

# De Nationale Coalitie Digitale Duurzaamheid

Inge Angevaare  
*coördinator*

*121 – 12 mei 2009*

## Wat is de NCDD?

Een 'bottom-up' samenwerkingsverband van een tiental grote nationale organisaties uit de publieke sector die digitale informatie beheren (of bij dat beheer betrokken zijn als financier of adviseur)

rond de gemeenschappelijke uitdaging:

Hoe garanderen we **duurzame toegang** tot de digitale informatie in de publieke sector in Nederland ?

## Ontstaan NCDD

- Negentiger jaren: bewustzijn groeit dat digitaal echt anders is dan gedrukt – bijzondere kwetsbaarheden
- 2000-2003: NDIPP (VS), DPC (UK), nestor (Du)
- 2003: Koninklijke Bibliotheek opent digitaal e-Depot
- 2005: oprichting DANS (archief onderzoeksgegevens  $\alpha$  &  $\gamma$ )
- 2007: Eerste gesprekken 2007 vanuit de Koninklijke Bibliotheek en DANS met andere partners in de publieke sector
- 2007: informele oprichtingsvergadering
- 2008: Stichting NCDD, coördinator, bureau

## Leden van de NCDD

- 3TU Datacentrum
- Centraal Bureau voor de Statistiek - CBS
- Digitaal Erfgoed Nederland - DEN
- Data Archiving and Networked Services - DANS
- Koninklijke Bibliotheek - KB
- Koninklijke Nederlandse Akademie van Wetenschappen - KNAW
- Ministerie van Binnenlandse Zaken
- Nationaal Archief - NA
- Nederlands Instituut voor Beeld & Geluid - NIBG
- Nederlandse Organisatie voor Wetenschappelijk Onderzoek - NWO
- SURFfoundation

## De missie van de NCDD

‘Een **organisatorische** en technische infrastructuur tot stand brengen voor duurzame toegang tot digitale informatie in Nederland.’

## De missie van de NCDD

‘Een **organisatorische** en technische infrastructuur tot stand brengen voor duurzame toegang tot digitale informatie in Nederland.’

Geraamde looptijd: vijf jaar ...

## Filosofie (1)

Techniek van digitale duurzaamheid krijgt al veel aandacht en die komt er ook wel, maar daarnaast is nodig:

- duurzaam beleid
- duurzame organisatie
- duurzame afspraken over rollen en verantwoordelijkheden
- duurzame financiering.

## Filosofie (2)

- Gemeenschappelijk probleem in alle sectoren
- Technisch complex, veel geld en veel expertise nodig

ergo:

- Onderzoeken wat sectoroverschrijdend kan worden aangepakt
- Gezamenlijk plan bij de overheid indienen

Startpunt:

Inventarisatie van de stand van zaken in Nederland:

- Wat voor rol speelt digitale informatie?
- Waar is duurzaamheid verzekerd, waar niet?
- Wie zijn de actoren, wat zijn hun drijfveren?
- Wat weten we van kosten, financiering?
- Wat kun je samen doen, wat moet je alleen doen?
- Hoe zou een infrastructuur eruit kunnen zien?

## ‘Nationale Verkenning Digitale Duurzaamheid’:

### Drie sectoren

- Wetenschap
- Cultureel erfgoed
- Overheid/archieven

### Methodiek:

- Drie onderzoekers uit NCDD partners + projectleiding
- Semi-gestructureerde interviews
- Desk research
- Looptijd zes maanden

Uitkomsten:

Onderzoek loopt nog.

Speciaal voor I2I: wat voorlopige conclusies:

- Algemene trends
- De situatie per sector
- Wat gebeurt er in het buitenland?
- Wat zijn de vooruitzichten voor gezamenlijke aanpak?

## Algemene trends:

- Qua organisatie is digitaal echt anders dan papier/gedrukt
- Papier: lineaire organisatie, duidelijke overdrachtsmomenten, tien jaar onopgemerkt in een kast is meestal geen ramp, en: we weten inmiddels hoe we papier moeten conserveren
- Digitaal: cyclische of zelfs ‘fuzzy’ organisatie, records continuüm, overdrachtsmomenten wel wettelijk geregeld, maar in praktijk veel vragen, zorg van wieg tot graf nodig (‘benign neglect’ werkt niet meer), en we weten eigenlijk nog niet hoe we digitaal materiaal goed kunnen behouden.
- Voor duurzame toegang is invloed op creatieproces nodig
- Enorme schaalvergroting
- Alles is kopie – waar blijft authenticiteit?

## Algemene trends:

- Duurzame toegang is vooral probleem van de publieke sector: bedrijfsleven heeft geen economisch belang bij langetermijnbewaring
- Uitgevers: 'langetermijntoegang is vooral taak van overheid c.q. (nationale) bibliotheken'.
- Farmaceutische industrie moet wel langdurig gegevens bewaren, maar gaf geen nadere informatie voor ons onderzoek
- Bedrijfsleven acteert heel pragmatisch, kijkt nooit langer dan vijf jaar vooruit en vertrouwt op marktwerking voor noodzakelijke (conversie)technieken.
- 'Look and feel' van een document alleen in de publieke sector relevant

## Algemene trends:

- In algemene zin is er nauwelijks iets zinnigs te zeggen over de kosten
- Opslag steeds goedkoper, maar dat is slechts een fractie: het beheer rondom de opslag – en vergeet de energierekening niet
- Kosten worden vooral bepaald door de toegangsregimes – van *deep archiving* (€80 per Tb/jr) tot volledige, snelle online access en doorzoekbaarheid (zomaar een bedrag: €18.000 per Tb/jr).
- Ook bij ingest grote verschillen: diversiteit in bestandsformaten, simpele tot gecompliceerde objecten.



## Trends uit de drie sectoren

## Sector overheid/archieven:

- Waarom bewaren?: rekenschap overheid + cultuurhistorie
- Algemene situatie: rapport 'De dementerende overheid'
- Archiefwet nog te veel gezien als cultuurhistorie, niet als relevant voor het primaire proces van de staat
- Hoewel processen digitaal zijn, is het te bewaren archiefstuk nog erg vaak een printje met natte handtekening
- Rijksoverheid maakt actief beleid ('Informatie op Orde'), lagere overheden wachten vooral af

## Sector overheid/archieven:

### Pluspunten (qua duurzame toegang):

- Archiefwet goed kader als uitgangspunt
- Traditie van selectie voorhanden (digitaal aanpassen)
- Nationaal te besturen c.q. (re)organiseren
- Theorievorming relatief goed ontwikkeld (records continuüm, keteninformatisering)
- Eindverantwoordelijke = tevens duurzame financier (overheid)

## Sector overheid/archieven:

- Programma 'Informatie op Orde' van BZK en OCW
- Nationaal Archief krijgt in 2009 e-depot
- Stadsarchieven Rotterdam/Amsterdam op weg naar e-depot
- LOPAI ontwikkelt eerste normen voor digitaal archief (ED3)
- Tot nu toe twee echt digitale gemeentes: Oss en Deventer.
- Best practice: JustID met CDD+: alle informatie uit de strafrechtsketen beschikbaar via één digitaal loket – terwijl eigenaars toch zelf regie over informatie houden

## Sector wetenschap

- Waarom bewaren?: verificatie van onderzoeksresultaten en hergebruik ten behoeve van nieuw onderzoek
- Digitaal werken is de norm
- Internationaal werken is de norm – maakt organisatie lastig
- Groot verschil tussen publicaties enerzijds en (ruwe) onderzoeksdata anderzijds.
- Publicaties: productie, selectie, distributie, duurzame bewaring naar analogie van gedrukte publicaties goed geregeld (duurzaam archief = Koninklijke Bibliotheek); relatief robuuste bestandsformaten (PDF)
- Ruwe onderzoeksgegevens & software: enorme hoeveelheden, grote diversiteit aan bestandsformaten, data internationaal verspreid over het web, geen selectiemechanismen

## Sector wetenschap

- Grote verschillen per discipline: waar onderzoekers zelf longitudinale data nodig hebben is duurzame bewaring snel op gang gekomen en door onderzoekers zelf georganiseerd
- Elders: onderzoekers hebben vooral belang bij *publicatie* (citaties, prestige), langdurig bewaren van data alleen maar 'extra werk'.
- Waarde van hergebruik wordt vooral gezien door financiers/samenleving, niet door degenen die er het werk voor moeten doen
- Angst bij onderzoekers dat hun data door anderen misbruikt worden.
- Paradox: betaalbare duurzame toegang vraagt om schaalgrootte; wetenschappelijk bedrijf vraagt om heel specifieke datacuratie.

## Sector wetenschap

In Nederland:

- Koninklijke Bibliotheek is depotarchief voor wetenschappelijke publicaties – ook internationale uit de repositories van de universiteiten (DARE) en grote internationale uitgevers
- DANS ingericht door NWO en KNAW om de onderzoeksdata van alfa en gammawetenschappen op te slaan. NWO verplicht een contract met DANS
- Max-Planck-Instituut Nijmegen grootste taalkundige archief
- 3TU.federatie ontwikkelt data-archief voor technische wetenschappen – inclusief ondersteunende diensten die het onderzoeker gemakkelijk moeten maken data *bij creatie* goed te beheren

## Sector wetenschap

- IISG ontwikkelt eigen digitaal depot – data vaak vertrouwelijk.
- SURFfoundation faciliteert ontwikkelingsprojecten.
- Exacte wetenschappen: vooral internationale datacentra waar we nationaal weinig invloed op hebben. ‘Duurzame’ bewaring zeldzaam – NCDD betrokken bij Europese Alliance for Permanent Access to the Records of Science.

## Sector cultureel erfgoed

- Waarom bewaren?: cultuurhistorisch belang
- Meeste aandacht voor fysiek object
- Digitalisering vaak gezien als 'eindstation' van conserveringstraject
- Duurzaam beheer digitale gegevens nog in kinderschoenen – vaak metadata over fysieke collecties.
- Beperkte budgetten nopen tot keuzes: dus veel aandacht voor directe toegang, weinig voor duurzaamheid

## Sector cultureel erfgoed

- NIBG grootste audiovisuele archief, alle programma's publieke omroepen. Duurzame toegang moet nog ontwikkeld worden.
- Een aantal grote digitaliseringsprogramma's en initiatieven: Beelden voor de Toekomst, Metamorfoze, Het Geheugen van Nederland – duurzame toegang via NIBG en KB?
- 'Nederlands Erfgoed: Digitaal!' digitaliseringsproject met tien partners, hoge ambities, maar financiering onzeker
- Digitaal Erfgoed Nederland (DEN) = kenniscentrum en adviescentrum voor erfgoedsector. Duurzaamheid opgenomen in 'De Basis'

## Gesignaleerde behoeftes (1)

- Bewustzijn van het probleem heeft nog steeds veel aandacht.
- Een aantal grote instellingen in Nederland is goed op weg (KB, NA, NIBG, DANS), maar achter de schermen is het werk nog lang niet gedaan – techniek staat nog in de kinderschoenen
- Kleine instellingen zijn zich nog niet echt bewust van het probleem of wachten af.
- Grote instellingen hebben behoefte aan meer en *structurele* financiering i.p.v. projectmatige financiering.

## Gesignaleerde behoeftes (2)

- Kleine instellingen vragen vooral om praktische tools en diensten: Waar moet ik beginnen? Wat zijn de goede bestandsformaten? Wie kan mijn back-up verzorgen? Wie kan mijn data migreren?
- Grote instellingen zijn wel bereid om nationale diensten uit te rollen, maar kleine instellingen stappen niet zo maar in die structuren.
- Overal: roep om kennis, expertise en overzicht over wie wat doet en weet
- Behoefte aan een goed opleidingskader – kennis moet continu bijgehouden worden.



Samenwerken?

In het buitenland diverse sectoroverschrijdende initiatieven:

## VS: National Digital Information Infrastructure and Preservation Program (NDIIPP)

- Opgericht 2000, door het Congres, Library of Congress aangewezen als leider
- Doel 'nationale strategie' – realisatie: her en der intersectorale projecten
- M\$100 budget, maar slechts helft toegewezen
- Momenteel M\$7 budget structureel per jaar
- National Science Foundation krachtige speler in wetenschap met Datanet.

## UK: Digital Preservation Coalition (DPC)

- Oprichting 2002
- Zelfde opbouw als NCDD, inmiddels meer leden
- 2006: invloedrijk 'Mind the Gap' rapport, eerste echte beschrijving van het probleem
- na 2006: 'fill the gap' – nog niet goed van de grond gekomen
- Wel actief in kennisuitwisseling – What's New; DP Handbook
- UK invloedrijk in R&D en beleidsvorming: Digital Curation Centre, JISC, Research Information Network

Duitsland: nestor

- Oprichting 2003
- ‘Kompetenznetzwerk’ gericht op kennisuitwisseling; gedegen organisatie d.m.v. ‘Arbeitsgruppen’
- Handboek digitale duurzaamheid, ‘Kriterienkatalog’
- Project loopt af in 2009, wordt voortgezet door leden

- Australië en Nieuw Zeeland zeer actief in R&D, relatief gemakkelijke organisatie
- Frankrijk erg op zichzelf – centralistisch aangestuurd.
- Denemarken – pogingen tot nationale coalitie mislukt.
- Veel EU R&D projecten (Planets, Shaman, Caspar, DPE, PARSE, enz.) – maar er zijn vragen de bij duurzaamheid van de opgeleverde resultaten
- Alliance for Permanent Access = Europese coalitie voor wetenschappelijke informatie

## Vooruitzichten voor samenwerking in Nederland

- Geen kant en klaar plaatje
- Ook ervaringen in het buitenland: een gemeenschappelijk probleem leidt niet zomaar tot gezamenlijke oplossingen
- Heel veel afko's en organisaties. Wie doet wat nu precies? Overzicht ontbreekt.
- De materie loopt dwars door bestaande structuren en organisatievormen heen
- Samenwerking wordt bemoeilijkt doordat de instellingen ook concurrenten zijn: ze strijden om de middelen en moeten zichzelf profileren

## Vooruitzichten voor samenwerking in Nederland

- Belangrijkste slag die eerst gemaakt moet worden ligt *binnen* de sectoren: archiveren begint bij creatie
- Selectie, selectie, selectie!
- Uiterste' scenario's lijken weinig kansrijk (één datacentrum voor Nederland versus iedere organisatie zijn eigen e-depot)
- Waarschijnlijker: netwerk van grotere e-depots die datadiensten leveren voor kleine organisaties
- Echter: die grotere organisaties hebben nog hun handen vol aan hun eigen ontwikkelingen (met name beperkte menskracht en expertise)

## Kansrijke samenwerking

- Bewustzijn kweken van langetermijnbelang – PR campagnes
- Delen van expertise, ontwikkelen van nieuwe technieken en vooral procedures, opleidingskader
- Leren van elkaar: wetenschap is ver in ontwikkelen van persistent identifiers; archivistiek is ver in theorievorming en selectie, enz.
- Afstemmen collectieprofielen, dienstenpakketten
- Gezamenlijk blijven werken aan een landelijk dekkend netwerk van duurzame toegang

## Randvoorwaarden

- Lange adem nodig – meer dan vijf jaar
- Vertrouwen nodig om ook de missers met elkaar te delen – oude rivaliteiten moeten worden afgebroken, dat kost veel tijd.
- Een bindend orgaan dat de partijen steeds weer om de tafel zet (NCDD?)
- Bereidheid van overheid om mede te investeren in infrastructuur – geld en mensen.

## Slotopmerking

- Jeff Rothenberg (1995): we hebben een technisch probleem met digitale informatie
- David Rosenthal (2009): technisch flinke vooruitgang geboekt, maar met name de *schaalgrootte* is een probleem dat we nog niet onder de knie hebben.

Schaalgrootte en gebrek aan lineaire structuren dé uitdagingen voor de menselijke geest

Meer informatie:

[www.ncdd.nl](http://www.ncdd.nl)

<http://digitaalduurzaam.blogspot.com>

[inge.angevaare@kb.nl](mailto:inge.angevaare@kb.nl)