestromen, wet- en regelgeving, inrichting van zaaktypen, zaken, acties en transacties, de benodigde applicaties voor bewerking van informatie, data- en informatie-analyse, datawarehouses, data- en documentmining, databasemanagementsystemen, workflow management, business intelligence, e.d.
- *Structurering*: de toepassing van (ongeacht welke) ordeningssystemen in digitale archieven realiseert, voornamelijk op basis van metadatasystemen en -methoden. Hiervoor is abstractievermogen nodig om op metaniveau structuren te ontdekken in grote hoeveelheden informatie, die aansluiten bij de organisatorisch gekozen wijzen van werken en bij het gedefinieerde informatiemanagement. Kennis en vaardigheden zijn nodig betreffende (automatische) classificatie, meta-analyse, databasemanagementsystemen, metadatasystemen en -methoden, e.d.
- *Publicatie*: de publicatie van informatie faciliteert via de daartoe geëigende en organisatorisch bepaalde kanalen. Het vereist kennis en vaardigheden betreffende de beschikbare publicatiekanalen, intellectueel eigendomsrecht, digital rights management, rechten en plichten inzake de gepubliceerde informatie, wet- en regelgeving, web content management, sociale media, website-ontwikkeling, e.d.
- *Gebruik*: het gebruik faciliteert van de informatie in de tijd. Dit impliceert kennis van gebruikers en gebruikersgroepen, hun wensen en belangen, de wijzen waarop zij informatie aangeleverd wensen te krijgen, de systemen

- die zij gebruiken voor de raadpleging van informatie, internettechnologie, het ontwikkelen van queries, information retrieval technologie, eDiscovery, sociale media, e.d.
- *Behoud*: het voortbestaan van betrouwbare informatie realiseert, door de ontwikkeling, strategische inpassing en toepassing van geëigende behoudstrategieën, waarbij de economische haalbaarheid voor de organisatie voortdurend in de afwegingen wordt betrokken. Dit vereist inzicht in de informatietechnologische mogelijkheden, (open) standaarden, kosten- en batenanalyse, best practices, behoudmogelijkheden, automatische conversie en migratie, inkapseling, e.d.
 - *Toegankelijkheid*: de toegankelijkheid van informatie faciliteert gedurende het gehele bestaan van de informatie. Het vereist kennis van toegankelijkheidssystemen en -mogelijkheden, afstemming van informatiepresentatie, metadata en user generated content, information retrieval technologie, e.d.
 - *Distributie*: de logistiek van informatie analyseert en op basis van autorisaties, wensen en behoeften betrouwbare informatie ter beschikking stelt binnen (of buiten) organisaties. Hier is kennis nodig van informatielogistiek, procesanalyse, wet- en regelgeving, privacybepalingen, communicatietechnologie, beveiliging, internettechnologie, enterprise resource planning, e.d.
 - *Waardering*: de waarde van informatie bepaalt op basis van de afweging van de verschillende organisatorische belangen, wet- en regelgeving en relevantie. Kennis en vaardigheden zijn vereist betreffende het analyseren van organisatorische belangen, het onderkennen van pragmatische kwaliteit, wet- en regelgeving, het samenstellen van selectielijsten en het implementeren van deze lijsten in de bestaande metadatasystemen en -methoden, het identificeren en selecteren van informatie ter vernietiging, acquisitie door erfgoedinstellingen en behoud, e.d.
 - *Vernietiging*: de informatie (op basis van procedures) vernietigt op basis van de vastgestelde selectielijsten. Hier is kennis nodig van databasemanagementsystemen en wijzen van digitale vernietiging.
 - *Beveiliging*: de beveiliging waarborgt, zodat de inherente kwaliteit van informatie niet kan worden aangetast. Dit vereist kennis van informatiebeveiligingssystemen, authenticatie, encryptie (en technologie voor digitale handtekeningen), sleutelbeheer, autorisatiesystemen, detectiesystemen, privacy, wet- en regelgeving, e.d.
 - *Toetsing*: de toetsing waarborgt van de inherente kwaliteit van informatie, de vier dimensies van informatie, de mogelijkheid om het verleden te reconstrueren en de kwaliteit van bedrijfsprocessen en de processen van de informatiewaardeketen. Dit vereist kennis en vaardigheden betreffende

IT- en operational auditing, kwaliteitssystemen, informatiekwaliteit, proces- en informatieanalyse, systeemanalyse, wet- en regelgeving, e.d.

Archiving & Compliance vereisen nogal wat van de *digital archivist*, waarbij ik ervan overtuigd ben dat ik bij lange na niet compleet ben. Ik ben geneigd het met Zeller eens te zijn als hij stelt dat een *digital archivist* niet gauw één persoon kan zijn, gezien de benodigde kennis en vaardigheden om dat beroep goed uit te kunnen voeren.¹⁵⁹ Duranti voorspelt om dezelfde reden de ondergang van de generalist, en kiest als alternatief voor een basisopleiding met een daaropvolgende specialisatie.¹⁶⁰ Dat is een aantrekkelijk model, dat ook binnen de Hogeschool van Amsterdam aandacht verdient. Er zijn niet veel generalisten, die al de processen van de informatiewaardeketen kunnen overzien. Om de doelstellingen van organisaties te bereiken zal dat echter wel nodig zijn.

Ketelaar definieerde een aantal uitgangspunten die iedere beroepsbeoefenaar als archivistische basiskennis zou moeten omhelzen: de onschendbaarheid van de bewijswaarde van archieven, het *records continuum*, het organische karakter van de archiefvorming, de hiërarchie van archiefbescheiden en hun beschrijvingen en het *respect des fonds*.¹⁶¹ Het zijn valide uitgangspunten, denk ik, al heb ik mijn twijfels over die hiërarchie. In een digitale omgeving is een hiërarchie 'platgeslagen' in een database en blijkt een dergelijke relatie enkel uit de beschrijvingen in de metadata. Wellicht dat het beter is om dit uitgangspunt te vervangen door de vier dimensies van informatie? Er is weinig in deze uitgangspunten dat niet door andere informatiespecialisten zou kunnen worden onderschreven. Het hoeft dus niet vanzelfsprekend te zijn dat een archivaris *digital archivist* is, wel dat iemand in deze functie de bovenstaande uitgangspunten kent, herkent, begrijpt en kan implementeren. Als dat zo is kunnen ook informatici, bedrijfskundigen, IT-auditors of accountants als *digital archivist* functioneren. Het zou een verrijking zijn. Deze uitgangspunten integreren in deze (en andere) opleidingen is dan ook een van de grootste uitdagingen van het komende decennium.

9. 'Archiving should be just like an Apple^(tm)'

Ik kom tot een afsluiting. Ik zal er weinig misverstand over hebben laten bestaan dat zowel *archiving* als *compliance* cruciaal zijn voor organisaties. Allebei hebben ze een grote impact op de wijze waarop wordt gewerkt in bedrijfsprocessen. Allebei vereisen ze betrouwbare informatie die snel toegankelijk en beschikbaar is. Allebei vereisen ze een reconstructie van het verleden. Allebei hebben ze impact op de termijn waarop informatie beschikbaar is. Dat kan

‘eeuwig’ zijn, als de informatie van maatschappelijke of cultuurhistorische betekenis is. ‘Eeuwig’ de beschikking hebben over betrouwbare informatie! Informatie die voldoet aan de gedefinieerde criteria van kwaliteit, waarvan bekend is wat de relevantie was (en is) op het moment van ontstaan, waarvan context van ontstaan, gebruik en beheer bekend is (en blijft), en waarvan het voortbestaan is gewaarborgd. Dat betekent nogal wat. De informatiewaardeketen tracht dit te realiseren, maar is in de meeste organisaties slechts marginaal geïmplementeerd. Het is een hybride construct, waarin mensen en computers processen uitvoeren zonder veel coördinatie.

Mijn visie is dat een volledige, geautomatiseerde waardeketen *Archiving & Compliance* realiseert, zodanig dat gebruikers zelf geen of nauwelijks handelingen hoeven te verrichten en dat het intuïtieve processen worden. Mijn streven naar *fully automated archiving environments* komt uit dat denkbeeld voort. Medewerkers in organisaties dienen te kunnen werken met betrouwbare informatie, zonder zich af te hoeven vragen hoe die wordt opgeslagen en gearchiiveerd. Ze moeten erop kunnen vertrouwen dat de informatie er is op het moment dat ze die nodig hebben (ongeacht wanneer dat is), voorzien van alle metadata om de context, relevantie en kwaliteit te kunnen beoordelen. Datzelfde geldt op termijn voor alle informatie die als cultuurhistorisch erfgoed bewaard is. Gebruikers daarvan moeten op diezelfde intuïtieve manier de benodigde informatie kunnen vinden en gebruiken, voorzien van alle context die nodig is om de gevraagde informatie te kunnen interpreteren. Betrouwbare informatie moet als het ware met een druk op de knop beschikbaar zijn! Ik wil op zoek naar ‘lerende systemen’ die dat voor ons mogelijk maken. Er moet nog heel veel werk verzet worden om zo’n omgeving te realiseren. Maar *Archiving & Compliance* moeten op eenzelfde gebruiksvriendelijke en intuïtieve manier gaan werken als een Apple^(tm). Dan wordt het leuk! Een uitdaging dus, waar studenten hun tanden in kunnen zetten. Een uitdaging ook, waarvan de uitkomsten baanbrekend kunnen zijn voor de manier waarop mensen met informatie werken. Een uitdaging die het ‘nieuwe werken’ pas echt mogelijk maakt!

Ik sluit af met een woord van dank aan het College van Bestuur van deze hogeschool voor het instellen van dit bijzonder lectoraat, aan de Koninklijke Vereniging van Archivarissen in Nederland voor het mogelijk maken ervan, aan mijn collega’s binnen het Kenniscentrum Create-IT (waarvan ik vooral Hans Henseler wil noemen voor zijn niet-aflatende steun bij de totstandkoming van dit lectoraat), aan de leden van mijn kenniskring (Jos Vrolijk, Maarten van der Burg, Kees Stelling, Paul Rijnierse, Simone Keizer en Jim Voorzee), aan de leden van mijn geassocieerde kenniskring, aan de domeinvoorzitter Geleyn Meijer, de opleidingsmanagers Irene Sparreboom en Kees Rijsenbrij en

aan de teammanagers, docenten en studenten van de opleiding MIC voor hun stimulerende enthousiasme.

Ik heb gezegd.

Noten

1. Zie hiervoor bijvoorbeeld: R.C.H. van Otterlo, *HBO-rechten, het onderzoek(en) waard!* (Amsterdam: Hogeschool van Amsterdam, 2012) en L.W.C. Tavecchio, *Systematisch én flexibel. Methodiekontwikkeling in het spanningsveld tussen evaluatie en beleid* (Amsterdam: Hogeschool van Amsterdam, 2008).
2. F. Boudrez, H. Dekeyser, J. Dumortier, *Digital Archiving. The new challenge. Legal and archival issues* (Mont Saint Guibert: IRIS, 2005).
3. Ik gebruik de term 'informatie' in deze rede vrijelijk voor wat eigenlijk gegevens, documenten en archiefdocumenten zijn. Voor een uitgebreide bespreking van deze begrippen: G.J. van Bussel, F.F.M. Ector, *Op zoek naar de herinnering... Verantwoordingssystemen, content-intensieve organisaties en performance* (Helmond: Van Bussel Document Services, 2009), blz. 67-86.
4. R.A. Kahn, B.T. Blair, *Information Nation. Seven keys to information management compliance* (Indianapolis: Wiley Publishing, Inc., 2009²).
5. Bussel, Ector, *Op zoek naar de herinnering*, blz. 24-27.
6. In Bussel, Ector, *Op zoek naar de herinnering*, blz. 167-172, wordt deze keten als 'content value chain' aangeduid. Ook: N.S. Koutsoukis, G. Mitra, *Decision modeling and information systems. The information value chain* (Berlijn: Springer, 2003) en S.L. Cisco, K.V. Strong, 'The value added information chain', *Information Management Journal*, 33 (1999), nr. 1, blz. 4-15.
7. Bussel, Ector, *Op zoek naar de herinnering*, blz. 172-175.
8. A. Toffler, *Future Shock* (New York: Random House, 1970).
9. M. Stefik, *The Internet Edge. Social, technical, and legal challenges for a networked world* (Cambridge (Ms.): The MIT Press 2000); K. Cukier, 'Data, data everywhere', *The Economist. A special report on managing information* (27 februari 2010), blz. 1-3.
10. J. Armitage, J. Roberts (red.), *Living with Cyberspace. Technology and society in the 21st century* (Londen, New York: Continuum, 2002); J. Manyika, e.a., *Big Data. The next frontier for innovation, competition and productivity* (McKinsey Global Institute, 2011); G. Shroff, *Enterprise Cloud Computing. Technology, architecture, applications* (Cambridge: Cambridge University Press, 2010).
11. P. Lyman, H.R. Varian, *How much information?* (Berkeley: School of Information Management and Systems, University of California, 2003).
12. IDC heeft een serie *white papers* gepubliceerd: *The expanding digital universe* (maart 2007); *The diverse and exploding digital universe* (maart 2008); *As the economy contracts, the digital universe expands* (mei 2009); *The digital universe decade. Are you ready?* (mei 2010) en *Extracting value from chaos* (juni 2011). Alle *white papers* zijn beschikbaar via: www.emc.com/leadership/programs/digital-universe.htm (geraadpleegd op 31 juli 2012).
13. M. Hilbert, P. López, 'The world's technological capacity to store, communicate, and compute information', *Science* Vol. 332, no. 6025 (april 2011), blz. 60-65, en tabel 1, blz. 63.

14. S. Low, e.a., *Quarterly Statistics: disk array storage, all regions, all countries, 1Q12 update* (Stamford (CT): Gartner, 2012).
15. K.R. Gray, L.A. Frieder, G.W. Clark, *Corporate scandals. The many faces of greed: the great heist, financial bubbles, and the absence of virtue* (St. Paul (Min.): Paragon House, 2005); E. Engelen, 'Bazel III is een lachertje en bankiers lachen het hardst', *Me Judice* 3 (2010), 14 december; N. Roubini, S. Mihm, *Crisis Economics. A crash course in the future of finance* (New York: The Penguin Press, 2010); B. Holzner, L. Holzner, *Transparency in global change. The vanguard of the Open Society* (Pittsburgh (Pa.): University of Pittsburgh Press, 2006).
16. *The true costs of compliance. Benchmark study of multinational organizations* (Traverse City (Mi.): Ponemon Institute, 2011).
17. A. Braganza, A. Franken, 'SOX, Compliance and Power Relationships', *Communications of the ACM* 50 (2007), nr. 9, blz. 97-102.
18. A. Willis, 'Corporate governance and management of information and records', *Records Management Journal* 15 (2005), nr. 2, blz. 86-97.
19. S.N. Abdullah, M. Indulska, S. Sadiq, 'A study of compliance management in information systems research', S. Newell, e.a. (red.), *Information Systems in a globalising world. Challenges, ethics and practices. The 17th European Conference on Information Systems, Verona, Italy, 8-10 June 2009* (Verona: ECIS, 2009), blz. 1-10. Aangevuld met: A. Cleven, R. Winter, 'Regulatory compliance in Information Systems research. Literature analysis and research agenda', T. Halpin, e.a. (red.), *Enterprise, business-process and information systems modeling (LNBIP 29). Proceedings of the 10th International Workshop, BPMDS 2009, and 14th International Conference, EMMSAD 2009, held at CAiSE 2009, Amsterdam, The Netherlands, June 8-9, 2009* (Berlijn, Heidelberg: Springer, 2009), blz. 174-186.
20. Bussel, Ector, *Op zoek naar de herinnering*, blz. 14-27; G.J. van Bussel, 'Reconstructing the past for organizational accountability', *The Electronic Journal of Information Systems Evaluation* 15 (2012), nr. 1, blz. 127-137.
21. G.J. van Bussel, "'Green Computing": IT en energieverbruik' [online], geraadpleegd op 31 juli 2012, <http://www.vbds.nl/2010/03/09/green-computing-it-en-energieverbruik/>.
22. S. Murugesan, 'Harnessing Green IT. Principles and practices', *IT Professional* 10 (2008), nr. 1, blz. 24-33.
23. J.G. Koomey, 'Worldwide electricity used in data centers', *Environmental Research Letters* 3 (2008), no. 034008. 23 september 2008 [online], geraadpleegd op 31 juli 2012, <http://stacks.iop.org/1748-9326/3/034008>.
24. J.G. Koomey, *Growth in data center electricity use 2005 to 2010* (Oakland (Ca): Analytics Press, 2011); G. Cook, *How clean is your cloud? Catalysing an energy revolution* (Amsterdam: Greenpeace International, 2012), blz. 11.
25. Hilbert, López, 'The world's technological capacity to store, communicate, and compute information', tabel 1, blz. 63.
26. Murugesan, 'Harnessing Green IT', blz. 25.
27. 'Groen computergebruik: energie besparen en helpen de aarde te beschermen' [online], geraadpleegd op 31 juli 2012, http://h41131.www4.hp.com/be/nl/stories/fs_emea_02_08.html.

28. Zie o.a. X. Zhang, Z. Xiao-nan, L. Yi, Z. Lei-jie, 'Key technologies for green data center', Q. Li, F. Yi, Y. Liu (red.), *Third International Symposium on Information Processing, 12-14 November 2010, Qingdao (China)* (Los Alamitos (Ca.): IEEE Computer Society, 2010), blz. 477-480; Bussel, 'Green Computing'.
29. C.W. Bailey, Jr., *Digital Curation Bibliography. Preservation and stewardship of scholarly works* (Houston (Tx.): Digital Scholarship, 2012).
30. *Innovation Intelligence. Verkenning creatieve industrie* ('s-Gravenhage: Min. OC&W, 2009).
31. M. Castells, *The Information Age. Economy, society and culture* Vol. I, II en III (Oxford: Blackwell, 1996, 1997, 1998); L.Z. Karvalics, 'Information Society – what is it exactly? (The meaning, history and conceptual framework of an expression)', R. Pintér (ed.), *Information Society. From theory to political practice* (Budapest: Network for Teaching Information Society, 2007).
32. K. Kumar, *From Post-Industrial to Postmodern Society* (Oxford: Blackwell, 1995), blz. 154; F. Webster, 'The Information Society revisited', L.A. Lievrouw, S. Livingstone, *Handbook of new media* (Londen: Sage, 2002), blz. 255-266.
33. N. Negroponte, *Being Digital* (New York: Vintage, 1996); Castells, *The Information Age; Armitage, Roberts, Living with Cyberspace*.
34. O. Volkoff, D.M. Strong, M.B. Elmes, 'Technological embeddedness and organizational change', *Organization Science* 18 (2007), nr. 5, blz. 832-848.
35. T.H. Davenport, J.G. Harris, *Competing on analytics. The new science of winning* (Boston: Harvard Business School Press 2007); R. Moss Kanter, *Evolve! Succeeding in the digital culture of tomorrow* (Boston: Harvard Business School Press, 2001).
36. K.E. Weick, K.M. Sutcliffe, *Managing the unexpected. Assuring high performance in an age of complexity* (San Francisco (Ca.): John Wiley & Sons, 2007).
37. P. Ziegler, K.R. Dittrich, 'Three decades of data integration. All problems solved?', *Proceedings of the 18th IFIP World Computer Congress (WCC 2004)* (Toulouse: IFIP, 2004), Vol. 12, blz. 3-12.
38. G.J. van Bussel, 'Enterprise 2.0., Accountability and the necessity for digital archiving', W.F. Riekert, I. Simon, *Information in e-motion. Proceedings of Bobcatss 2012, 20th International Conference on Information Science, Amsterdam, 23-25 January 2012* (Bad Honnef: Bock + Herchen, 2012), blz. 82-86.
39. Bussel, Ector, *Op zoek naar de herinnering*, blz. 18-34; P.M. Toebak, *Records Management. Gestaltung und Umsetzung* (Baden: Hier + Jetzt, 2010), blz. 55-134.
40. T. Hulme, 'Unlocking the business value of information', *Business Information Review* 26 (2009), nr. 3, september, blz. 170-181.
41. M. Kooper, R. Maes, E.R. Lindgreen, 'On the governance of information. Introducing a new concept of governance to support the management of information', *International Journal of Information Management* 31 (2011), nr. 3, juni, blz. 195-200.
42. Bussel, Ector, *Op zoek naar de herinnering*, blz. 24-27. Ook: G.J. van Bussel, F.F.M. Ector, P.M.A. Ribbers, G.J. van der Pijl, 'Work- or document flow? The document revolution: documents as trigger for process efficiency', J.P. Heje, C. Ciborra, K. Kautz, e.a. (red.), *Proceedings of the 7th European Conference on Information Systems Copenhagen Business School, 23-25 June 1999* (Kopenhagen: ECIS, 1999), I, blz. 55-69; G.J. van Bussel, F.F.M. Ector, P.M.A. Ribbers, G.J. van der Pijl, 'Building

- the record keeping system (RKS). Process improvement triggered by management of archival documents', J.P. Sprague (red.), *Proceedings of the 34th Hawaii International Conference on System Sciences* (Maui: HICSS, 2001) (CD-ROM).
43. K. Barata, P. Cain, 'Information, not technology, is essential to accountability: electronic records and public-sector financial management', *The Information Society* 17 (2001), blz. 247-258.
 44. M.J. Eppler, *Managing information quality. Increasing the value of information in knowledge-intensive products and processes* (Berlijn-Heidelberg: Springer, 2006³), blz. 17-39; P.T. Groth, *The origin of data. Enabling the determination of provenance in multi-institutional scientific systems through the documentation of processes* (Southampton: University of Southampton, 2007).
 45. Boudrez, Dekeyser, Dumortier, *Digital Archiving*, blz. 75.
 46. S. Spear, H.K. Bowen, 'Decoding the DNA of the Toyota production system', *Harvard Business Review* 77 (1999), nr. 5, blz. 97-106.
 47. J.P. Walsh, G.R. Ungson, 'Organizational memory', *The Academy of Management Review* 16 (1999), nr. 1, blz. 57-91; F. Wijnhoven, 'Organizational memory and information technology. The missing link', J.D. Coelho e.a. (red.), *Proceedings of the Fourth European Conference on Information Systems Lisbon, 2-4 July 1996* (Lissabon: ECIS, 1996), blz. 571-584.
 48. K.E. Weick, *The social psychology of organizing* (New York: McGraw-Hill, 1979²).
 49. Bussel, Ector, *Op zoek naar de herinnering*, blz. 4-10.
 50. Zie de noten 44 en 45. Aangevuld met: Bussel, *Reconstructing the past*.
 51. Bussel, Ector, *Op zoek naar de herinnering*, blz. 172-175.
 52. S.L. Star, K. Ruhleder, 'Steps toward an ecology of infrastructure. Design and access for large information spaces', *Information Systems Research* 7 (1996), nr. 1, blz. 111-133, blz. 113.
 53. De verstoring kan leiden tot het bekend worden van informatie buiten de organisatie: dossiers of bestanden die 'op straat komen'. Als voorbeelden uit vier maanden dit jaar: de medische dossiers van bedrijfsartsen via Humannet (april 2012), de Dienst Vastgoed van het Ministerie van Defensie zette te vernietigen dossiers in containers aan de straat (juni 2012), wachtwoorden (mei 2012) en Outlook-gegevens via onbeveiligde verbindingen bij LinkedIn (juni 2012), gegevens van 84.000 Bol.com-klanten staan op een niet goed beveiligde server (juli 2012), etc.
 54. M. Nyrhinen, *IT Infrastructure. Structure, properties and processes* (Helsinki: Helsinki School of Economics, 2006), Working papers W-403, blz. 2.
 55. B. Bygstad, 'Information Infrastructures as organization. A critical realist view', *Proceedings of the 29th International Conference of Information Systems (ICIS)* (Parijs: ICIS, 2008), blz. 774-786.
 56. P. Weill, M. Broadbent, *Leveraging the new infrastructure. How market leaders capitalize on information technology* (Cambridge (Ms.): Harvard Business School Press 1998).
 57. P.N. Edwards, S.J. Jackson, G.C. Bowker, C.P. Knobel, 'Understanding infrastructure. Dynamics, tensions, and design' [online], geraadpleegd op 31 juli 2012, <http://deepblue.lib.umich.edu/bitstream/2027.42/49353/3/UnderstandingInfrastructure2007.pdf>

58. J. Law, 'Actor-Network theory and material semiotics', B.S. Turner (red.), *The New Blackwell Companion to social theory* (Oxford: Blackwell Publishing, 2008), blz. 141-158 voor een compacte uitleg van deze theorie en een uitgebreid overzicht van wetenschappelijke literatuur.
59. O. Hanseth, K. Lyytinen, 'Design theory for dynamic complexity in information infrastructures: the case of building internet', *JIT. Journal of information technology* 25 (2010), nr. 1, blz. 1-19, blz. 4.
60. J.P. Pironti, 'Key elements of a threat and vulnerability management program', *Information Systems Control Journal* 7 (2006), nr. 3, blz. 52-56, vooral blz. 54-55.
61. Star, Ruhleder, 'Steps toward an ecology of infrastructure', blz. 111-114; Hanseth, Lyytinen, 'Design theory', blz. 2-5; E. Monteiro, O. Hanseth, *Understanding Information Infrastructure* [online], geraadpleegd op 31 juli 2012, <http://heim.ifi.uio.no/~oleha/Publications/bok.pdf>.
62. Star, Ruhleder, 'Steps toward an ecology of infrastructure', blz. 113.
63. Bygstad, 'Information Infrastructures as organization, blz. 780.
64. L. Groth, *Future organizational design. The scope for the IT-based enterprise* (Chichester: John Wiley & Sons, 1999), blz. 349-426.
65. J.R. Beniger, *The Control Revolution. Technological and economic origins of the Information Society* (Cambridge (Ms.): Harvard University Press, 1986), blz. 434-435.
66. Eppler, *Managing information quality*, blz. 17-39.
67. M. Callon, 'Techno-economic networks and irreversibility', J. Law (red.), *A sociology of monsters. Essays on power, technology and domination* (Londen: Routledge, 1991), blz. 132-161.
68. Voor andere zwakke punten: W.B. Arthur, 'Competing technologies, increasing returns and lock-in by historical events', *The Economic Journal* 99 (1989), nr. 394, maart, blz. 106-131.
69. Monteiro, Hanseth, *Understanding Information Infrastructure*, blz. 160-161.
70. Bussel, Ector, *Op zoek naar de herinnering*, blz. 267-272.
71. C.A. Lynch, 'When documents deceive: Trust and provenance as new factors for information retrieval in a tangled web', *Journal of the American Society for Information Science* 52 (2001), nr. 1, januari, blz.12-17; Groth, *The origin of data*.
72. D. Artz, Y. Gil, 'A survey of trust in computer science and the semantic web', *Web Semantics: Science, services and agents on the World Wide Web* 5 (2007), nr. 2, juni, blz. 58-71.
73. K. Kelton, K.R. Fleischmann, W.A. Wallace, 'Trust in digital information', *Journal of the American Society for Information Science and Technology* 59 (2008), nr. 3, blz. 363-374, blz. 366-367.
74. D. Druckman, J.E. Singer, H. Van Cott, *Enhancing organizational performance* (Washington: The National Academies Press, 1997), blz. 39.
75. R.M. Pirsig, *Zen and the art of motorcycle maintenance. An inquiry into values* (New York: William Morrow & Company, 1974), blz. 213.
76. D.A. Garvin, *Managing quality. The strategic and competitive edge* (New York: Free Press, 1988), blz. 43.

77. W.J. Bellows, 'Conformance to specifications, zero defects, and six sigma quality. A closer look', *International Journal of Internet and Enterprise Management* 2 (2004), nr. 1, blz. 82-95.
78. Garvin, *Managing quality*, blz. 46.
79. J. Baudrillard, *Simulations* (Los Angeles: Semiotext(e), 1983), blz. 146.
80. L.P. English, *Improving data warehouse and business information quality. Methods for reducing costs and increasing products* (New York: John Wiley & Sons, 1999), blz. 22-27.
81. Eppler, *Managing information quality*, blz. 17-39; English, *Improving data warehouse and business information quality*, blz. 3-13, voor een uitgebreid overzicht van non-quality-gevallen.
82. Voor een overzicht: R. Wang, J. Funk, Y. Lee, Y., L. Pipino, *Journey to Data Quality* (Cambridge (Ma.): MIT Press, 2009).
83. G.J. van Bussel, *Documentbeheer bij overheidsinstanties. Bouwen van een archiveringsstelsel* (Eindhoven: Euroforum Uitgeverij, 2006⁶).
84. Voor het volgende overzicht zie: Bussel, Ector, *Op zoek naar de herinnering*, blz. 188-209.
85. E. Brynjolfsson, L.M. Hitt, 'Beyond computation. Information technology, organizational transformation and business performance', *Journal of Economic Perspectives* 19 (2000), nr. 4, p. 23-48.
86. S. Greenberg, 'Context as a dynamic construct', *Human Computer Interaction* 16 (2001), nrs. 2-4, blz. 257-268; V. Akman, 'Rethinking context as a social construct', *Journal of Pragmatics* 32 (2000), nr. 6, blz. 743-759.
87. B. Dervin, 'Given a context by any other name. Methodological tools for taming the unruly beast', P. Vakkari, R. Savolainen, B. Dervin (red.), *Information seeking in context: Proceedings of an international conference on information needs, seeking, and use in different contexts* (Londen: Taylor Graham Publishing, 1997), blz. 13-38, blz. 14.
88. H. Clark, T. Carlson, 'Context for comprehension', J. Long, A. Baddeley (red.), *Attention and performance* (Hillsdale (NJ): Lawrence Erlbaum Associates, 1981), IX, blz. 313-330, blz. 313.
89. Zie voor een overzicht van de literatuur: Bussel, Ector, *Op zoek naar de herinnering*, blz. 215-260. Voor een (gedeeltelijk) online overzicht: <http://philpapers.org/browse/the-nature-of-context?sqc=&cn=the-nature-of-context> (geraadpleegd op 31 juli 2012).
90. G. Hirst, 'Context as a spurious concept', A. Gelbukh (red.), *Proceedings of the International Conference CICLing-2000: Conference on Intelligent Text Processing and Computational Linguistics* (Mexico City: IPN, 2000), blz. 273-287.
91. C. Penco, 'Objective and Cognitive Context', P. Bouquet, L. Serafini, P. Brézillon, M. Benerecetti, F. Castellani (red.), *Modeling and Using Context. Proceedings of the Second International and Interdisciplinary Conference, CONTEXT'99, Trento, Italy, September 1999* (Berlijn: Springer-Verlag, 1999), blz. 270-283, blz. 270.
92. P. Dourish, 'What we talk about when we talk about context', *Personal and Ubiquitous Computing* 8 (2004), nr. 1, blz. 19-30, blz. 21-23; Weick, *The social psychology*, blz. 119-232.

93. Respectievelijk: D. Bearman, 'Record Keeping Systems', *Archivaria* 36 (1993), blz. 16-36, blz. 22; D. Bearman, *Electronic Evidence. Strategies for managing records in contemporary organizations* (Pittsburgh: Archives & Museum Informatics, 1994), blz. 43-44.
94. D. Bearman, R.H. Lytle, 'The Power of the Principle of Provenance', *Archivaria* 21 (1985), blz. 14-27, blz. 16. Zie ook: L. Duranti, 'The archival bond', *Archives and Museum Informatics* 11 (1997), nrs. 3-4, blz. 213-218, blz. 217.
95. D. Bearman, 'Archival principles and the electronic office', A. Menne-Haritz (red.), *Information handling in offices and archives* (New York: Saur, 1993), blz. 177-193, blz. 178.
96. R.M. Turner, 'Context-mediated behavior for intelligent agents', *International Journal of Human-Computer studies* 48 (1998), nr. 3, blz. 307-330. Voor een modelmatige benadering van context zie ook: Bussel, Ector, *Op zoek naar de herinnering*, blz. 223-252.
97. C. Brevoord, *De informatiehuishouding. Grondslagen en metamethode* (Deventer: Kluwer, 1991), blz. 79; N.J.T.A. Kramer, J. De Smit, *Systeemdenken. Inleiding tot begrippen en concepten* (Leiden-Antwerpen: Stenfert Kroese, 1982³), blz. 54; K.J. Devlin, 'Oracles in situation semantics', J. Barwise, J.M. Gawron, G. Plotkin, S. Tutiya (red.), *Situation theory and its applications*, II (Stanford (Ca.): CSLI Publications, 1991), blz. 41-49.
98. Groth, *The origin of data*, blz. 17.
99. Zie voor een overzicht van benodigde metadata: Bussel, Ector, *Op zoek naar de herinnering*, blz. 369-407.
100. Bussel, Ector, *Op zoek naar de herinnering*, blz. 356-368. B.E. Bargmeyer, D.W. Gillman, 'Metadata Standards and Metadata Registries: An Overview' (2009), [online], geraadpleegd op 31 juli 2012, <http://www.bls.gov/ore/pdf/st000010.pdf>.
101. R. Groenewald, A. Breytenbach, 'The use of metadata and preservation methods for continuous access to digital data', *The Electronic Library* 29 (2011), nr. 2, blz. 236-248.
102. L. Schamber, M.B. Eisenberg, M.S. Nilan, 'A re-examination of relevance. Toward a dynamic, situational definition', *Information Processing & Management* 26 (1990), nr. 6, blz. 755-776, blz. 755.
103. T. Saracevic, 'Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance', *Journal of the American Society for Information Science and Technology* 58 (2007), nr. 3, blz. 1915-1933, blz. 1917-1918.
104. Saracevic, 'Relevance: A review', blz. 1918.
105. B. Hjørland, F. Sejer Christensen, 'Work tasks and socio-cognitive relevance: a specific example', *Journal of the American Society for Information Science and Technology* 53 (2002), nr. 11, blz. 960-965.
106. A. Abecker, A. Bernardi, K. Hinkelmann, O. Kuhn, M. Sintek, 'Context-aware, proactive delivery of task-specific information: the KnowMore Project', *Information Systems Frontiers* 2 (2000), nrs. 3-4, blz. 253-276.
107. E. Cosijn, P. Ingwersen, 'Dimensions of relevance', *Information Processing and Management* 36 (2000), nr. 4, blz. 533-550; C. Brouard, J.-Y. Nie, 'Relevance as

- resonance: A new theoretical perspective and a practical utilization in information filtering', *Information Processing and Management* 40 (2004), nr. 1, blz. 1-19.
108. B. Hjørland, 'Towards a theory of aboutness, subject, topicality, theme, domain, field, content ... and relevance', *Journal of the American Society for Information Science* 52 (2001), nr. 9, blz. 774-778. De toepassing van het begrip op organisaties is mijn verantwoording.
 109. P. Wilson, 'Situational relevance', *Information Storage and Retrieval* 9 (1973), nr. 8, blz. 457-471. De toepassing van het begrip op organisaties is mijn verantwoording.
 110. R. Glazer, 'Measuring the value of information. The information-intensive organization', *IBM Systems Journal* 32 (1993), nr. 1, blz. 99-110. Voor een overzicht: D.B. Lawrence, *The economic value of information* (New York: Springer, 1993).
 111. H.E. Thiele, 'Appraisal, provenance, and the computer revolution. An examination of organizational records in the electronic age', *University of Pittsburgh Katharine Sharp Review*, 6 (Winter 1998), [online], geraadpleegd op 31 juli 2012, <http://mirrored.ukoln.ac.uk/lis-journals/review/review/6/thiele.pdf>.
 112. R.J. Cox, H.W. Samuels, 'The archivist's first responsibility. A research agenda to improve the identification and retention of records of enduring value', *The American Archivist* 51 (1988), Winter/Spring, blz. 28-42.
 113. E. Shepherd, G. Yeo, *Managing records. A handbook of principles and practice* (Londen: Facet Publishing, 2003), blz. 147-162.
 114. B.L. Craig, *Archival appraisal. Theory and practice* (München: Walter de Gruyter, 2004).
 115. L. Duranti, 'The concept of appraisal and archival theory', *The American Archivist* 57 (1994), Spring, blz. 328-345.
 116. T. Huskamp Peterson, 'Archival principles and records of the new technology', *The American Archivist* 47 (1984), Fall, blz. 383-393, blz. 386.
 117. C. Bailey, 'Archival theory and electronic records', *Archivaria* 29 (1989), Winter, blz. 180-196.
 118. R. Brown, 'Macro-appraisal theory and the context of the public records creator', *Archivaria* 40 (1995), Fall, blz. 121-172.
 119. P. Horsman, 'Appraisal on wooden shoes. The Netherlands PIVOT Project', *Janus* (1997), nr. 2, blz. 35-41.
 120. Bussel, Ector, *Op zoek naar de herinnering*, blz. 308-311.
 121. Archiefwet, artikel 3.
 122. D. Paknad, 'A new framework for defensible disposal', *CIO Update* 27 oktober 2010 [online], geraadpleegd op 31 juli 2012, <http://www.cioupdate.com/insights/article.php/3910271/A-New-Framework-for-Defensible-Disposal.htm>. Zie ook: J. Henseler, *E-Discovery. Op zoek naar de digitale waarheid* (Amsterdam: HvA Publicaties, 2010), blz. 22-26.
 123. J.-F. Blanchette, D.G. Johnson, 'Data Retention and the Panoptic Society. The social benefits of forgetfulness', *The Information Society* 18 (2002), blz. 33-45.
 124. V. Mayer-Schoenberger, 'Useful Void: The Art of Forgetting in the Age of Ubiquitous Computing', Working Paper RWP07-22 (Harvard University, 2007) [online], geraadpleegd op 31 juli 2012, [http://ksgnotes1.harvard.edu/Research/wpaper.nsf/rwp/RWP07-022/\\$File/rwp_07_022_mayer-schoenberger.pdf](http://ksgnotes1.harvard.edu/Research/wpaper.nsf/rwp/RWP07-022/$File/rwp_07_022_mayer-schoenberger.pdf).

125. Mayer-Schoenberger, 'Useful Void', blz. 5-6.
126. H.J.W. van Heerde, *Privacy-aware data management by means of data degradation: making private data less sensitive over time* (Enschede: CTIT, 2010).
127. Nieuwssites wordt regelmatig gevraagd oude artikelen aan te passen of te verwijderen, omdat de informatie verouderd of onwelgevallig is. Het geheugen van kranten is echter belangrijker dan de privacy van personen. Dat is wat de Rechtbank Amsterdam op 19 februari 2010, zaaknummer 448962, rolnummer KG ZA 10-139 WT/CN, bepaalde. Het vonnis in deze zaak is het eerste waarin een rechtbank het belang van integrale archieven afweegt tegen de privacy van betrokkenen. Ik kan mij overigens voor specifieke gevallen wel voorstellen dat het bericht verwijderd wordt uit de index van de zoekmachines, zodat het vinden ervan moeilijker wordt.
128. C. Dollar, *Archival theory and information technologies. The impact of information technologies on archival principles and methods* (Macerata: University of Macerata, 2000); J. Van Bogart, 'Storage Media Life Expectancies' [online], geraadpleegd op 31 juli 2012, <http://nssdc.gsfc.nasa.gov/nost/isoas/dads/presentations/VanBogart/VanBogart-Unabridged.ppt>. Op zich is de levensduur van media niet belangrijk, omdat de informatie zelf leesbaar moet blijven; daarvoor is de aanwezigheid van geschikte software vereist en is hardware nodig, die de gebruikte media toegankelijk kan maken. Ook al blijft het medium 25 jaar in goede staat, dan wil dat niet zeggen dat de informatie die erop staat, kan worden gelezen.
129. Bailey, *Digital Curation Bibliography*. Er zijn ook online overzichten, zoals: 'Preserve bibliography on digital preservation' [online], geraadpleegd op 31 juli 2012, <http://preserv.eprints.org/Preserv-bibliography.html> (bijgehouden tot en met 2009); L. Duranti, R. Preston (red.), 'Bibliography' (2008) [online], geraadpleegd op 31 juli 2012, http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_bibliography.pdf.
130. Boudrez, Dekeyser, Dumortier, *Digital Archiving*, blz. 79-80.
131. D. Bearman, 'Moments of risk: Identifying threats to electronic records', *Archivaria* 62 (2007), blz. 15-46.
132. Boudrez, Dekeyser, Dumortier, *Digital Archiving*, blz. 95-96.
133. Boudrez, Dekeyser, Dumortier, *Digital Archiving*, blz. 80.
134. P. Maniatis, M. Baker, 'Enabling the archival storage of signed documents', D.D.E. Long (red.), *Proceedings of the FAST '02 Conference on File and Storage Technologies, January 28-30, 2002* (Monterey (Ca.): USENIX, 2002), blz. 31-46.
135. M.W. Storer, K.M. Greenan, E.L. Miller, K. Voruganti, 'POTSHARDS: secure long-term storage without encryption', *Proceedings of the 2007 USENIX Annual Technical Conference, June 17-22, 2007* (Santa Clara (Ca.): USENIX, 2007), blz. 142-156.
136. T. Hendley, *Comparison of methods and costs of digital preservation* (Londen: British Library Research and Innovation Centre, 1998), blz. 11-13.
137. In navolging van: K.H. Lee, O. Slattery, R. Lu, X. Tang, V. McCrary, 'The state of the art and practice in digital preservation', *Journal of Research of the National Institute of Standards and Technology* 107 (2002), nr. 1, blz. 93-106.
138. Voor een uitgebreide beschrijving: Bussel, Ector, *Op zoek naar de herinnering*, blz. 317-339.

139. S. Kowalczyk, 'Digital preservation by design', M. Raisinghani (red.), *Handbook of research on Global Information Technology* (Hershey (Pa.): IGI-Global, 2007), blz. 405-431.
140. J. Rothenberg, 'Ensuring the longevity of digital documents', *Scientific American* 272 (1995), nr. 1, blz. 42-47; J. Rothenberg, *Avoiding technological quicksand. Finding a viable technical foundation for digital preservation* (Washington: CLIR, 1999).
141. D. Bearman, 'Reality and chimeras in the preservation of electronic records', *D-Lib Magazine* 5 (1999), nr. 4 [online], geraadpleegd op 31 juli 2012, <http://www.dlib.org/dlib/april99/bearman/04bearman.html>; A. Waugh, R. Wilkinson, B. Hills, J. Dell'oro, 'Preserving digital information forever', *Proceedings of the fifth ACM conference on digital libraries* (DL '00) (New York: ACM, 2000), blz. 175-184; M. Lu, T. Chiueh, *Challenges of long-term digital archiving. A survey* (Stony Brook University (NY), 2006), Technical Report 205; E.H. Dooijes, 'Preface to J. Rothenberg: 'Avoiding technological quicksand. Finding a viable technical foundation for digital preservation' (2000) [online], geraadpleegd op 31 juli 2012, http://www.science.uva.nl/museum/pdfs/Rothenberg_preface.pdf.
142. S. Granger, 'Emulation as a digital preservation strategy', *D-Lib Magazine* 6 (2000), nr. 10 [online], geraadpleegd op 31 juli 2012, <http://www.dlib.org/dlib/october00/granger/10granger.html>.
143. J. van der Hoeven, B. Lohman, R. Verdegem, 'Emulation for digital preservation in practice: The results', *International Journal of Digital Curation* 2 (2007), nr. 2, blz. 123-132.
144. S. Creemers, *Emulatie van videogames. Onderzoek naar het gebruik en de kennis van emulatie bij de Leuvense studenten. Verhandeling aangeboden tot het verkrijgen van de graad van Licentiaat in de Communicatiewetenschappen* (Leuven, 2003), blz. 111-112.
145. K. Russell, 'Digital Preservation and the Cedars Project Experience', *New Review of Academic Librarianship* 6 (2000), blz. 139-154.
146. Hendley, *Comparison*, blz. 18-24.
147. Boudrez, Dekeyser, Dumortier, *Digital Archiving*, blz. 92-96.
148. M. Ferreira, A.A. Baptista, J.C. Ramalho, 'A foundation for automatic digital preservation', *Ariadne* 48 (2006), nr. 48 [online], geraadpleegd op 31 juli 2012, <http://www.ariadne.ac.uk/issue48/ferreira-et-al/>.
149. Lee, 'The state of the art', blz. 98.
150. Waugh, 'Preserving digital information forever', blz. 178.
151. Lee, 'The state of the art', blz. 98.
152. C. Wauters, B. Glashouwer, 'eDepot: van bedrijfsvoering tot digitale rustplaats', *De EDP Auditor* 5 (2009), nr. 1, blz. 16-19; B.F. Cooper, A. Crespo, H. Garcia-Molina, 'Implementing a reliable digital object archive', J.L. Borbinha, T. Baker (red.), *Research and Advanced Technology for Digital Libraries, 4th European Conference, ECDL 2000, Lisbon, Portugal, September 18-20, 2000*. Lecture Notes in Computer Science nr. 1923 (Berlijn: Springer, 2000), blz. 128-143
153. Wauters, 'eDepot', blz. 17.
154. J.-D. Zeller, 'Do we need Cyberarchivists? and what should be their professional profile. Work presented for obtaining a Certificate of continuing education in ar-

- chival sciences' [online], blz. 3/28, geraadpleegd op 31 juli 2012, <http://regardejanus.files.wordpress.com/2011/07/cyarfinalanglr.doc>.
155. Zeller, 'Do we need cyberarchivists', blz. 21-23/28.
156. 'Gartner identifies four information management roles IT Departments need to remain effective' (persbericht) [online], geraadpleegd op 31 juli 2012, <http://www.gartner.com/it/page.jsp?id=1282513>.
157. L. Duranti, 'Educating the eXtreme records professional: a proposal', E. Hokke, T. Laeven (red.), *Archivaris. Professie Professional Professionaliteit Professionalisering* ('s-Gravenhage 2010), blz. 198-207, blz. 203-204.
158. F. Van Kan, 'De archivaris moet emanciperen', E. Hokke, T. Laeven (red.), *Archivaris. Professie Professional Professionaliteit Professionalisering* ('s-Gravenhage 2010), blz. 276-285, blz. 283.
159. Zeller, 'Do we need cyberarchivists', blz. 21-23/28.
160. Duranti, 'Educating the eXtreme records professional', blz. 204.
161. F.C.J. Ketelaar, 'Wat maakt een goede archivaris?', *Archievenblad* 110 (2006), blz. 16-17. Het *records continuum* (of informatiecontinuüm) is het concept dat informatie in de tijd verschillende belangen voor verschillende belangengroepen kan dienen – al dan niet gelijktijdig – en dus betrouwbaar beheer behoeft, ongeacht wie dat uitvoert. Het *respect des fonds* wijst erop dat informatie behoort tot het informatie- of archiefbestand, waarin het bij ontvangst of creatie is opgenomen, en dat het daarin ook dient te blijven berusten.

Curriculum vitae

Geert-Jan van Bussel (1960) voltooide onderwijzers- en lerarenopleidingen, studeerde middeleeuwse geschiedenis en volgde met succes de archiefopleiding. Daarna volgden enkele bedrijfskundige en informatiekundige opleidingen en cursussen. In 2009 promoveerde hij aan de Universiteit van Amsterdam op een proefschrift over organisatorische verantwoordingsystemen en de relaties daarvan met de digitale informatievoorziening en archivering. Vanaf 1993 is hij directeur-eigenaar van Van Bussel Document Services, een adviesbureau voor strategisch informatie- en compliancemanagement. Voor de oprichting van dit bedrijf heeft hij jarenlang gewerkt in het bedrijfsleven en het basis- en voortgezet onderwijs. Vanaf 1994 was hij verbonden aan de Rijksarchiefschool, vanaf 1997 aan de Stichting Archiefschool – Nederlands Instituut voor Archiefonderwijs en -onderzoek. Als zodanig was hij tevens docent aan de Universiteit van Amsterdam en de Hogeschool van Amsterdam. Vanaf 2009 is hij verbonden aan de Hogeschool van Amsterdam. Hij is als gastdocent en *visiting professor* verbonden (geweest) aan verschillende binnenlandse en buitenlandse universiteiten (Erasmus Universiteit Rotterdam, Universiteit van Tilburg, Katholieke Universiteit Leuven, Universiteit van Jyväskylä (Finland)).

Vanaf 1 januari 2012 is Van Bussel bijzonder lector Digital Archiving & Compliance. Hij heeft een groot aantal publicaties over informatie- en archiefmanagement op zijn naam staan. Hij is gespecialiseerd in de ontwikkeling, inrichting, organisatie en implementatie van digitale informatievoorziening, met de nadruk op archivering, privacy, beveiliging en kwaliteit. De afgelopen jaren heeft hij zich ontwikkeld tot een vooraanstaand auditor. Zijn benadering van deze problematiek (een combinatie van bedrijfskundige, bestuurlijk informatiekundige en archiefkundige inzichten) was bij de introductie daarvan in 1993 uniek in de markt. Sindsdien is deze benadering door andere partijen overgenomen. Hij was zowel lid als voorzitter van de Bijzondere Commissie Archieven van de Raad voor Cultuur en is de ‘geestelijk vader’ van de praktijkrichtlijn Bewaren en Bewijzen van het NEN uit 1998. In 2007 heeft hij als hoofdredacteur een beslissende rol gehad bij de herziening van deze richtlijn door EdiForum en NEN. Hij was betrokken bij het samenstellen van het referentiemodel voor RMA-software-specificaties ReMANO 2004. Om bedrijven in de gelegenheid te stellen hun applicaties te toetsen aan dit referentiemodel heeft hij de Stichting Certificatie ReMANO opgericht. Namens deze Stichting heeft hij als ReMANO Auditor het gros van de RMA-applicaties op de Nederlandse markt getoetst. Hij is nu een veelgevraagd auditor voor de NEN-normen 15489 (informatie- en documentenbeheer) en 2082 (archiveringsfunctio-

naliteiten in software) en de ISO-norm 16175 (functional requirements for digital records management systems).

Voor meer informatie, een uitgebreider profiel, publicaties, tools en veel actueel nieuws zie: www.vbds.nl en blog.gjvanbussel.nl. Voor het bijzonder lectoraat zie: www.digitalarchiving.nl.